Three Essays in Empirical Asset Pricing

by

Stephen Szaura
Finance Area
Desautels Faculty of Management
McGill University

A thesis submitted to McGill University in partial fulfillment
of the requirements of the degree of Doctor of Philosophy

Montreal, Quebec, Canada

Abstract

This thesis comprises three essays in empirical asset pricing. My first essay entitled "Are stock and corporate bond markets integrated? A Big Data Approach" I document the existence a growing Factor Zoo of discovered characteristics and factors that predict the cross-section of corporate bond returns and generate a significant high minus low portfolio alpha. I determine a higher statistical benchmark, by accounting for those characteristics and factors that have been discovered in published and working papers and find that in cross-sectional regressions and portfolio sorts of over a hundred characteristics and factors, on average 2.4% predict the cross-section of corporate bond returns when adjusting for higher benchmarks. A multivariate horse-race of all characteristics and factors in cross-sectional regressions finds a higher number of corporate bond, rather than stock, characteristics and factors that predict the cross-section of corporate bond returns when adjusting for higher benchmarks. In addition to the lower number of corporate bond characteristics and factors that predict the cross-section of stock returns, my results show that the stock and corporate bond markets are more segmented than previously documented.

My second essay is based on a joint working paper entitled "Do Option Implied Measures of Stock Mispricing Find Investment Opportunities or Market Frictions" where we find that existing option implied stock mis-pricing measures, the portfolios identified as being the most mispriced (highest quintile), typically have the highest shorting fee. When those stocks are omitted, the average abnormal returns of the long-short stock portfolios are insignificant or greatly reduced in economic magnitude. We propose a new measure, IPD, using a novel intra-day options trades data set, circumvents this and does not require shorting hard to borrow firms.

My third essay is based on a joint working paper entitled "Accounting Transparency and the Implied Volatility Skew". We show theoretically and empirically that firms with higher accounting transparency have an implied volatility smirk that is more sensitive to leverage (vice versa). The more clear the accounting information the more skewed the implied volatility smirk. Our theoretical predictions rely on extending the Duffie and Lando [2001] credit risk model to stock option pricing whereby incomplete accounting information and the risk of bankruptcy together act as an economic source of jump risk for stocks. Empirical tests confirm the theoretical predictions of the model and the model can be solved in closed form solution up to Bivariate Standard Normal Cumulative Distribution Function.

Resume

Cette thèse consiste de trois essaies de finance empirique. Dans la premiére essaie, je documente l'éxistence d'un Factor Zoo de caractéristiques et de facteurs qui prédisent la coupe transversale des obligations d'entreprise et génère une long-short alpha portefeuille significatif. Je détermine un point de référence statistique plus élevé, en tenant compte de tous les caractéristiques et facteurs qui ont été découverts dans les journeau academique de finance. Dans les régressions transversales et les exercises portefeuille, en utilisant plus d'une centaine de caractéristiques et de facteurs, en moyenne 2.4% de ces caractéristiques et de facteurs prédisent la coupe transversale des obligations d'entreprise en utilisant une ajustement statistique pour tenir compte de tous les caractéristiques et de facteurs que je teste. Une horse-race de plus qu centaine de caractéristiques et facteurs dans les régressions transversales pour prédire la coup transversale des obligations d'entreprise, révèle que la plupart qui sont significatif vient de la donnée des d'obligations d'entreprise et pas la donnée des d'actions, quand utilisant une ajustement statistique pour tenir compte de tous les caractéristiques et de facteurs que je teste. En plus, les caractéristiques des obligations d'entreprise n'aide pas a prédire la coupe transversale des actions, alors mes résultats montre que les marchés des actions et des obligations d'entreprise sont plus segmentés que ce qui avait été précédemment documenté.

Dans mon deuxieme essaie, nous découvrons que les mesures existante de tarification erronée des actions (actions mal évalués) qui sont construit en utilisant d'infomation des dérivés d'actions, identifiés les actions qui sont difficile a faire du vent de flash et pas les actions qui sont mal évalués. Lorsqu'on omises les actions qui sont le plus cher de faire le vent de flash, les alphas des long-short portefeuilles d'actions sont insignifiants ou considérablement réduits. Nous proposons une nouvelle mesure, IPD, qui utilise un nouvel ensemble de données des transactions de dérivés, qui fonctionne bien pour prédire les retours des actions quand on omises les actions qui sont le plus cher de faire le vent de flash.

Dans mon troisième essaie, nous montrons théoriquement et empiriquement que les entreprises avec une transparence comptable plus élevée ont un sourire de volatilité implicite qui est plus sensible à l'effet de levier (vice versa). Plus les informations comptables sont claires, plus le sourire de volatilité implicite est biaisé. Nos résultats théoriques reposent sur l'extension du modèle de risque de crédit Duffie and Lando [2001] à la tarification des options d'achat d'actions, dans laquelle des informations comptables incomplètes et le risque de faillite agissent ensemble comme une source économique de risque de saut pour les actions.

Contribution of Authors

My first essay entitled "Are stock and corporate bond markets integrated? A Big Data Approach" I am the sole author and contributor. My second essay is based on a joint working paper entitled "Do Option Implied Measures of Stock Mispricing Find Investment Opportunities or Market Frictions" coauthored with my advisor Ruslan Goyenko (McGill University) and co-authors Martijn Cremers and Paul Schultz from Mendoza College of Business, University of Notre Dame. Ruslan, Martijn, Paul, and myself all contributed equally to the paper. My third essay is based on a joint working paper entitled "Accounting Transparency and the Implied Volatility Skew" coauthored with my committee member Jan Ericsson (McGill University) and co-authors Hitesh Doshi (Bauer School of Business, University of Houston) and Fan Yu (Claremont McKenna College). Jan, Hitesh, Fan, and myself all contributed equally to the paper.

Acknowledgements

First I would like to thank my thesis advisor, Professor Ruslan Goyenko, for being an exceptional advisor and mentor. It was a pleasure to complete this dissertation under his guidance and supervision. I would also like to thank professors Jan Ericsson, Guillaume Roussellet, and Professor Christian Dorion (HEC Montreal University) for accepting to be part of my thesis committee. I would also like to thank professors Pascal Francois (HEC Montreal University) and Andriy Shkilko (Wilfred Laurier University) for valuable feedback comments in the written examination reports.

Next I would like to take the time to thank my co-authors and collaborators. My second essay is based on a joint working paper with my advisor Ruslan Goyenko and co-authors Martijn Cremers and Paul Schultz from Mendoza College of Business, University of Notre Dame. I was extremely fortunate to be able to work with such amazing and dedicated co-authors as well as being able to spend an academic visit under Paul Schultz at the University of Notre Dame. I am beyond grateful for Paul's assistance in completing the paper as well as agreeing to be a letter reference writer when I went on the finance job market. My third essay is based on a joint working paper with my committee member Jan Ericsson and co-authors Hitesh Doshi (University of Houston) and Fan Yu (Claremont McKenna College). I thank them for their help in completing the paper. Additionally I am gratefull for Paul Schultz recommending and setting me up to visit and work with Rene Stulz at Fischer College of Management, Ohio State University. I really appreciate the time I was able to spend discussing my research with Rene Stulz as well as his feedback on my job market paper.

I would like to take the time to thank several of the McGill Finance, McGill Economics, HEC Montreal University Department of Finance, and University of Montreal Economics department for offering and teaching some outstanding courses that I greatly benefited from. In particular I would like to thank Patrick Augustin, Benjamin Croituro, Adolfo De Motta, Jan Ericsson, Jiro Kondo, and Sergei Sarkissian all from the McGill Finance area; as well as Saraswata Chaudhuri and Victoria Zinde-Walsh from the McGill Economics department and Rene Garcia from the University of Montreal Economics department.

I am extraordinarily grateful for the financial support that was provided by: National Bank Financial Group, the Montreal Derivatives Exchange (through the TMX Canadian Derivatives Scholarship), the Society of Actuaries (through the Hickman Doctoral Fellowship and renewals) and Desautels Faculty of Management for my PhD student stipend but also two McGill Travel Grants that allowed for me research visits to The Mendozza School of Business at The University of Notre Dame and Fischer College of Management at Ohio

State University. Without all of this financial support this thesis would not have been possible.

I would also like to thank many of my friends and classmates for their support and help over the years of the PhD. In particular I would like to thank: Ghazaleh Aghili, Amir Akbari, Miron Avidan, Linxiao (Francis) Cong, Fillipo D'Allio, John Fast, Feng Jiao, Evan Jo, June Lee, Yiliu (Lucie) Liu, Rojin Mansouri, Diego Mastroianni, Wadih Reno, and Chengyu Zhang. I would also like to thank my friend Yoontae Jeon for many helpful discussions.

Additional thanks go to the Desautels IT department for outstanding computer and technical support. In particular I would like to thank Rohan Bailey, Pierre Cambron, and David Ilacqua for their outstanding support in maintaining the computer system in the Desautels Faculty of Management. They were very prompt, always helpful and friendly with any computer support I required while working remotely. Truly very valuable support. I would also like to thank the McGill High-Power Computing (HPC) department and support team, as well as those from Compute Canada and Calcul Quebec, for using their cloud computing and SSH shell servers. Their software infrastructure made analyzing many of my high-frequency data sets much easier and faster. Lastly, I would also like to thank the SAS technical support team for the prompt and very helpful replies with any questions I had about different SAS functions and SAS installations.

Additionally, I would like to thank Peter Christoffersen for introducing me to empirical asset pricing as well as to my future advisor at McGill. I was deeply saddened to learn of his passing and that he could not have read this dissertation. I'm sure he would be proud and have made outstanding comments about it.

Finally, I would also like to thank my parents Sonia and Vladimir Szaura and my brother Michael Szaura for their constant loving support and encouragement!

Dedication

This thesis is dedicated to my parents Sonia and Vladimir Szaura and my brother Michael Szaura for their constant loving support and encouragement!

Table of Contents

List of Tables

List of Figures

Chapter 1

Thesis Roadmap

My thesis comprises three essays based on three separate working papers each of which is written as a separate chapter in this dissertation. Each of my three essays is rooted in the empirical asset pricing literature in finance. The common theme of all three essays is to analyze the relationship and interaction between two related financial markets.

Chapter 2 of this dissertation is based on my job market paper Szaura [2020]. A literature has emerged in finance that demands a higher hurdle for whether characteristics and factors predict the cross-section of stock returns. I observe a growing Factor Zoo of discovered characteristics and factors that predict the cross-section of corporate bond returns and generate a significant high minus low portfolio alpha. Once determining a higher benchmark, by accounting for those characteristics and factors that have been discovered, many are no longer significant. In cross-sectional regressions and portfolio sorts of over a hundred characteristics and factors, on average 2.4% predict the cross-section of corporate bond returns when adjusting for higher benchmarks. A horse-race of all characteristics and factors in cross-sectional regressions finds a higher number of corporate bond, rather than stock, characteristics and factors that predict the cross-section of corporate bond returns when adjusting for higher benchmarks. In addition to the lower number of corporate bond characteristics and factors that predict the cross-section of stock returns, my results suggest that the stock and corporate bond markets are more segmented than previously documented.

Chapter 3 of this dissertation is based on the working paper Cremers et al. [2019]. Many measures of option implied stock mis-pricing have been proposed in the academic literature. In this paper we ask: *Do these option implied stock mis-pricing (informed trading (private information)) measures capture investment opportunities or simply market frictions?*

Answer: Most existing measures capture market frictions, however, we do propose a new measure (IPD) using a novel data set of options trades in which abnormal trading profits are driven by taking a long position in larger firms where there is more option trading and does not require shorting hard to borrow firms. We find that the average abnormal returns that existing option implied stock mis-pricing measures require the investor to short a portfolio of stocks that are hard to borrow (highest quintile of shorting fee). When sorting stocks into quintile portfolios by existing option implied stock mis-pricing measures, the portfolios identified as being the most mis-priced (highest quintile) have the highest shorting fee. When those stocks that have the highest quintile of shorting fees are omitted, the average abnormal returns of the long-short stock portfolios of most existing measures are insignificant or greatly reduced in economic magnitude, however, not for IPD.

Chapter 4 of this dissertation is based on the working paper Doshi et al. [2020]. In this paper we ask: *How does the quality of firm accounting information change the impact of firm leverage on the implied volatility skew?* We show theoretically and empirically that firms with higher accounting transparency have an implied volatility smirk that is more sensitive to leverage (vice versa). The more clear the accounting information the more skewed the implied volatility smirk. Our theoretical predictions rely on extending the Duffie and Lando [2001] credit risk model to stock option pricing whereby incomplete accounting information and the risk of bankruptcy together act as an economic source of jump risk for stocks. Our model can be solved in closed form up to Bivariate Normal Cumulative Distribution Function.

Chapter 5 concludes with future avenues of research for each of the chapters of research.

Chapter 2

Are stock and corporate bond markets integrated? A Big Data Approach

2.1 Introduction

Recently the cross-section of U.S. corporate bond returns has seen a plethora of signals discovered that predict the cross-section of corporate bond returns.[1] We know that corporate bonds are linked to the firm stock through the structure of the firm and will share variation in stock risk premia. Through this channel, signals which are significant in predicting future stock returns, i.e. firm stock risk premium, can be informative about corporate bond risk premium. Studying the corporate bond return premium, however, is a complex task compared to the stock premium. Some of the reasons for the challenging nature due to the fact that corporate bonds trade in a decentralized over-the-counter dealer market (whereas stocks trade in a liquid centralized limit-order book market), multiple underlying bonds for the same firm, different times to maturity, different credit ratings, and different clientele in that bond market investors are largely institutional investors who have long-term buy-and-hold investment strategies, and other reasons. Despite the challenges, the

[1] Despite the fact that characteristics are not risk factors, if firm characteristics are correlated with the cross-section of corporate bond returns, a long-short portfolio can typically be formed in order to proxy for the unknown risk factor and hence in this form the characteristic can be considered as a risk factor. Hence we refer to both characteristics and factors that predict the cross-section as signals for simplicity for the remainder of the paper.

cross-section of corporate bond returns has had signals discovered from many different risk categories.[2]

Recently the empirical cross-sectional analysis of returns literature has been criticized for not adjusting their testing methods to incorporate all the signals that have already been discovered in published and working papers in the literature (see Harvey et al. [2016]) as well as selective reporting of only those signals that are statistically significant (see Green et al. [2017] and Chordia et al. [2020]).[3] Motivated by the increase in the number of recently discovered signals that predict the cross-section of corporate bond returns, as well as the recent criticisms in the cross-sectional empirical analysis of returns in this paper I conduct a comprehensive analysis of the integration between the risk premia of the corporate bond and stock markets by asking: (1) Do signals that originate from the corporate bond literature perform better than those that originate from the stock return literature in predicting the cross-section of corporate bond returns? (2) Do signals that originate from the corporate bond literature predict the cross-section of stock returns?

To assess whether signals that originate from the corporate bond literature perform better than those that originate from the stock return literature in predicting the cross-section of corporate bond returns, I first analyze which set of signals have lower false rejections in published papers to see historically which signals are better identified to predict the cross-section of returns.[4] I then test which are significant signals in classical empirical asset pricing tests (multivariate cross-section regressions and double sorts) adjusting the statistical significance level for the number of signals simultaneously being tested and using a set

[2]For example, in terms of the market microstructure risk class, bondholders earn higher returns for more illiquid bonds and illiquid stocks (Lin et al. [2011]), for bonds that have higher financial intermediary risk exposure and bond supply both positively predicts future bond returns (He et al. [2017] and Goldberg and Nozawa [2019] respectively). Another example, from return based signals, would be bonds that have a higher exposure to systematic risk bond market return, or credit risk, or liquidity risk, and downside risk, all earn higher bond returns (Bai et al. [2019b]). I present a comprehensive study of the signals that have been already been discovered to predict the cross-section of corporate bond returns in Section 2.4.1.

[3]Harvey et al. [2016] advocate for adjusting the level of significance in empirical cross-sectional asset pricing tests based on the number of signals that have already been discovered in published and working papers in the literature. This is due to the fact that the same null hypothesis, that a signal does not predict the cross-section of returns, has been tested many times in those published and working papers in the literature. Harvey et al. [2016] show that using multiple hypothesis testing techniques from statistics is necessary in order to reduce the probability of getting at least one signal to be statistically significant result due to random change increases as the number of signals tested in predicting the cross section of returns increases.

[4]I analyze whether corporate bond or stocks signals that have been already been discovered to predict the cross-section of corporate bond returns have lower false discovery rates (by applying the empirical simulation method from Harvey et al. [2016]) once adjusting the significance level for all the of discovered signals.

of over a hundred signals which are an unbiased of selective reporting as well as representative of different risk categories (stock illiquidity, stock risk, equity options, balance sheet variables, and corporate bond based signals) and signals that have already been discovered in published and working papers in the literature. Adjusting the significance level for the number of signals tested, fewer signals will be statistically significant (as well as including all signals in multivariate regressions) specifically those signals that are weakly identified (having a lower t-statistic) in predicting the cross-section of corporate bond returns. Once adjusting the significance level for the number of signals tested in set of over a hundred signals which are an unbiased of selective reporting, the number of signals that originate from the corporate bond literature compared to those that originate from the stock return literature in predicting the cross-section of corporate bond returns is how I measure how well integrated the the risk premia of the corporate bond and stock markets are.

When considering a higher threshold of discovered signals in published papers or empirical asset pricing tests, the lower false discoveries in the corporate bond based signals over stock based signals suggest that corporate bond signals, over stock signals, better predict the cross-section of corporate bond returns. A horse-race of all signals in a multivariate cross-sectional regression to predict the cross-section of corporate bond returns finds a higher frequency of corporate bond, rather than stock, signals that drive return predictability of the cross-section of corporate bond returns. Results are driven by bond return skewness, bond age, and bond return standard deviation. In a second horse-race, using all signals to predict the cross-section of stock returns finds none of the corporate bond signals drive return predictability of the cross-section of stock returns when considering higher benchmarks.

In order to measure the ability of combinations of signals in predicting corporate bond returns (and stock returns), I sort corporate bond returns based on pairs of the all signals into conditional double sorted quintile portfolios four corner high minus low alphas. When imposing higher thresholds to the over $40,000$ high-minus-low portfolios, 0.13% of double sort combinations of signals are statistically significant in predicting the cross-section of corporate bond returns. Of those double sorted portfolios who exceed the multiple hypothesis testing benchmark, over 60% of them involve sorting on a corporate bond signal. Similarly, in order to measure the ability of combinations of signals in predicting stock returns (for firms with corporate bonds outstanding), I sort stock returns based on pairs of the all signals into conditional double sorted quintile portfolios four corner high minus low alphas. When imposing higher thresholds to the over $40,000$ high-minus-low portfolios, only one of double sort combinations of signals are true discoveries in predicting the cross-section of stock returns, and does not use a corporate bond signal. Both sets of horse-racing results are robust to different factor specifications, value-weighting by firm size (not to overweight micro-cap firm corporate bond returns), and different multiple hypothesis testing methods.

The empirical evidence shows that corporate bond signals drive the predictability in the cross-section of corporate bond returns and that none of the corporate bond signals, in either multivariate regressions and double sorts, drive return predictability of the cross-section of stock returns. These findings are in line with the lower false rejections of corporate bond signals in historical tests that predict the cross-section of corporate bond returns and show that the risk premia of the corporate bond and stock markets are not well integrated. Numerous published papers claim that various signals from the corporate bond and stock return predictability literature play an important role for understanding the risk premia of the corporate bond and stock markets, however, once accounting for biases in selective reporting as well as the number of tests of predictability that have likely been tried only a few corporate bond signals remain important in predicting the corporate bond returns and they do not help in predicting the cross section of stock returns.[5]

The results of my paper have direct implications for the trading and portfolio management strategies of institutional investors whose assets under management comprise large amounts of corporate bonds such as: fixed income hedge funds, pension funds, insurance companies, and mutual funds. Classic textbook portfolio trading and portfolio management strategies teach funds to make use of different signals to earn higher returns for their portfolios.[6] Empirical evidence in Asness et al. [2013] and Moskowitz et al. [2013] show that the same signal, or variations of them, have been found to exist in asset classes beyond U.S stocks, such as currencies, bond futures, options, and across markets. Trading on signals that are not significant can have adverse and undesirable consequences for the portfolio and fund returns.

The remainder of this paper is organized as follows: Section 2.2 outlines the related literature, Section 2.3 provide the details of the data sets used, Section 2.4 presents my main empirical results, Section 2.5 details my robustness tests performed and Section 2.6 concludes.

[5]See Lin et al. [2011], Chordia et al. [2017], Choi and Kim [2018], Avramov et al. [2019], Chung et al. [2019], Bali et al. [2020], and Sandulescu [2020] empirically study various well known signals that predict the cross-section of stock returns to see if they predict the cross-section of corporate bonds.

[6]See for example Chapter 7 of Ang [2014] or Section 3.4 of Pedersen [2015].

2.2 Literature Review and Theoretical Predictions

Amongst the first papers to study corporate bond returns, Fama and French [1993], create two bond market factors: a default risk factor and a term factor.[7] The construction of the bond market factors is based on corporate bond index returns and not individual bond transaction prices.[8] The Bai et al. [2019b] four factor model was constructed using individual bond transaction prices in order to capture both the systematic and idiosyncratic variation of the individual corporate bond trades as opposed to using bond indices where returns are already pooled across quote and trade prices, maturity, and credit rating.

Theoretically the classical structural credit risk model of Merton [1974] is one of the simplest frameworks to analyze the joint stock and credit risky bond model of the firm, with a integrated risk premia of stocks and corporate bonds, under the rigid assumptions such as the investors have common information set, no transaction costs, etc. Many of the frictionless market assumptions in the model of Merton [1974] are violated in reality: corporate bond trading costs are higher than stocks, there are differences in the trading structure of equities and corporate bonds since equities trade in centralized exchange whereas corporate bonds trade in a relationship decentralized structured over-the-counter market, corporate bonds are held for long term investment horizon typically by insurance companies and pension plans to match expected cash flows whereas stocks held for shorter term, also there are differences in investor information sets of the investors of ewuities and corporate bonds since corporate bond investors are sophisticated investors (”smart style” trading) and stock investors are typically unsophisticated. When decomposing stock and corporate bond each into cash flow and discount rate risk components, as in the classic Campbell and Schiller [1988] decomposition, Chen et al. [2013] show that stocks are more sensitive to cash flow risk. Lochstoer and Tetlock [2020] find that signals that drive stock return predictability is largely driven by the cash-flow risk channel and hence cash-flow risk predictors. Nozawa [2017] shows that corporate bond credit spreads are largely driven by changes in discount rate risk and that discount rate shocks are well explained by corporate bond signals such as credit rating, changes in interest rates, and duration. Based on the empirical evidence that corporate bond returns are less driven by cash flow shocks, I would expect that stock signals are less important in predicting the cross-section of corporate bond returns.

My paper contributes to three strands of academic literature. First I contribute to

[7]See Keim and Stambaugh [1986] and Fama and French [1989] for initial studies that test what signals predict corporate bond index returns.

[8]See also Elton et al. [1995], Gebhardt et al. [2005a], and Gebhardt et al. [2005b] that use corporate bond index returns and not individual bonds.

the literature where signals that predict the cross-section of stock returns are tested as to whether or not they predict the cross-section of corporate bond returns and vice versa (cross market integration of stock and corporate bond risk premia). Lin et al. [2011], Avramov et al. [2019], Chung et al. [2019], empirically study whether illiquidity, momentum, volatility (constructed from corporate bond and stock returns) to see if they are important to predict the cross-section of stock and corporate bond returns as well as Chordia et al. [2017], Choi and Kim [2018], and Cao et al. [2020], show that some of the well known stock signals (leverage, book-to-market ratio, stock momentum, equity option implied volatility, etc.) that have been documented to predict the cross-section of stock returns can predict the cross-section of corporate bonds.[9]

In a related paper, using a large set of equity and bond signals, Bali et al. [2020] find that there is no difference in the predicted corporate bond return spread of machine learning methods when adding stock signals to the corporate bond signals. My paper complements their recent findings as I find that more corporate bond signals, as oppose to stock signals, predict future corporate bond returns when adjusting for multiple hypothesis testing methods. Hence the results of Bali et al. [2020] and this paper both find, using different methods, that stock signal predictive power is economically insignificant whereas corporate bond signals are important in predicting corporate bond returns. In a similar vein, I find that corporate bond signals do not predict future stock returns when adjusting for multiple hypothesis testing. The result is consistent with the additional findings of Bali et al. [2020] that corporate bond signals do not provide any incremental predictive power beyond equity signals in predicting stock returns.

The main difference is in the set of signals that drive the predictability. I find the corporate bond return predictability to be largely driven by bond age, bond time to maturity, bond return skewness, bond return volatility, and amount outstanding whereas the ML techniques used in Bali et al. [2020] find the corporate bond return predictability to be driven by bond Beta uncertainty (UNC), bond market beta, bond return downside risk ($VaR5$ and $VaR10$). Also I find that conditional double sorts that exceed multiple hypothesis testing higher t-statistic thresholds, are driven by conditionally sorting on corporate bond age, which highlights the importance of the corporate bond age as a signal, whereas Bali et al. [2020] find corporate bond age to be of relatively low importance compared to all other signals used. I evaluate the abnormal high-low portfolio alpha (with respect to the Bai et al. [2019b] four factor model) of a variety of machine learning methods in predicting the cross-section of corporate bond returns when using multiple hypothesis testing higher

[9] In a related but different adaptation of stock and corporate bond market analysis, Campello et al. [2008], Elkamhi and Ericsson [2008], and Zundert and Driessen [2017] use implied stock prices from structural models estimated from corporate bond yields data to see if they predict stock returns.

t-statistic thresholds.[10]

This leads to the second strand of literature that I contribute to: the understanding which signals predict the cross-section of corporate bond returns. Thanks to advances in computing power and data storage, economists can now test signals faster than ever before, from more data sets (publicly and privately available) and use a 50 year history Factor Zoo of signals from different risk classes.[11] Many signals from different risk classes have already been discovered as predictors of the cross-section of corporate bond returns. There are, for example: bond return signals (bond return momentum,Jostova et al. [2013], bond short and long term reversal, Bai et al. [2019a], bond and stock volatility, Chung et al. [2019]), bond and stock microstuctural (illiquidity, Lin et al. [2011], financial intermediary bond risk exposure He et al. [2017], bond supply Goldberg and Nozawa [2019]), balance sheet (Chordia et al. [2017], Choi and Kim [2018], and Chichernea et al. [2019]), equity option (Cao et al. [2020]), and others.[12] I add to this list several signals, that are well known to predict the cross-section of stock returns which are statistically significant and have not been discovered to predict the cross-section of corporate bond returns. I find the change in stock momentum (Gettleman and Marks [2006]) is a positive predictor of future corporate bond returns and the dispersion of analyst beliefs (Diether et al. [2002]) is a negative predictor of future corporate bond returns.

Lastly, I contribute to a growing literature in finance which uses multiple hypothesis testing and accounting for data-snooping methods in the evaluation of signals in the cross-section of asset returns.[13] In the determinants of the cross-section of corporate bond return literature, my paper is the first to apply multiple hypothesis testing methods in evaluating whether signals are statistically significant or not both to the t-statistics of those signals that have already been discovered and in empirical tests of additional signals in the cross section of corporate bond returns.[14]

[10] I provide a description of each of the machine learning methods used in my empirical analysis in Appendix Section A.1.5 and a description of the performance attribution of the machine learning methods in Appendix Section A.1.4.

[11] For a running list of the over 300 signals that have been discovered in predicting the cross-section of stock returns see Harvey and Liu [2019].

[12] I provide a comprehensive list of all signals discovered in the cross section of corporate bond returns in publications or working papers see Section 2.4.1, Table A.1.5.

[13] For applications of multiple hypothesis testing in finance see: Shanken [1990] Ferson and Harvey [1991] Boudoukh et al. [2007] and Patton and Timmermann [2010], Harvey et al. [2016], Yan and Zheng [2017], Green et al. [2017], Chordia et al. [2020], and Harvey and Liu [2020]. For applications of data-snooping methods in finance see: Lo and MacKinlay [1990], Foster et al. [1997], Sullivan et al. [2001], Conrad et al. [2003], Cooper and Gulen [2006], and Lynch and Vital Ahuja [2012].

[14] For applications of multiple hypothesis testing in other areas of finance see: Barras et al. [2010], An-

2.3 Data

2.3.1 Corporate Bond Data

For the corporate bond data I use the transaction records that are reported in the TRACE and Enhanced TRACE reporting system for the sample period July 1, 2002 to October 31, 2019. The TRACE data set offers the best quality of corporate bond transactions with intra-day trades information which include price, trading volume, buy/sell indicators. I use the TRACE (and Enhanced TRACE) data cleaning procedure and SAS code made publicly available in Dick-Nielsen [2009] (and Dick-Nielsen [2012]). I then merge the TRACE data with the Mergent Fixed Income Securities data set in order to obtain bond specific details such as bond offering date, offering amount, maturity date, coupon rate, coupon type, interest payment frequency, bond type, bond rating, bond option features, and any issuer specific information.

I adopt the following standard filtering criteria for the TRACE intraday trades data:

- Remove bonds that are not listed or traded in the U.S. public market, this includes bonds that are private placements, issued under 144A rule, bonds that do not trade in U.S. dollars, and bond issuers not in the U.S. jurisdiction.
- Remove bonds that are structured notes, mortgage backed, asset backed, agency backed, and equity linked.
- Remove bonds that are convertible
- Remove bonds that have floating rate coupons (just include fixed rate).
- Remove bonds that have less than one year to maturity
- Bonds must have prices between 5 and 1000

The monthly corporate bond returns at time t (denoted, $R_{i,t}$) are computed as:

$$R_{i,t} = \frac{P_{i,t} + AI_{i,t} + C_{i,t}}{P_{i,t-1} + AI_{i,t-1}} - 1 \tag{2.3.1}$$

drikogiannopoulou and Papakonstantinou [2019] in the mutual fund literature, Mitton [2019] and Mulherin et al. [2018] for corporate finance, Holland et al. [2010] for accounting literature, and Heath et al. [2020] for natural experiments.

where $P_{i,t}$ is the transaction price, $AI_{i,t}$ is accrued interest, and $C_{i,t}$ is the coupon payment, if any, of corporate bond i in monthly t. I compute the bond excess return ($r_{i,t} = R_{i,t} - r_{f,t}$) by subtracting the risk free rate ($r_{f,t}$) which is a proxy for by using the one-month U.S. Treasury Bill Rate. With using TRACE intraday bond transaction data I calculate the daily clean price as the volume-weighted trading average of intra-day prices to minimize the effect of bid-ask spreads in trade prices as per Bessembinder et al. [2009].

Two scenarios are used in order to compute monthly frequency corporate bond returns to account for the infrequent trading of corporate bonds. The first scenario to compute a bond return at time t is (1) when a bond has traded at the end of month $t-1$ and at the end of month t. I take the closest to the last trading of the month as long at is within five days of the last trading day of the month. The second scenario to compute a bond return at time t is (2) when a bond has traded at the beginning of month t and at the end of month t. I take the closest to the first trading of the month as long at is within five days of the first trading day of the month.

I follow Bai et al. [2019b] (BBW, henceforth) to construct four bond specific factors based on the individual bond returns from transaction data. The BBW four factor corporate bond return model is a linear asset pricing factor model which encompasses a bond market return risk factor ($BMKT_t$), a corporate bond return liquidity risk factor (LRF), a corporate bond return credit risk factor (CRF_t), and a corporate bond return downside risk factor (DRF_t). In order to construct the corporate bond return factors I compute the individual bond characteristic values of bond illiquidity, credit quality, and downside risk. Monthly bond return factors are calculated by sorting individual bond returns each month into quintiles based on their characteristic values.

Bond illiquidity is measured on an individual bond basis by computing the Roll [1984] measure on corporate bond transaction data over the month. The credit ratings is used as a measure of credit quality of individual corporate bonds. Bond-level rating information is obtained through the Mergent Fixed Income Securities Database (FISD) of historical bond credit ratings. All credit ratings are assigned a integer number from 1 (AAA rating) to 21 (CCC rating) in order to facilitate credit risk comparability across bonds. Corporate bond return downside risk is measured as the individual bond's second lowest monthly return observation over the past 36 months (then multiplied by negative one for interpretation). Since the computation of the corporate bond return downside risk measure requires using 36 months of historical data our sample construction for this variable is calculated for the time period July 1, 2005 to October 31, 2019.

The corporate bond return factors are constructed by sorting individual bond returns each month into quintiles based on their characteristic values of bond illiquidity, credit quality, and downside risk. The bond market excess return factor is computed as the value-weighted average returns of all corporate bonds in the sample less the one-month Treasury bill rate. The downside risk factor for corporate bond returns is constructed for each month from July 1, 2005 to October 31, 2019 by forming bivariate portfolios by independently sorting bonds into five quintiles based on their credit rating and five quintiles based on their downside risk (measured by a 5% monthly Value at Risk (VaR)). I then compute the downside risk factor (DRF_t) is the amount-outstanding value weighted average return difference between the highest-VaR portfolio and the lowest-VaR portfolio across the rating portfolios. The credit risk factor (CRF_t) is the value-weighted average return difference between the lowest-rating (highest credit risk) portfolio and the highest rating (lowest credit risk) portfolio across the VaR portfolios. The liquidity risk and the return reversal factors are constructed in a similar fashion using independent sorts. The liquidity risk factor (LRF_t), is the value-weighted average return difference between the highest-illiquidity and the lowest illiquidity portfolios across the rating portfolios.

INSERT TABLE A.1.1 **HERE**

The final sample includes an average of $5,864$ bonds bond-month returns observations during the sample period July 1, 2002 to October 31, 2019. In Table A.1.1 we report the pooled, across the time-series and cross-sectional, bond returns and characteristics. The sample contains a total number of 1.035 million bond return observations (N) with an average monthly return of 0.58%. The average time to maturity is 9.49 years with an average age of the bond (time since the bond was issued) of 4.74 years. The average amount outstanding of the our sample of corporate bonds is 376 million dollars. Corporate bond returns in Table A.1.1 are windsorized at the 1% and 99% percentiles.

2.3.2 Cross-Sectional Signal Data

The cross-sectional asset pricing literature has largely been developed to determine signals that help understand, explain, and predict the cross-section of U.S. common stock returns. Many of the signals have been constructed from stock return, stock microstuctural, accounting/balance sheet ratio data and equity option trade price, signed volume, and order flow data. Analogous versions of the stock signals have been constructed using corporate bond returns and other corporate bond data.

Intra-day stock transactions and signed trade volume data is obtained from TAQ. Daily equity options end of day quotes data is obtained from OptionMetrics. OptionMetrics also provides non-parametric calculations of Black-Scholes option implied volatility, delta, gamma, and vega. I remove contracts from which the absolute Black-Scholes deltas for calls and puts in which the are above 0.98 and below 0.02 in addition to those that have missing open-interest and missing trade volume. I also remove option contracts that have less than 10 days to maturity due to rollover of option contracts. Intra-day equity option trade price data are obtained from LiveVol which provides intra-day options trades data and Black-Scholes deltas for calls and puts. I filter out those trades in which the implied volatility are above 0.98 and below 0.02 in addition to those that have missing trade volume and those have less than 10 days to maturity. These option filters are standard in the literature see Christoffersen et al. [2018]. End of day equity option signed volume data (new open/ close buy and sell positions) for each contract are obtained from CBOE/ISE Exchanges. Monthly stock return, price, and volume data are obtained from CRSP. I only consider common stocks with CRSP share codes of 10 or 11. Quarterly and Annual accounting data are obtained from the merged CRSP/COMPUSTAT files from the year July 1, 2002 to October 31, 2019. In order to adjust for the availability of the accounting information at the time of reporting, I lag the quarterly (annual) accounting data 3 (6) months.

I augment the set of signals with a larger more comprehensive set that have historically been discovered as predictors of the cross-section of stock returns (the signals set used in Green et al. [2017]). I compute 103 signals (of the 120) from Green et al. [2017] (GHZ, henceforth).[15]. The firm characteristics are constructed using all firms with common shares that are listed on the AMEX, NYSE, or NASDAQ, that have end of month value on CRSP, quarterly and annual balance sheet reporting on COMPUSTAT and earnings information reported to the $I/B/E/S$ data. The GHZ data set is available beginning from January 1, 1980, however, I only require data beginning from July 1, 2002 when TRACE corporate bond reporting began. The list of the 103 signals and their sample statistics that overlap in the CS of firms that have corporate bond outstanding are listed in Tables A.1.2 and A.1.4. I provide a complete and detailed list of the 143 signals and their summary statistics in Table A.1.3.

INSERT TABLES A.1.2, A.1.3, and A.1.4 **HERE**

[15] I thank Jeremiah Green for making the SAS code to construct the data set freely available on his website. The GHZ data has been used in other well known published studies such as Gu et al. [2020]

2.4 Main Empirical Results

2.4.1 Signal Discovery Rates in the Cross-Section of Corporate Bonds using Discovered t-statistics

Panel A of Figure A.1.1 below shows the history of signals that have been discovered (in published and in working papers) to predict the cross-section (CS, henceforth) of U.S. corporate bond returns. The number of discovered signals has steadily increased since the introduction of the centralized TRACE reporting system for U.S. corporate bond and fixed income reporting system in July 1, 2002. Papers published prior to the introduction of TRACE typically used either smaller subsets of bond data collected from the Wall Street Journal, Lehman Brothers Fixed Income database, or National Association of Insurance Companies (NAIC) bond transactions. Since 2015 the number of discovered signals that have been tested to predict the CS corporate bond returns has increased from less than 10 per year to over 15 per year in 2019.

INSERT FIGURE A.1.1 **HERE**

Overall I find a total of 56 papers with a total of 111 signals that have been discovered to predict the CS of corporate bond returns. Signals have been tried from different risk categories such as stock/bond/equity option return based (return, momentum, volatility, skewness), accounting/firm balance sheet ratios and levels (size, book-to-market ratio, investment), behavioural signals (earnings surprises, positive/negative earnings), and market microstructural (illiquidity measures, dealer inventory, effective spread trade measures). So far, in the cross-section of corporate bond returns it has been documented that: corporate bond and stock illiquidity are positive predictors of future bond returns (Lin et al. [2011]), financial intermediary risk exposure and bond supply both positively predicts future bond returns (He et al. [2017] and Goldberg and Nozawa [2019] respectively), systematic risk bond market return, a credit risk, a liquidity risk, and a downside risk factors all positively predict future bond returns (Bai et al. [2019b]), as well as many stock and balance sheet variables (Chordia et al. [2017], Choi and Kim [2018], and Chichernea et al. [2019]), and many others that have been discovered. A complete list of the papers and discovered signals that predict the cross-section of corporate bonds are reported in Table A.1.5.

INSERT TABLE A.1.5 **HERE**

Panel B of Figure A.1.1 shows a distribution of the absolute value of the signal's reported t-statistic. The majority of the absolute values of the t-statistics do in fact exceed the single hypothesis test (SHT, henceforth) level of 1.98 benchmark of a 5% significance (non-significant signals are typically those where the paper provides multiple different signals only some of which are significant). In my sample, a t-statistic of 1.98 corresponds to a $\alpha_{LOS} = 5\%$ significance level with a 210 degrees of freedom (nearly 18 years of monthly data, starting point of July 1, 2002) and a t-statistic of 2.60 corresponds to a $\alpha_{LOS} = 1\%$ significance level with a 210 degrees of freedom. In total I have 78 observations of signals with a t-statistic greater than 1.98 (with 57 greater than 2.60), 21 observed t statistics in the interval $(1.98, 2.60)$, 15 observed t statistics in the interval $(2.60, 3.12)$, 19 observed t statistics in the interval $(3.12, 3.87)$, and 23 with t-statistics greater than 3.87.[16]

In order to assess whether signals that originate from the bond literature perform better than those that originate from the stock return literature in predicting corporate bond returns, I need to evaluate whether the discovered t-statistic of the signal (in the original first paper claiming discovery) is statistically significant when considering a higher threshold. I establish the higher threshold by simulating t-statistics from a modified empirical distribution of that presented in Panel B of Figure A.1.1 and then applying standard statistical adjustments for multiple hypothesis testing. The empirical simulation technique I use follows from Harvey et al. [2016]. The framework assumes that the discovered t-statistic for a signal follows an Exponential distribution where the set of t-statistics is truncated at a particular statistical point (single hypothesis test level of a 5% level of significance) which is the statistical hurdle that the researcher needs to surpass in order to publish the discovered signal. There will be under-representation of smaller discovered t-statistics in the corresponding SHT t-statistics between the 5% and 1% level of significance. Due to this observation, an additional assumption is made that the sample of discovered t-statistics is only partially represented in this interval range of t-statistics. Hence the simulation of the benchmark t-statistics is done under independence sampling of the t-statistics of a sample size that also includes a fraction (which is termed sampling ratio (S.R.)) of all signals in the corresponding SHT t-statistics between the 5% and 1% level of significance.

I present the empirically simulated multiple hypothesis testing (MHT, henceforth) t-statistic benchmarks in Table A.1.6 using the following standard adjustments for MHT that have been used in the finance literature: Bonferroni [1936] (Bonf/Bonferroni henceforth), Benjamini and Yekutieli [2001] (BHY henceforth) and Holm [1979] (Holm, henceforth). I

[16] For the remainder of the paper I denote the statistical level of significance (as well as the Family-Wide Error Rate and False Discovery Rate error rate (Type I error rate)) for a SHT as α_{LOS}. The use of the subscript LOS (level of significance) is to differentiate the α_s from that used as the intercept in the portfolio sorts in equation 2.4.1.

provide a brief overview of MHT techniques in Section A.1.3. Benchmark t-statistics are presented assuming different sampling ratios of under-representation of smaller discovered t-statistics. For example a sampling ratio of $r = 2$ implies that the fraction $1/r = 1/2$ of the corresponding SHT t-statistics between the 5% and 1% level of significance were assumed to not been reported and hence needs to be counted for an additional $r = 2$ times. Upper and lower 10% confidence intervals (C.I.) are presented around each of the benchmark t-statistic estimates.

INSERT TABLE A.1.6 **HERE**

Assuming a sampling ratio of $r = 2$, I find that the number of discovered signals that have been falsely rejected, rejected under a SHT framework (57 with an absolute t-statistic over 2.60) but not rejected under the MHT framework benchmark, are 34 under the Bonferroni, 34 under Holm, 33 under BHY (at a $\alpha_{LOS} = 1\%$ and 15 under BHY (at a $\alpha_{LOS} = 5\%$). These results imply false discovery rates of signals in the CS of corporate bond returns of 59.6% under Bonferroni, 59.6% under Holm, 57.9% under BHY (at a $\alpha_{LOS} = 1\%$), and 26.3% under BHY (at a 5%).

Harvey et al. [2016] find that many of the stock signal discoveries are in fact false discoveries under their MHT framework. Of the 296 published signals they find false discovery rates of 53% under Bonferroni, 48% under Holm, 45% under BHY (at a $\alpha_{LOS} = 1\%$), and 27% under BHY (at a $\alpha_{LOS} = 5\%$).[17] The false discovery rates that I find for signals discovered in the CS of corporate bond returns are quite similar to those in Harvey et al. [2016].

Under the framework presented in this section, some examples well known discovered signals that are deemed true signals, i.e. predict the CS of corporate bond returns that have been rejected under BHY (at a $\alpha_{LOS} = 5\%$), are: the stock return momentum, stock return reversal, expected default frequency measure of Bharath and Shumway [2008], short term bond reversal discovered in Chordia et al. [2017], the stock and bond illiquidity factor of Pastor and Staumbaugh [2003] (and bond illiquidity measured using Amihud [2002]) all discovered in Lin et al. [2011], the credit risk factor of Bai et al. [2019b], and others. Examples of well known discovered signals that predict the CS of corporate bond returns that are found to be false discoveries under this framework, i.e. have not been rejected under BHY (at a 5% significance level), are: the default risk factor of Gebhardt et al.

[17]Of the 296 published factors (requiring an absolute t-statistic to be over 2.57) they find the number that are false discoveries: 158 under the Bonferroni, 142 under Holm, 132 under BHY (at a $\alpha_{LOS} = 1\%$) and 80 under BHY (at a $\alpha_{LOS} = 5\%$).

[2005a], the bond permanent component adaptation of Sadka [2006] in Lin et al. [2011], seasonal difference in quarterly earnings in Easton et al. [2009], the intermediary factor of He et al. [2017], the total monthly stock return volatility of Chung et al. [2019], the $CVOL$ measure of An et al. [2010] in Cao et al. [2020], and others.

Of the 111 signals, 52 (59) are corporate bond (stock) based signals of which 28 (29) exceed the single hypothesis testing benchmark of 2.62 (at a level of significance of 1%). Of the 28 corporate bond signals 21, 13, and 12 are found to exceed the multiple hypothesis testing t-statistic benchmarks of 3.12, 3.67 and 3.87 (benchmarks found in Table A.1.5) respectively. Corresponding of the 29 stock signals, 21, 11, and 11 are found to exceed the the multiple hypothesis testing t-statistic benchmarks 3.12, 3.67, and 3.87, respectively. These frequencies of rejections lead to false discovery rates of 25%, 54% and 57% (using multiple hypothesis testing t-statistic benchmarks 3.12, 3.67, and 3.87) for corporate bond signals and 28%, 62%, and 62% (using multiple hypothesis testing t-statistic benchmarks 3.12, 3.67, and 3.87) for the stock signals. The lower false discoveries in the corporate bond based signals over stock based signals (at each of the different hypothesis testing methods) suggest that corporate bond signals, over stock signals, better predict the cross-section of corporate bond returns.

2.4.2 Signal Discovery Rates in the Cross-Section of Corporate Bonds: Empirical Evidence from Portfolio Sorts and Cross-Sectional Regressions

In this section I test the empirical performance of 143 different signals in univariate cross-sectional regressions and portfolio sorts. My goal is to see if how many of the signals are statistically significant signals in predicting the cross-section of future corporate bond returns as well how many are statistically significant after incorporating MHT benchmarks. I briefly outline the univariate cross-sectional regressions and portfolio sort techniques before proceeding to the main empirical results in Section 2.4.3.

Every month corporate bond returns are sorted into value-weighted (by principal amount outstanding) decile portfolios based on a one month lagged signal $s_{i,t-1}$ and hold the portfolio for one month. I choose to re-balance the portfolios on a monthly frequency as oppose to annually or quarterly in order to account for the effects of short term impacts of signals such as corporate bond illiquidity and short term reversal return. In order to evaluate the realized portfolio alphas I use the corporate bond return linear factor model of Bai et al. [2019b] (BBW henceforth) which includes a bond market return factor ($BMKT_t$), a credit risk factor (CRF_t), a liquidity risk factor (LRF_t), and a downside risk factor (DRF_t).

$$r_{s,t} = \alpha_s + \beta_s^{BMKT} \cdot BMKT_t + \beta_s^{CRF} \cdot CRF_t + \beta_s^{LRF} \cdot LRF_t + \beta_s^{DRF} \cdot DRF_t + u_{s,t} \tag{2.4.1}$$

The long-short hedge portfolio as define as taking a long position in the decile 10 and short in decile 1, hence the long-short portfolio return can be either positive or negative. I then run a time-series regression of the bond long-short portfolio excess returns $r_{s,t}$ on the BBW factor returns and record the t-statistic of the long-short portfolio alpha denoted (t_{α_s}). The t-statistic of the long-short portfolio alpha has standard errors that are adjusted for heteroskedasticity and 3 lags of autocorrelation using the method Newey and West [1987].

Additionally, I evaluate the ability of the signal to predict corporate bond returns by estimating the following univariate cross-sectional regression[18] on signal $s_{i,t-1}$ for each month:

$$R_{i,t} - r_{f,t} - \hat{\beta} \cdot F_t = \lambda_0 + \lambda_s \cdot s_{i,t-1} + \lambda_z \cdot Z_{i,t-1} + e_{i,t} \tag{2.4.2}$$

where $F_t = \{BMKT_t, CRF_t, LRF_t, DRF_t\}$, where $s_{i,t-1}$ represents the signal and $Z_{i,t-1}$ are control variables that include the coupon amount, credit rating (from Moody's provided in TRACE), log of the age of the bond since issuance, and log of the remaining time to maturity. The corporate bond excess returns are adjusted as per Brennan et al. [1998] with the four factor corporate bond factor model of BBW (the same bond factor model is used to calculate the replicating decile portfolio alphas). In the estimation of the bond return betas I use 48 months of historical return bond data with at least 12 bond returns. All bond betas are windsorized at the 1 and 99 percentile levels. I then calculate the average λ_s regression slope coefficient using the Fama and Macbeth [1973] (FM henceforth) regression methodology and it's heteroskedastic adjusted t-statistic (t_{λ_s}) where standard errors are adjusted for 3 lags of autocorrelation using the Newey and West [1987] adjustment.

I evaluate 3 different specifications of the high minus low portfolio alphas using the time-series regression of the hedge returns in equation 2.4.1: (1) the high-minus-low portfolio is formed using deciles and portfolio returns are adjusted using the BBW corporate bond factor model (Decile PS (BBW), henceforth). (2) the high-minus-low portfolio is formed using quintiles and portfolio returns are adjusted using the BBW corporate bond factor model (Quintile PS (BBW), henceforth). (3) the high-minus-low portfolio is formed using quintiles and portfolio returns are adjusted using the corporate bond return linear

[18] I refer to equation 2.4.2 as a univariate cross-sectional regression since I am only interested in the magnitude of the signal coefficient (λ_s) even though I add the control variables: coupon amount, credit rating, log of the age of the bond since issuance, and log of time to maturity.

factor model of BBW augmented with the stock market linear factor model of Fama and French [2015] (Quintile PS (BBW+FF5), henceforth) which includes a stock market factor ($SMKT_t$), size (SMB_t), value (HML_t), profitability (RMW_t), and investment (CMA_t).

I also evaluate three different specifications of the univariate FM regression equation 2.4.2: (1) the ability of the signal to explain corporate bond returns by estimating the univariate cross-sectional regression for each month as in equation 2.4.2 (FM (BBW) henceforth). (2) I also re-compute the univariate λ_s of the FM regression equation 2.4.2 where the corporate bond excess returns are adjusted as per Brennan et al. [1998] using the BBW corporate bond return linear factor model augmented with the stock market linear factor model of Fama and French [2015] (FM (BBW+FF5), henceforth). (3) I also estimate the FM equation 2.4.2 based on value weighted least squares weighting bonds on the market capitalization (size) of the company common shares (VW FM (BBW) henceforth) since empirical evidence from discovered signals in the CS of stock returns Green et al. [2017] and Hou et al. [2020] that the existence of predictability has typically been associated with over-weighting portfolios or stocks that have small market capitalization. In each of the FM regressions I use the same control variables (coupon amount, credit rating, log of the age of the bond since issuance, and log of time to maturity).

2.4.3 Empirical Evidence from Univariate Portfolio Sorts and Cross-Sectional Regressions

The estimated λ_s and α_s for each of the 143 signals (and their corresponding t-statistics t_{λ_s} and t_{α_s}) for each of the three types of univariate FM regressions and for each of the three PS regressions are individually reported in Table A.1.7. Of the list of 143 signals used in my empirical analysis, several are discovered to predict the CS of corporate bond returns. In particular, I note that the idiosyncratic stock volatility negatively predict the CS of corporate bond returns (as discovered in Chung et al. [2019]. However, Chordia et al. [2017] did not find any statistical significance. The CVOL, PVOL, and change in implied volatility measure of An et al. [2010] (Cao et al. [2020]), the idiosyncratic bond volatility measure of Bai et al. [2020c] (different from the idiosyncratic bond volatility measure of Chung et al. [2019]) and the EDF measure (Vassalou and Xing [2004] and Bharath and Shumway [2008]) all negatively predict the CS of corporate bond return. The systemic uncertainty risk of Jurado et al. [2015] (Bai et al. [2020a]), and change in non-current operating accruals (Chichernea et al. [2019]) and one month stock return reversal (Chordia et al. [2017]) are all positive predictors of the CS of corporate bond returns.

Several signals, that are well known to predict the CS of stock returns, that are statistically significant under SHT for FM and PS that have not been discovered to predict the CS of corporate bond returns (the original paper where the signal is discovered in the CS of stock returns is in parentheses) are: change in stock momentum (Gettleman and Marks [2006]), dispersion of analyst beliefs (Diether et al. [2002]), and change in analyst forecast (Hawkins et al. [1984]). Despite the fact that I do evaluate some signals that have already been discovered in the cross-section of corporate bond returns, my results should not be taken as an attempt to directly replicate the published bond market anomalies in their original studies (as was done in the stock market anomalies literature Hou et al. [2020]). There are three possible reasons for this: (1) The time period of the data will differ from the original papers since several of the papers use monthly corporate bond quotes data from the Lehman Brothers Fixed Income database or Thompson Reuters Tick Data which provide corporate bond quote data from the late 1970's onwards, or National Association of Insurance Companies (NAIC) bond transactions data which provides a subset of corporate bond transaction prices made by insurance companies during the 1980s and 1990s. The Lehman Brothers Fixed Income and NAIC datasets are no longer publicly available for purchase. (2) my FM and PS results use the BBW factors of Bai et al. [2019b] which have only recently been published and made available as of 2019, most prior studies use the stock and bond factors (term and default) of Fama and French [1993] (or Fama and French [2015]) augmented with stock momentum factor of Cahart [1997] and the market liquidity factor of Pastor and Staumbaugh [2003] (3) Several of the original papers do not present FM regressions adjusted using the Brennan et al. [1998] technique.

Table A.1.15 displays the out-of-sample R-squared (R^2_{OS}) for the machine learning method, as well as the high-minus-low portfolio α_s (and corresponding t_{α_s} and monthly Sharpe Ratio), under each of the machine learning methods for each of the three PS regressions.[19] Table A.1.15 Panel A presents the results for all corporate bonds and Panels B/C present the results separately for investment grade and speculative grade respectively. Only those t-statistics t_{α_s} of the LASSO and the Enet methods are significant when considering all corporate bonds but are only marginally significant when separating into investment grade and speculative grade bond subsets.

INSERT TABLE A.1.7 **HERE**

Table A.1.8 summarizes the statistical distribution properties (mean, median, standard deviation, and percentiles) for each of the three types univariate FM regressions and for

[19] I provide a description of each of the machine learning methods used in my empirical analysis in Appendix Section A.1.5 and a description of the performance attribution of the machine learning methods in Appendix Section A.1.4. Diebold-Mariano test statistics are calculated for all corporate bonds in Table A.1.16.

each of the three PS regressions that are individually reported in Table A.1.7. Table A.1.8 Panel A shows the distribution of the three sets of λ_s and α_s whereas Panel B shows the distribution of the three sets of the t_{λ_s} and t_{α_s}. The mean/median monthly α_s for each of the three PS are close to zero (ranging from -0.025% to 0.00% per month), consistent with signals earning 0.00% per month on average. The standard deviation of the signal α_s is between 0.18% and 0.29% per month and tail values ranging from -1.33% to 0.84% per month. The mean/median monthly λ_s ranges from -0.5% bps to 0.19%. In general, each of the three FM regression and PS mean/median t_{λ_s} and t_{α_s} are quite close to zero (between -0.27 and 0.08), however, tail values of t_{λ_s} and t_{α_s} range (in absolute value) from over 2.82 up to 7.64 in the three FM regressions and from 3.31 to 4.89. Figure A.1.2 displays the histograms of the λ_s (and corresponding t_{λ_s}) for each of the three types of univariate FM regressions and Figure A.1.3 displays the histograms of the α_s (and corresponding t_{α_s}) for each of the three PS regressions. The two distributions are centered around zero and appear to be symmetrically Normally distributed which is consistent with the sample statistics.

INSERT TABLE A.1.8 **HERE**

INSERT FIGURE A.1.2 and A.1.3 **HERE**

The distribution of t_{λ_s} and t_{α_s} range from over -16.57 up to 7.51 in the three FM regressions and from -4.56 to 3.96 in the three sets of PS. From Panel B in Table A.1.8 it is clear that of the signals in my sample a larger fraction than 5% exceed the SHT level. I find that the fraction of null hypotheses that are rejected under SHT for each of the three PS regressions are 53/143 = 37% (Decile PS (BBW)), 38/143 = 34% (Quintiles PS (BBW)), and 48/143 = 26% (Quintiles PS (BBW+FF5)). Similarly, the fraction of null hypotheses that are rejected under SHT for each of the three FM regressions are 30/143 = 21% (FM (BBW)), 32/143 = 24% (FM (BBW+FF5)), 24/143 = 17% (VW FM (BBW)). In general, there are considerably more signals that reject the null hypothesis, that the signal does not predict the CS of the corporate bond returns, under the three PS regressions as opposed to the three FM regressions. The fewest rejections occur under the size-weighted FM CS regression. Nevertheless when considering each of the signals individually under SHT, the percentage of rejections that is observed is far beyond the desired 5% level.

INSERT TABLE A.1.9 **HERE**

The left hand side panel of Table A.1.9 displays the benchmark t-statistics for each of the different specifications of the three univariate FM regressions and each of the three PS regressions under SHT and under the Bonferroni, Holm, and BHY MHT methods. Depending on the specification of the FM and PS, the benchmark t-statistics can range up to from 3.35 to 3.9 but are on average about 3.53. The middle panel of Table A.1.9 shows the fraction of the 143 signals whose t-statistics, of testing the null hypothesis that the signal does not predict the CS of corporate bond returns, exceed the SHT and Bonferroni, Holm, and BHY MHT benchmark t-statistics. The right hand side panel of Table A.1.9 computes the ratio of the number of t-statistics that exceed the SHT threshold but are not rejected under the MHT thresholds (single not multiple rejection, SnM or false discovery rates). For each of the three FM regressions and three PS, when imposing the Bonferroni, Holm, and BHY MHT tests results in rejection rates, of the null hypothesis of no return predictability, between $0 - 4.2\%$ (on average about 2.49%) which leads to false discovery rates of $76 - 100\%$. Those signals whose t-statistics are higher than the MHT benchmark are identified as true discoveries of signals that predict the cross-section of corporate bond returns.

2.4.4 Integration of Stock and Corporate Bond Markets: Evidence from Multivariate FM Regression Horse Race in the Cross-section of Corporate Bond Returns

My results in Section 2.4.3 highlight that many of the corporate bond and stock signals have limited or no ability to predict the cross-section of corporate bond returns when considering higher statistical thresholds once accounting for all other signals being tested simultaneously and discovered in the literature. In this section, I ask do corporate bond signals better predict the cross-section of corporate bond returns than stock signals?

To assess whether corporate bond signals better predict the cross-section of corporate bond returns than stock signals I run a horse race of a total of 129 corporate bond and stock signals by simultaneously including 129 of the 143 signals in a multivariate FM cross-sectional predictive regression in the cross-section of the corporate bond returns.[20] In order to be able to include as many signals as possible, I follow Green et al. [2017] and set missing values of the signal to the cross sectional average value of that signal.

I present the results in six multivariate FM cross-sectional predictive regressions which includes using three different dependent variables (allowing for different risk adjustment

[20]Not all of the 143 signals were able to be included due to a shorter data history or limited cross-sectional variability.

in excess returns) and value-weighted least-squares estimates of each of those three specifications. First I use excess corporate bond return as a baseline but I also present risk adjusted excess corporate bond returns adjusted for the four factor corporate bond model of Bai et al. [2019b], in order to account for the different systematic risk exposures that corporate bond returns face. Additionally I present results using excess corporate bond returns risk adjusted for corporate bond return and stock return factors (using both the Bai et al. [2019b] and Fama and French [2015] five factor stock model) in order to adjust for systematic risk exposures in corporate bonds and stocks.[21] Finally I present results using value weighted least squares which weight bonds by the market capitalization (size) of the company common shares. This is done to control for the over-weighting of regression estimates in firms which have small market capitalization and whose signal predictability might end up only existing in a small subset of the market.

INSERT TABLE A.1.10 **HERE**

Each two columns in Table A.1.10 report the average regression slope coefficient (λ_s) and corresponding t-statistic (t_{λ_s}) for each of the 129 signals in the six multivariate FM cross-sectional one-month-ahead predictive regressions of the corporate bond returns when simultaneously including all 129 signals.[22] In Table A.1.10 columns 2 and 3 use excess corporate bond returns, columns 6 and 7 are risk adjusted using the Bai et al. [2019b] four factor corporate bond return model, columns 10 and 11 are risk adjusted using the Bai et al. [2019b] and the Fama and French [2015] factors.[23]

INSERT TABLE A.1.11 **HERE**

[21]As shown in Bao and Hou [2017], there are cases where corporate bond returns can behave, or co-move, like the firm's stock return based on the time to maturity, credit riskiness, and de-factor seniority of the bond. Hence I believe that it is important to control for systematic exposure to the stock market risk factors.

[22]Note that all signals are lagged by one month to the left-hand-side corporate bond return variable and t-statistics are computed using standard errors are adjusted for three lags of autocorrelation using the Newey and West [1987] method.

[23]In Table A.1.10 columns 4 and 5 represent the corresponding value-weighted (VW), by one month lagged firm market capitalization, least squares estimated multivariate FM regression using excess corporate bond returns as a dependent variable, columns 8 and 9 represent the corresponding VW least squares estimate multivariate FM regression using risk adjusted corporate bond returns (risk adjusted using Bai et al. [2019b]) as a dependent variable, and columns 12 and 13 represent the corresponding VW least squares estimate multivariate FM regression using risk adjusted corporate bond returns (risk adjusted using Bai et al. [2019b] and Fama and French [2015] five factor stock model) as a dependent variable. The risk-adjusted corporate bond excess returns are calculated using the method of Brennan et al. [1998].

For each of the six multivariate regressions shown in Table A.1.10 I compute the benchmark t-statistics, using different multiple hypothesis testing methods (Bonferroni, Holm, and BHY methods), which depend on the t-statistics of all signals used in the multivariate regression.[24] Resulting multiple hypothesis testing benchmark t-statistics for a signal in the multivariate regressions range from 3.48 to 4.09, which are the required benchmark that the signal t-statistic must exceed in order to be deemed a true discovery.

Of the 129 signals 17 (or 22 of the 134 signals including the control variables) are constructed using only corporate bond specific information (data from TRACE). The remaining 112 signals are those that have been discovered to predict the cross-section of stock returns. Of the 129 signals it can be seen that the stock short term return reversal, log (age), log (TTM), log (amount outstanding), corporate bond return volatility, and corporate bond return skewness are the only signals that have t-statistics that exceed the multiple hypothesis testing benchmarks t-statistics in each of the different multiple hypothesis testing benchmark methods in at least one of their respective multivariate regressions.

Corporate bond return skewness negatively predicts future corporate bond returns in each of the six multivariate regressions. In 5 of the 6 multivariate regressions, the corporate bond return skewness t-statistic (with absolute t-statistics of -4.38 to -7.08) exceeds the multiple hypothesis testing benchmark for each of the different multiple hypothesis testing benchmark methods. In the case of the regression in columns 12 and 13 (VW least squares estimate multivariate FM regression using risk adjusted corporate bond returns as per BBW and FF5) the absolute t-statistics of -3.5 does not exceed the required benchmark of 4.08 under any of the three different MHT methods. Stock short term return reversal, corporate bond return volatility, log (age) and log (TTM) each positively predict future corporate bond returns in each of the multivariate regressions. However, each of the signals only exceeds the multiple hypothesis testing benchmark once or twice in each of the six multivariate regressions.

Hence in each of the six multivariate predictive regressions, there is a higher ratio of corporate bond signals (at highest 3/17, at lowest 0/17) instead of stock signals (at highest 1/17, at lowest 0/17), that predict the cross-section of corporate bond returns. The predictability is driven by corporate bond return skewness (with the exception of regression in columns 12 and 13). My results, as well as, Bali et al. [2020], find that there is no difference in the predicted corporate bond return spread of machine learning methods when adding stock signals to the corporate bond signals. Hence the results of Bali et al.

[24] The left hand side panel of Table A.1.11 displays the benchmark t-statistics for each of the different MHT method specifications (and the middle and right-hand side panels show the fraction of the signals whose t-statistics exceed the single and multiple hypothesis testing benchmark t-statistics and the single and not multiple rejection ratios, respectively).

[2020] and this paper both find, using different methods, that stock signal predictive power is economically insignificant whereas corporate bond signals are important in predicting corporate bond returns. This result confirms my hypothesis in Section 2.2.

2.4.5 Integration of Stock and Corporate Bond Markets: Evidence from Multivariate FM Regression Horse Race in the Cross-section of Stock Returns

To assess whether corporate bond signals help in predicting the cross-section of stock returns I run a horse race by simultaneously including 129 of the 143 signals in a multivariate FM cross-sectional predictive regression on the cross-section of the stock returns. In order to be able to include as many signals as possible, I follow Green et al. [2017] and set missing values of the signal to the cross sectional average value of that signal. I present the results in four multivariate FM cross-sectional predictive regressions which includes value-weighted least-squares estimates (by the market capitalization (size) of the company common shares) in two of those specifications (see columns 2 and 4 of Table A.1.12) as well as conditioning the analysis on only firms whose stock price is greater than $5 (see columns 3 and 4 of Table A.1.12) in order to avoid the effect being driven by small micro-cap firms.

INSERT TABLE A.1.12 **HERE**

INSERT TABLE A.1.13 **HERE**

The left hand side panel of Table A.1.13 displays the benchmark t-statistics for each of the different multiple hypothesis testing method specifications (and the middle and right-hand side panels show the fraction of the signals whose t-statistics exceed the single and multiple hypothesis testing benchmark t-statistics and the single and not multiple rejection ratios, respectively). Resulting multiple hypothesis testing benchmark t-statistics for a signal in the multivariate regressions range from 3.43 to 3.9.

Of the 129 signals tested in Table A.1.12 none have t-statistics that exceed the multiple hypothesis testing benchmarks t-statistics in each of the different multiple hypothesis testing benchmark methods in any of the multivariate regressions. Bali et al. [2020] find that corporate bond signals do not provide any incremental predictive power beyond equity signals when combining stock and bond signals. In a similar vein, I find that corporate bond signals do not predict future stock returns when adjusting for multiple hypothesis testing. The result is consistent with the additional findings of Bali et al. [2020] that corporate bond signals do not provide any incremental predictive power beyond equity signals in predicting stock returns.

2.4.6 Integration of Stock and Corporate Bond Markets: Evidence from Conditional Double Sorts in Predicting the Cross Section of Corporate bond and Stock Returns

Muller and Schmickler [2020] study double sort combinations of 102 stock signals and find hundreds of profitable signals in excess of transaction costs and whose performance is on par with recent machine learning strategies. Of particular interest would be the performance of corporate bond specific signals versus stock signals in predicting the cross-section of corporate bond returns (and separately using the same set of test signals to predict the cross-section of stock returns).

I perform monthly 5×5 conditional double sorts (value-weighted using the amount outstanding of the corporate bonds) of all combinations of 143 signals in predicting the cross-section of corporate bond returns. Each month, I sort corporate bonds into quintile portfolios based on a signal S_1 and then within each quintile sort into quintiles on S_2.[25] This results in $\binom{143}{2} = 10,153$ combinations of 5×5 conditional double sorts. I explicitly focus on the four corner portfolios of the 5×5 matrix of conditionally sorted portfolios. I denote the corner portfolios as HH, HL, LH, and LL depending on whether the corporate bond return is assigned to the high or low quintile based on the ranking of the first signal and on the second (conditionally on it's ranking within the first signal).[26] For example, a corporate bond return is assigned to the high-low portfolio if it has a rank greater than 0.8 on signal S_1 and a rank less than 0.2 on signal S_2. Once accounting for all four combinations of corner high minus low portfolios, this results in a total of $4 \times 10,153 = 40,612$ high-low portfolios.

For each of the four high-minus-low quintile portfolios, I compute the portfolio alpha (α_{S_1,S_2}) using a time-series regression on the four factor corporate bond return factor model of Bai et al. [2019b] and t-statistics ($t_{\alpha_{S_1,S_2}}$) using the Newey and West [1987] method for computing standard errors adjusted for three lags.[27] When imposing standard multiple

[25] In unreported results I do consider the cases of sorting first on a signal S_2 and then within each quintile sort into quintiles on S_1, the set of resulting statistically significant combinations of signals is similar.

[26] For ease of exposition a diagram of my naming convention is presented in Figure A.1.4.

[27] Additionally I compute the portfolio alpha adjusting for corporate bond return and stock return factors using both the Bai et al. [2019b] and Fama and French [2015] five factor stock model.

hypothesis testing methods (Bonferroni, Holm, and BHY) to the $40,612$ portfolios, 52/ $40,612 = 0.13\%$ of double sort combinations of signals are true discoveries when applying higher thresholds.[28]

INSERT TABLE A.1.14 **HERE**

Table A.1.14 shows the 52 double sorted portfolios whose t-statistics are higher than the multiple hypothesis testing benchmark (I use the benchmark t-statistic of 5.00). The portfolios in Table A.1.14 are denoted by the first signal sorted on by S_1 then the second signal S_2, then the high-minus-low quintile portfolio corner (PF), the monthly portfolio Sharpe Ratio (SR), the portfolio alpha (α_{S_1,S_2}), and the corresponding t-statistics ($t_{\alpha_{S_1,S_2}}$). Of the 52 double sorted portfolios who exceed the multiple hypothesis testing benchmark, 31 of them involve sorting on the age of the corporate bond. The 31 double sorted portfolios involved sorting on the age of the corporate bond, the Sharpe Ratios of these signals are all positive and range from 0.04 to 0.24 monthly.

As a robustness test, I compute the portfolio alpha adjusting for corporate bond return and stock return factors using both the Bai et al. [2019b] and Fama and French [2015] five factor stock model for the $40,612$ portfolios. When imposing standard multiple hypothesis testing methods to the $40,612$ portfolios, using the benchmark t-statistic of 5.00, $64/40,612 = 0.16\%$ of double sort combinations have t-statistics that exceed the higher thresholds. Of the 64 double sorted portfolios who exceed the multiple hypothesis testing benchmark, 32 of them involve sorting on the age of the corporate bond and the Sharpe Ratios of these signals are all positive and range from 0.04 to 0.24 monthly as in my main results in Table A.1.14.

I perform monthly 5×5 conditional double sorts (value-weighted using the firm size, market capitalization) of combinations of 143 signals in the cross-section of stock returns(for firms with corporate bonds outstanding). I sort stocks into quintile portfolios based on a signal S_1 and then within each quintile sort into quintiles on S_2 For each of the four high-minus-low quintile portfolios, I compute the portfolio alpha (α_{S_1,S_2}) using a time-series regression on the four factor corporate bond return factor model of Bai et al. [2019b] and t-statistics ($t_{\alpha_{S_1,S_2}}$) using the Newey and West [1987] method for computing standard errors adjusted for three lags.[29] When imposing higher thresholds to the over $40,000$ high-minus-low portfolios, only one of double sort combinations of signals are true

[28]Bonferroni, Holm, and BHY benchmark t-statistics range from 4.58 to 5.00, with rejection rates, of the null hypothesis of no return predictability, of $0.1-0.2\%$, which leads to false discovery rates of $99-99.4\%$.

[29]Additionally I compute the portfolio alpha adjusting for corporate bond return and stock return factors using both the Bai et al. [2019b] and Fama and French [2015] five factor stock model.

discoveries in predicting the cross-section of stock returns and it does not rely on sorting on a corporate bond signal. Table A.1.14 shows that corporate bond signals are important in predicting the cross-section of corporate bond returns but do not help in predicting the cross-section of stock returns.

2.5 Robustness Tests

A particular concern that may be affecting the cross-section of corporate bond returns is that a firm could have many bond issues of varying amount outstanding which could impact our testing procedure. Bao et al. [2011] show that the firm's most recently issued bond and the firm's bond with the shortest maturity are the most liquid bonds.

In this section I provide several robustness tests, motivated by the findings Bao et al. [2011], to support our main empirical findings in the Internet Appendix. Robustness tests of the univariate FM regressions and PS using: (1) a firm's bond with the shortest time to maturity, (2) a firm's bond with lowest age since issuance, and (3) average bond return across all bonds for the same underlying firm. The resulting estimated λ_s and α_s (and corresponding t-statistics t_{λ_s} and t_{α_s}) for each signal for each of the three univariate FM regressions and for each of the three PS regressions are individually reported for the firm's bond with the shortest time to maturity, firm's bond with lowest age since issuance, and the average bond return across all bonds for the same underlying firm. For each of the three robustness tests mentioned, I compute the frequency of null hypotheses (no future predictability of the characteristic) that are rejected under SHT, under each of the MHTs, and the fraction of false discoveries. The results for each of the three robustness tests are similar to those in Table A.1.9 where we find a higher percentage of SHT rejections in the three different portfolio sorts as oppose to the three FM regressions, benchmark t-statistics and SnM rejection rates under each of the Bonferroni, Holm, and BHY multiple hypothesis test methods tend to be similar to those in Table A.1.9. Results are presented for the firm's: average bond return across all bonds for the same underlying firm, for firm's bond with lowest age since issuance, and bond with the shortest time to maturity.

2.6 Conclusion

The infamous Figure 2 in Harvey et al. [2016] shows the increase in the number of signals being discovered that predict the cross-section of stock returns and the increasing rate in which they have been discovered over the recent years. Thus coining the phrase from

Cochrane [2011] Factor Zoo of stock characteristics and factors that predict the cross-section of stock returns. I notice a similar increase in the number and rate of discovered signals that predict the cross-section of U.S. corporate bond returns.

I determine higher statistical benchmarks for discovering a signal that predicts the cross-section of corporate bond returns based on the signals in published/working papers as well as in empirical tests using Fama and Macbeth [1973] cross-sectional regressions and portfolio sorts using the bond factor model of Bai et al. [2019b]. When applying higher benchmarks to the t-statistics of signals in published/working papers of the cross-section of corporate bond returns under the multiple hypothesis testing framework I find a comparable higher benchmark to that documented in Harvey et al. [2016] in the cross-section of stock returns. When applying higher benchmarks to a sample set of 143 signals in cross-sectional regressions and portfolio sorts under the multiple hypothesis testing framework I find roughly 2.4% are true discoveries. Rates of true discoveries and benchmark t-statistics using the bond with the shortest age, shortest time to maturity, and average across all bonds for each firm are quantitatively similar.

In a horse-race using a multivariate cross-sectional regression using all signals to predict the cross-section of corporate bond returns, I find a higher number of corporate bond signals, rather than stock signals, that drive return predictability of the cross-section of corporate bond returns. Results are driven by bond skewness, and bond age. In a second horse-race, using all signals to predict the cross-section of stock returns, I find that no corporate bond signals drive stock return predictability. Both sets of horse-racing results are robust to different factor specifications, value-weighting by firm size (not to overweight micro-cap firm corporate bond returns), and different multiple hypothesis testing methods. My results have implications for the trading and portfolio management signals that corporate bond institutional investors (pension plans, mutual funds, and fixed income hedge funds) use.

Chapter 3

Do Option Implied Measures of Stock Mispricing Find Investment Opportunities or Market Frictions?

3.1 Introduction

A number of studies, including An et al. [2010], Cremers and Weinbaum [2010], Hu [2014], Johnson and So [2012], Manaster and Rendleman [1982], Muravyev et al. [2020], Pan and Poteshman [2006], and Xing et al. [2010] propose option-based measures of stock mispricing. These measures are different from each other, but fall into three categories. Some use differences between implied and actual stock prices, others rely on differences in implied volatilities across different options or over time, and still other measures are based on trading volume. Recent research demonstrates that these measures predict abnormal stock returns. To date however, nobody has compared what these measures capture and when they do and do not work. In this paper, we use nine option-based measures of stock mispricing to predict stock returns. Much of the predictability comes from illiquid or hard-to-borrow stocks, but some of the measures appear to generate significant positive abnormal returns from long-only strategies.

There are two necessary conditions for options to predict stock returns. First, at least some informed investors must trade options. There are good reasons for them to do so. Informed investors with positive information can obtain far greater implicit leverage by purchasing calls than they could get by buying stock on margin. Similarly, by purchasing puts, informed investors who believe a stock is overpriced can take a bearish position that requires far less collateral than is required to short the stock directly. For informed investors, buying puts provides additional advantages over shorting stock. If the stock price increases, an investor with a short position must post additional margin or close part of his position. No additional money is required from a put holder. A short seller must borrow shares to short and it is sometimes costly or difficult to do so. If the borrowed shares are recalled and the short seller is unable to locate new shares, the short position may have to be closed prematurely. Put buyers, on the other hand, can maintain their position until the option expires.

The second condition for options to predict stock returns is that stock prices must adjust slowly to information in options. In general, economists attribute slow adjustment of prices to information to market frictions or limits of arbitrage. These include the costs of trading, particularly for illiquid securities, and the risk that prices can move in the wrong direction before finally incorporating information. There are additional important frictions that can prevent stock prices from responding quickly to negative information. Many institutional investors, like mutual funds, are restricted to long positions. They cannot short stocks. In addition, some stocks are very difficult or expensive to short. These hard-to-borrow stocks are often smaller companies. Market frictions are likely to be especially important in slowing the response of stock prices to information in options because market frictions may lead informed investors to trade options rather than the stock. For example, informed investors may choose to trade options because it is difficult to sell the underlying stock short.

We find that the stocks that option-based mispricing measures identify as mispriced are disproportionately small or hard to borrow. When we sort them into value-weighted, rather than equal-weighted portfolios, several of the measures appear to have little or no ability to find mispriced stocks. Some measures pick portfolios that generate significant negative alphas. After discarding hard-to-borrow stocks, however, the negative alphas either disappear or shrink dramatically. Option-based measures of stock mispricing, for the most part, find market frictions rather than investment opportunities.

It would seem that since all of these measures would contain the same information about future stock returns since all are derived from options. We find though, that correlations of portfolio placements across mispricing measures are low. This suggests that combining measures may produce portfolios with larger future abnormal stock returns. This is true, but only on the short side. The larger negative abnormal returns that are earned on paper by combining measures are, however, earned by portfolios that contain very high proportions of hard-to-borrow stocks. Combining measures does not produce portfolios with larger positive abnormal returns on the long side. For the most part, option-based measures of mispricing find market frictions. In some cases though, they appear to find investment opportunities.

The rest of this paper is organized as follows. Section 3.2 explains why informed investors may trade options rather than shares. Section 3.3 reviews option-based measures of stock mispricing and discusses how the measures are estimated in this paper. In Section 3.4 we estimate returns and alphas for portfolios created by sorting stocks on the basis of option-based measures of stock mispricing. Section 3.5 concludes.

3.2 Literature Review: Informed Trading using Options

Black [1975] and many subsequent papers, note that investors with information may choose to trade options rather than stock. For informed investors, trading options provides two significant advantages over trading shares. First, it is often easier and cheaper to buy put options or sell call options than it is to short-sell shares. Short-selling can be expensive, and it may be difficult to find shares to borrow. In addition, short-sellers face the risk of margin calls if the stock price increases. In those cases, buying puts or writing calls provide attractive alternatives for investors who believe that a stock is overvalued. In addition, options provide leverage and relax the investors borrowing constraints. When buying stock on margin, an investor cannot borrow more than 50% of the stocks cost. A basis of option pricing theory is that a call option can be replicated by borrowing money and buying shares. The implicit borrowing in an option purchase, particularly if the option is out-of-the-money, can be much greater than 50%. The implicit leverage means that for an informed investor, returns are much higher for an option position than a stock position.

The implicit leverage in call options can be seen in the Black-Scholes model. The value of a European call is

$$CALL = \left(S_0 - \sum_{i=1}^{I} D_i e^{-rt_i} \right) N\left(d_1\right) - Ke^{-rT} N\left(d_2\right) \tag{3.2.1}$$

where $d_1 = \frac{\log\left(\frac{S_0 - \sum_{i=1}^{I} D_i e^{-rt_i}}{K}\right) + \left(r + \frac{\sigma^2}{2}\right)T}{\sigma\sqrt{T}}$, $d_2 = d_1 - \sigma\sqrt{T}$, S_0 is the stock price, $\sum_{i=1}^{I} D_i e^{-rt_i}$ is the present value of dividends paid over the life of the option, K is the strike price of the option, r is the risk-free interest rate, σ is the stock volatility, T is the time to expiration, and $N\left(\cdot\right)$ is the standard Normal cumulative distribution function. Note that Black-Scholes option pricing is based on replicating a call option using a portfolio of shares of stock combined with riskless borrowing. In the Black-Scholes formula, $N(d_1)$ is the number of shares in the portfolio, and $\left(S_0 - \sum_{i=1}^{I} D_i e^{-rt_i}\right) N(d_1)$ is the total value of the shares. The second term in the formula is the amount borrowed. Hence, the leverage implicit in a call option, which is the amount borrowed divided by the value of the shares, can be calculated as

$$Leverage = \frac{Ke^{-rT}N(d_2)}{\left(S_0 - \sum_{i=1}^{I} D_i e^{-rt_i}\right) N(d_1)} \tag{3.2.2}$$

We demonstrate the implicit leverage in call options using LiveVol/CBOE intraday prices for all options over our sample period of 2004-2013. To calculate implicit leverages, we use the estimates of implied stock prices and implied volatilities from the analysis. We define in-the-money options as those with an absolute value for their delta of 0.625 to 0.875, at-the-money options as those with absolute values of deltas of 0.375 to 0.625, while out-of-the money options have absolute values of deltas in the range of 0.125 to 0.375.[1]

Panel A of Table B.1.1 shows the mean implicit leverage of in-the-money, at-the-money, and out-of-the-money calls in our sample for various times to expiration. Calls provide much more leverage than can be obtained by buying stocks on margin. For at-the-money calls with less than 30 days to expiration, the mean implicit leverage is 0.8835. That is, buying these calls is like buying shares and borrowing 88.35% of the cost of the shares. This leverage is much greater than the maximum margin of 50% available for buying shares. Put another way, an investor who buys at-the-money calls gets exposure to 1/(1-.8835) = \$8.58 worth of stock for \$1. The investor who purchases stock on margin gets exposure to, at most, \$2 worth of stock for \$1. As Panel A shows, leverage increases as options move further out of the money and as the time to expiration decreases. For example, for at-the-money options the implicit leverage decreases from 88.35% for options with less than 30 days to expiration to 83.1% for options with 30 to 179 days to expiration.[2] Even the options that are in the money and have more than 180 days to expiration do, however, have much greater implicit leverage than the 50% obtained by buying the stock on margin.

For puts, we calculate the proportion of money from the implicit short position that is invested at the riskless rate as collateral. When a stock is shorted, the short-seller has to

[1] This definition of moneyness is also used by Bollen and Whaley [2004] amongst others.

[2] Johnson and So [2012] use $S_0 N(-d_1)/C$ as a measure of leverage. Their measure can be thought of as the multiple of the stock return earned by investing in the option rather than the stock. So, $S_0 N(-d_1)/C = 2$ the option return is twice the share return. Our measure, on the other hand, gives the proportion of the position that is financed by borrowing. We use it because it makes for an easy comparison with margin requirements.

put up 150% of the short-sale proceeds as collateral. In other words, a short-seller needs to use all of the proceeds of the short-sale for collateral and put up an additional 50% of the proceeds out-of-pocket. The Black-Scholes model for puts is given by

$$PUT = Ke^{-rT}N\left(-d_2\right) - \left(S_0 - \sum_{i=1}^{I} D_i e^{-rt_i}\right) N\left(-d_1\right) \tag{3.2.3}$$

The put is priced as a short position in $N\left(-d_1\right)$ shares of the underlying stock and $Ke^{-rT}N\left(-d_2\right)$ in the riskless security. The proportion of proceeds that the investor implicitly puts up for collateral out-of-pocket when he purchases a put is the price paid for the put divided by the $\left(S_0 - \sum_{i=1}^{I} D_i e^{-rt_i}\right) N\left(d_1\right)$ proceeds from the short position. That is,

$$Collateral = \frac{Ke^{-rT}N\left(-d_2\right)}{\left(S_0 - \sum_{i=1}^{I} D_i e^{-rt_i}\right) N\left(-d_1\right)} \tag{3.2.4}$$

Panel B of Table B.1.1 shows the implicit out-of-pocket collateral requirements for puts. For at-the-money options, the average of the implicit collateral requirement is 0.3933 if the option has more than 180 days to expiration, 0.2363 for options with 30 to 180 days to expiration and 0.1416 for options with 10 to 30 days to expiration. Accordingly, buying puts allows investors to have far lower out-of-pocket collateral requirements than the 50% minimum in a short sale of stock.

So, regardless of whether informed investors buy puts or calls, trading options rather than stock of equal value allows investors to earn greater potential returns on their investments. There are, however, additional reasons why option trading is especially attractive for informed investors with unfavorable information about a stock. If a short sellers collateral falls below a maintenance level, he or she must post additional collateral. If the stock price moves against a put buyer, he just has a lower level of collateral. In addition, short sellers have to borrow shares to sell, and it is difficult or expensive to borrow shares of some stocks. Dealers who sell puts to investors may hedge by shorting the underlying stock, but they are typically better able to sell short than most investors. An investor who is able to borrow shares bears the risk that the borrowed shares will be recalled early and he will be forced to terminate the short position.

Figlewski and Webb [1993] provide evidence that options trading reduces the effect of constraints on short selling. Using S&P 500 stocks over 1974-1983, they show that short interest is higher for stocks with options than for other stocks. They attribute this to options market makers shorting to hedge their positions.[3]

This is not to say that all informed trading should take place in the options market. Easley et al. [1998] model trading by informed investors across options and shares. In the pooling equilibrium of their model, informed investors trade in both options and shares. More of the informed trading goes into options as depth in the options market increases and as the relative leverage of options increases. More trading takes place in shares as the depth of the stock market increases.[4]

Evidence suggests that most informed trading takes place in the stock market. Chakravarty et al. [2005] use vector autoregressions of stock prices and implied stock prices from options at one second intervals to estimate the proportion of information impounded in prices by option and stock prices. Across 60 stocks, they find that the information share of options varies from 11.8% to 23.5%. The information share is higher for out-of-the money options than for in-the-money or at-the-money options.

The information share of options increases with the ratio of option volume to stock volume, and decreases with the ratio of option effective spreads to stock effective spreads. Kacperczyk and Pagnotta [2019] provide direct evidence of option trading by informed investors. They hand collect a sample of 5,058 trades in 615 firms that were part of insider trading investigations by the SEC over 1995-2015. Stocks accounted for 67% of these trades and options accounted for 32%. Implied volatilities and volume from options were abnormally high while insiders were trading.

3.3 Option-Based Measures of Stock Mispricing

A number of option-based measures of stock mispricing have been proposed. We are the first to compare their performance under similar conditions and to show when they do and do not succeed in predicting stock returns. We are not interested in overnight returns

[3]Battalio and Schultz [2011] show that the imposition of short-sale bans during the financial crisis led to a sharp increase in bid-ask spreads for options on banned stocks. Synthetic share prices of banned stocks became significantly lower than actual share prices suggesting that dealers who could not hedge discouraged investors from shorting synthetically.

[4]Bid-Ask spreads for options are wide. In percentage terms, as measured by quoted spreads, it is typically much more expensive to trade options than the underlying shares. Muravyev and Pearson [2020] show, however, that trades are more likely to occur at the ask price when the fair value of the option is close to the ask price, and more likely to occur at the bid price when the options fair value is near the bid. Quoted spreads overstate trading costs.

generated by these measures as these returns may be due to microstructure noise. Instead, we focus on monthly returns that can be earned for up to a year after these measures are estimated.

We separate option-based measures of stock mispricing into three categories. The first category is measures based on the difference between implied and actual stocks prices. These measures use an options pricing model (e.g. Black-Scholes) or an arbitrage bound to generate an implied stock price. The second category is measures based on implied volatilities. These measures use an option pricing model to generate implied volatilities and compare them across options or over time. The third category is measures based on trading volume. These measures compare bullish and bearish trading volume, or trading volume across options and the underlying stock.

3.3.1 Measures based on differences between implied and actual stock prices

An early paper that used implied stock prices from options to predict stock returns is Manaster and Rendleman [1982]. Their data consists of daily closing prices for call options from April 26, 1973 June 30, 1976. They omit calls on days when the present value of the dividends to be paid over the life of the option exceeds the present value of interest foregone by early exercise. This removes all call options that could rationally be exercised early from their sample. They calculate implied stock prices each day by finding implied prices and volatilities that minimize the sum of the squared differences between dividend-adjusted Black-Scholes prices and market prices of all call options. Manaster and Rendleman [1982] rank stocks by the percentage difference between implied and actual stock prices, which they label Δ, and sort stocks into Δ quintiles. On average, the difference in the returns the next day for high and low Δ quintiles is over 18 basis points.

We implement the Manaster and Rendleman [1982] method in the following way. We jointly estimate the implied stock price and implied volatility using the extended Black-Scholes-Merton model for a European option on dividend paying stocks. For each option contract (identified as a call or put, by expiration date, and by strike price), over each 30 minute interval, we use an iterative process to solve for the values of $\{S_{IMP}, \sigma_{IMP}\}$ that minimize the sum of the squared differences between option trade prices (or quote midpoints) and Black-Scholes-Merton prices. Specifically, for call options we solve

$$\min_{S_{IMP},\sigma_{IMP}} SSE = \min_{S_{IMP},\sigma_{IMP}} \sum_{t=1}^{N_t} \left(\left(S_{IMP} - \sum_{i=1}^{I} D_i e^{-rt_i} \right) N(d_1) - Ke^{-rT} N(d_2) - C_t^{obs} \right)^2 \tag{3.3.1}$$

where C_t^{obs} is the observed option trade price, S_{IMP} is the estimated implied stock price, σ_{IMP} is the estimated implied volatility. $\{S_{IMP}, \sigma_{IMP}\}$ are calculated for put options in an analogous manner. We require that there be at least 3 different option trade prices within a 30 minute interval in order to uniquely identify the implied stock and volatility. Otherwise the interval is discarded.[5]

We then calculate daily relative implied prices (RIP) using the mean difference between implied prices and stock midpoints in each half hour interval. That is

$$RIP = \frac{\sum_i \frac{S_{IMP,i} - S_{MID,i}}{S_{MID,i}}}{N} \tag{3.3.2}$$

where i denotes the half-hour interval, $S_{IMP,i}$ denotes the implied stock price computed during the i-th half hour interval, $S_{MID,i}$ is the stock quoted midpoint in the end of i-th 30 minute interval, N is the number of 30-minute intervals for which implied prices can be calculated, and RIP is the relative implied price measure based on estimation using option trade prices.

Our measure of the implied price difference (IPD) of stocks, based on option trades and actual stock prices, is obtained by averaging RIP for each day across options on a stock using open interest as weights, and then averaging across all days in a month. IPD is calculated using only 30-minute intervals for which at least three option trade prices are available. Hence, it is only calculated during periods of price discovery. As such, we can think of it as the average percentage difference between option-implied stock prices and actual stock prices during periods of price discovery.

We also calculate RIP and IPD using option bid-ask midpoints at the time of trades rather than option trade prices. IPD calculated from bid-ask midpoints is highly correlated

[5] As a starting point for the iterative process, we use the stock quote midpoint (average of the stock bid and ask) and the option implied volatility (calculated in LiveVol using the Cox et al. [1979] tree method adjusted for dividends) of the last trade of the prior 30 minute interval. When solving for $\{S_{IMP}, \sigma_{IMP}\}$ we perform the iteration over the log starting values to avoid the situation where the implied stock and implied volatility become negative and we impose the condition in the joint estimation of $S_{IMP} - \sum_{i=1}^{I} D_i e^{-rt_i} > 0$ since the log $\left(\frac{S_{IMP} - \sum_{i=1}^{I} D_i e^{-rt_i}}{K}\right)$ results in an error when $S_{IMP} - \sum_{i=1}^{I} D_i e^{-rt_i} < 0$.

with IPD estimated from trades, but IPD estimated from trade prices is a somewhat stronger predictor of stock prices. Hence we focus on that measure in the empirical work to follow.

IPD is based on Manaster and Rendleman [1982], but there are some differences. First, we use option trades for an entire month to estimate IPD. We are interested in prediction of long-term stock returns. This requires that information must be incorporated slowly and hence it is appropriate to use a longer period to estimate IPD. Second, we use intraday trades. This allows us to match option trades with simultaneous stock quotes. It also increases the number of observations used in the estimation. Finally, we use puts as well as calls to estimate implied stock prices.[6]

In calculating IPD, we do not account for early exercise. It is difficult to simultaneously solve for implied volatility, implied stock price, and the early exercise boundary. In addition, IPD is calculated almost entirely from actively traded options. These tend to be at-the-money or out-of-the-money options that are unlikely to be exercised early. When IPD is estimated using only call options on stocks that do not pay dividends, results are similar to those presented here. As we will see, IPD works very well as a predictor of stock returns. It is possible that it would work even better if early exercise is incorporated in its estimation.

Muravyev et al. [2020] derive a related measure of mispricing. They use a non-linear transformation of put-call parity violations to estimate implied stock borrowing fees that short-sellers would expect to pay. This is a significant advantage of their measure it is based on an arbitrage bound rather than a model. They demonstrate that their measure of implied fees predicts future changes in indicative stock borrowing costs. Using data from July 2006 through August 2015, they calculate the median implied borrowing fee for each stock each day. Implied borrowing fees are a highly significant predictor of returns for the next week and for the following month. When they sort stocks into deciles by implied borrowing fees, the difference in four-factor alphas between high and low deciles over the next month is a highly significant 75 basis points. Muravyev et al. [2020] do not claim to estimate implied stock prices. Nevertheless, a put-call parity violation is observationally equivalent to a difference between implied and actual stock prices. There are many potential reasons why implied and actual stock prices can differ, but the implied borrowing fee measure assumes that differences between implied and actual stock prices occur because of short selling costs. They use pairs of puts and calls with the same strike

[6] Muravyev et al. [2013] use a similar technique. They predict returns for 39 stocks using differences between implied and actual stock prices. Their implied bid and ask prices are calculated from put-call parity.

price and time to expiration and attribute violations of put-call parity to borrowing fees. The estimated borrowing fee for a put-call pair is

$$h_{IMP}^{Q} = \frac{1}{\delta}\left(1 - \left(1 - \frac{S_t - C_t + P_t - PV(DIV) - PV(K)}{S_t}\right)^{1/k}\right) \tag{3.3.3}$$

where h_{IMP}^{Q} is the implied borrowing fee, δ is $1/(1+r)$, r is the daily discount factor, S_t is the stock price at time t, C_t is the call price at time t, P_t is the put price at time t, $PV(DIV)$ is the present value of the underlying stocks dividends over the life of the options, $PV(K)$ is the present value of the strike price and k is the time to expiration.

3.3.2 Measures based on option implied volatilities

If the Black-Scholes model is correct, implied volatilities from different options on a stock with the same time to expiration should be the same. Differences in implied volatilities across options indicate that some options have high prices relative to others. Bollen and Whaley [2004] provide evidence that differences in implied volatilities across options and time reflect differences in demand for the options by investors. Information about investor demand for various types of options as reflected in implied volatilities may be useful for predicting stock returns.

Cremers and Weinbaum [2010] (CW) show that differences between implied volatilities of calls and implied volatilities of puts predict stock returns. Using all pairs of calls and puts on a stock with the same strike price and expiration date, and weighting each pair by its open interest, they calculate weighted average differences between call and put implied volatilities. A higher (lower) implied volatility for calls than puts means that the call prices are high (low) relative to the put prices. They divide stocks into five quintiles based on these differences and calculate abnormal returns over the next week and the next four weeks for these portfolios. Portfolios with call implied volatilities that exceed put implied volatilities earn positive abnormal returns while those with larger put implied volatilities earn negative abnormal returns. A strategy of going long the portfolio with the greatest difference between call and put volatilities and short the portfolio with the smallest difference produces four factor alphas of 99 basis points over the following four weeks. Cremers and Weinbaum show further that the predictive power of option prices increases as options become more liquid or underlying shares become less liquid.

In our empirical work, CW is the difference between call and put implied volatilities as in Cremers and Weinbaum [2010]. We calculate CW daily for pairs of puts and calls with the same strike price and expiration date. A daily weighted average difference is calculated across option pairs using the open interest as weights. That is, for day t, $CW = \sum_i w_{i,t} \left(IV_{i,t}^{Call} - IV_{i,t}^{Put} \right)$ where i is the expiration date and strike price combination, the weight $w_{i,t}$ is the average of the put and call open interest, $IV_{i,t}^{Call}$ is the implied volatility of the call and $IV_{i,t}^{Put}$ is the implied volatility of the put. In calculating daily averages, option pairs are omitted if the implied volatility, option delta, open interest or option trade volume are missing, if the best bid or ask quotes are less than or equal to zero or if the ask quote is less than or equal to the bid. We omit options with absolute values of delta greater than 0.98 or less than 0.02. The average daily CW is averaged across days of the month weighting each day equally.

Xing et al. [2010] use skewness, defined as the difference between the implied volatility of out-of-the-money puts and the implied volatility of at-the-money calls, to predict future stock returns. Out-of-the-money puts maximize leverage for investors with negative news or bearish opinions about a stock. Strong demand for out-of-the-money puts will push up their prices and thus their implied volatilities. At-the-money calls are typically the most liquid options, hence subtracting the implied volatility of at-the-money-calls from the implied volatility of out-of-the-money puts provides a measure of the excess demand for puts. Using closing prices from 1996 2005, Xing et al. [2010] calculate weekly skewness for individual stocks by averaging daily skewness. They sort stocks into quintiles based on skewness and show that stocks with high skewness (large implied volatilities for out-of-the-money puts) underperform stocks with low skewness (small implied volatilities for out-of-the-money puts) by 15 to 20 basis points over the next week. Stocks with low skewness continue to outperform stocks with high skewness for up to six months after the portfolio formation.

An implicit assumption in the skewness measure is that informed trading of options takes place mainly by investors with negative information. This may be true if options are a way to get around short-sale constraints. Other measures based on option volume also assume that informed trading of options comes mostly from bearish investors. We estimate skewness as the difference between implied volatilities of puts with a delta of -0.2, and the average implied volatility from put and call contracts with an absolute value of delta of 0.5. Implied volatilities are obtained from the OptionMetrics volatility surface for options with 30 days to maturity. We calculate skewness daily and average daily values to compute a skewness measure for each month.

An et al. [2010] examine the power of changes in implied volatilities to forecast stock returns. Implied volatilities are obtained from the daily implied volatility surface calculated by OptionMetrics. Their empirical analysis uses end-of-month call and put implied volatilities for options with a delta of 0.5 and 30 days to maturity. An et al. sort stocks into decile portfolios each month based on the change in the stocks call and put implied volatilities. Larger increases in call volatilities are associated with larger stock returns the next month while larger increases in put volatilities are associated with lower stock returns in the following month. The difference in returns between the portfolio with the largest (typically positive) change in implied call volatility and the portfolio with the smallest (typically negative) change in implied volatility is about 1% over the next month. Differences in the next months abnormal returns, calculated with either the CAPM or the Fama-French three factor model, are also about 1%. Differences in returns and abnormal returns across portfolios of stock with the largest and smallest changes in put implied volatilities are about 0.5%.

We replicate the An et al. [2010] by calculating $\Delta CVOL$, the monthly change in implied volatilities of calls and $\Delta PVOL$, the month change in implied volatilities of puts. The data are from the OptionMetrics volatility surface. We remove observations with missing implied volatilities and deltas. Only at-the-money series ($|\Delta| = 0.5$) with 30 days to maturity are retained. We use the last available daily observation of the month for each call/put for each stock to compute the monthly changes.

3.3.3 Measures based on option volume and order flow

Johnson and So [2012] observe that, because of short-sale restrictions, investors with negative information are particularly likely to choose to trade options rather than shares.[7] They propose that O/S, the natural logarithm of the ratio of options volume to stock volume, may contain information about future stock returns. A high value of O/S is likely to reflect a large volume of trading by pessimistic investors who trade options rather than shares because of short-sale restrictions. Using data from 1996 2010, they calculate O/S using weekly volume shares of put and call options that expire between five and 35 days after the trade. They form decile portfolios based on O/S ratios and show that the four-factor alpha of the lowest O/S decile portfolio is about 34 basis points greater than the alpha of the highest decile portfolio for the week following portfolio formation.

Ge et al. [2016] confirm that high O/S ratios predict negative stock returns, but also provide evidence that implicit leverage, not short-sale restrictions, is behind the measures power to predict stock returns. They break down volume for puts and calls into buy volume and sell volume, and into trades that open and close positions. Both bullish and bearish volume predict stock returns with the strongest predictions coming from volume that opens call positions. Ge et al. [2016] state that high values of O/S are associated with negative

[7] Roll et al. [2010] originally found the option to stock volume ratio to predict stock returns.

future returns because *more components of options volume negatively predict returns than positively predict returns, due to the fact that trading volume stemming from the unwinding of bought call positions negatively predicts returns.*

We estimate the O/S measure used in Johnson and So [2012], and Ge et al. [2016] as the natural logarithm of the option to stock volume ratio. Unlike Johnson and So [2012], who use just short-term options, we use the total option volume across all strikes and maturities. We calculate O/S monthly. We measure stock volume in round lots of 100 to make it comparable to option contracts on 100 shares.

Pan and Poteshman [2006] (PP) show that the daily ratio of the number of put contracts purchased to the sum of put and call contracts provides information on stock mis-pricing. A large ratio, reflecting greater public purchases of puts than calls, implies negative private information while a small ratio, reflecting greater public purchases of calls than puts, implies that traders have bullish private information. Slope coefficient estimates obtained by regressing the next-day four-factor adjusted stock return on the ratio indicate that buying stocks with all volume coming from buys of calls and selling stocks with all volume coming from put buys yields a highly significant average return of over 50 basis points over the next day. Larger excess returns are produced when they use options that are out-of-the-money or close to expiration as these options provide more leverage. Pan and Poteshman [2006] note that their results are not incompatible with market efficiency. In their tests, returns are predicted using information that is not available to the public. Investors are not able to observe whether option trades, and thus option volume, is buyer initiated. We denote the put-to-call volume ratio as calculated by Pan and Poteshman [2006] as PP and calculate it this way

$$PP = \frac{OpenBuyPut}{OpenBuyPut + OpenBuyCall} \tag{3.3.4}$$

where *Open Buy Put* is the volume from purchases of puts that open put positions for customers while *Open Buy Call* is the volume from purchases that open call positions for customers. Open Buy and Sell trading volumes are obtained from the CBOE/ISE exchanges. We calculate PP daily and average the daily measure over the month.

Hu [2014] calculates the Options Order Imbalance (OOI) daily by summing the product of the volume for each option trade times the option trades delta. If most of the volume occurs in option trades with negative deltas (put purchases or call sales) OOI will be negative. If most volume is from positive alpha option trades (call purchases or put sales) OOI will be positive. Options order imbalance predicts stock returns the next day even after adjusting for order imbalances in the underlying stock. The return associated with options order imbalance does not appear to be reversed in succeeding days. Hu [2014] observes that options order flow contains significant information about stock values. We compute OOI as the difference between the synthetic positive and negative exposure to the underlying stock using the signed CBOE/ISE option volume weighted by the absolute value of the options delta and scaled by total option volume. OOI is calculated daily, averaged monthly, and is used to predict stock returns in future months.

3.4 Predicting Stock Returns

3.4.1 Data

In the empirical work that follows we use two sources of options data to estimate option-based measures of stock mis-pricing. CBOE/LiveVol provides intraday option trades and quotes for 2004-2013. We apply the following standard microstructure data filters to our option trades data. We remove option trades that occurred before 9:30 am or after 4:00 pm. We discard cancelled trades, trades in which the option trade price is greater than twice the contemporaneous option quote midpoint, and trades with missing prices or missing bid or ask quotes. Option trades are also deleted if contemporaneous option quote midpoints are less than ten cents, the options have zero trade volume, or have ten or fewer days to maturity. For inclusion, all option best bid and best ask quotes must satisfy the relationship $0 <$ Best Bid $<$ Best Ask $< 5\times$ Best Bid. Intraday data from CBOE/LiveVol is used to estimate IPD. Signed CBOE/ISE volume data are used to estimate PP, and OOI. Our second source of options data is OptionMetrics. OptionMetrics provides daily summary statistics for individual options. It also estimates an end-of-day implied volatility surface. We use data from OptionMetrics to estimate $\Delta CVOL$, $\Delta PVOL$, skewness, implicit borrowing fees, CW, and O/S.

Table B.1.2 provides summary statistics on trading volume in the CBOE/LiveVol data. The median number of contracts with 10 to 30 days to expiration that trade in one month is 226. That is a daily average of 10 contracts for 100 shares each. The distribution of trading volume is right-skewed with some options trading a lot and others very little. Options with greater times to maturity usually have lower volume, but still trade actively. For options with more than 90 days to expiration, the mean number of contracts traded per month is 288 and the median is 142. Some measures, like skewness, are estimated from heavily traded short-term options. Others, like IPD are estimated from several different options on the same stock. Options on the 500 largest stocks, which are especially important in calculating returns of value-weighted portfolios, trade more frequently than options on other stocks.

Table B.1.3 reports summary statistics for option-based measures of stock mis-pricing. Measures like IPD, CW, or OOI are typically small on average, as we would expect them to be. There are outliers in IPD, which could occur if the stock quote midpoint was mis-estimated. Likewise, O/S has outliers when stock volume is very low. We discard obvious errors, but in addition, our comparisons of returns and alphas across quintile portfolios mitigates the influence of outliers.

In this paper we show that stocks that are identified as overpriced by option-based measures are often hard to borrow to short. We use indicative fee data from Markit to identify hard-to-borrow stocks. This indicative fee reflects buy side end-users demand. Table B.1.4 summarizes the distribution of indicative fees by year for the 2004-2013 sample period. The median fee varies from 33.8 basis points per year in 2009 to 48.2 basis points in 2004. In each year, the difference between the 1st percentile and the median indicative fee is small, typically 5-7 basis points. We define hard-to-borrow stocks as those with indicative fees that are in the top 20%. The 80th percentile indicative fee is usually two to four times as large as the median fee. For example, in 2013 the median fee is 39.2 basis points while the 80th percentile fee is 1.72%. The 99th percentile of indicative fees are much higher than the median fees, and exceed 40% in 2011 and 2012.

3.4.2 Performance of Predicting Stock Returns

The nine option-based measures of mis-pricing used here have been shown to predict abnormal stock returns. This could be because informed investors trade options and the stock market incorporates the information from options slowly. In this case, it should be possible to earn abnormal returns in practice with the aid of these measures. On the other hand, it may be that informed traders trade options in part because of market frictions in the market for the underlying stock, and that these frictions prevent investors from profiting from mispricings in practice. In this section, we attempt to minimize the impact of market frictions and see if option-based measures of mis-pricing still yield abnormal returns.

We test whether abnormal returns can be earned using value-weighted portfolios sorted on option-based measures of mispricing. We test whether abnormal returns can be earned using value-weighted portfolios sorted on option-based measures of mispricing. Fama and French [2008] caution against using long-short returns from equal-weighted portfolios to examine anomalies. They note that the cross-sectional dispersion of anomaly variables is largest among very small firms. Hence, small, illiquid stocks that are expensive to trade dominate long-short returns from equal-weighted portfolios. Fama and French observe that a similar problem arises stock returns are regressed on anomaly variables. The extremes of the explanatory variables are likely to be small and illiquid firms. We would add that small stocks are also more likely to be difficult to borrow and sell short. Strategies that use value-weighted portfolios are more likely to be implementable in practice.

We are also concerned that the profiting from option-based measures of mispricing requires investors to short stocks that are costly or difficult to short in practice. The profitability of many anomalies that seem to promise abnormal returns depends on short-selling hard-to-borrow stocks. Stambaugh et al. [2015] construct value-weighted decile portfolios of stocks sorted on 11 anomaly variables over August 1965 through January 2008. Long-short portfolios produce statistically significant three-factor abnormal returns for each anomaly variable. For ten of the 11, the absolute value of the abnormal returns is larger on the short leg than long leg. The differences in abnormal returns are usually large. Stambaugh et al. [2015] show further that the poor (good) performance of overpriced (underpriced) stocks is especially strong for stocks with high idiosyncratic volatilities and hence greater arbitrage risk.

Jacobs [2015] examines the returns to 100 anomalies over August 1965 through January 2011. These anomalies are grouped into 19 meta anomalies like earnings surprise anomalies and long-term reversal anomalies. Jacobs [2015] notes (page 80) that most meta anomaly returns are effectively driven by the short-leg.

Using value-weighted portfolios reduces but does not eliminate the influence of hard-to-borrow stocks on the returns from strategies based on of options. We also examine the returns to option-based measures of mispricing after eliminating hard-to-borrow stocks. Finally, we examine the returns to long-only strategies that use value-weighted portfolios. The returns to these strategies are not diminished by the high trading costs of small stocks or by the costs of borrowing shares for short-selling.

Each month, we sort stocks into quintile portfolios by each of the following option-based mispricing measures: IPD, Implied Lending Fees, CW, Skewness, $\Delta CVOL$, $\Delta PVOL$, OOI, PP, and O/S. Each month, we calculate the average market capitalization of stocks in each quintile portfolio for each option-based mispricing measure. We then calculate grand averages across months. These average sizes, in billions of dollars are shown in Panel A of Table B.1.5.

For all option-based measures except Pan-Poteshman, both the low and high quintile portfolios tend to have smaller firms than the middle portfolios. For example, the average capitalization of low IPD stocks is just \$2.75 billion, while the average size of third quintile IPD stocks is \$13.82 billion. Similarly, the mean size of the stocks in the high implied fee portfolio is only \$1.7 billion, while the mean capitalization of stocks in the third implied fee quintile is \$10.16 billion. Likewise, the mean firm size for stocks in the high skewness portfolio is just \$2.65 billion while the mean size is \$8.36 billion for third skewness portfolio and \$13.58 billion for the second lowest skewness portfolio. Profiting from these option-based measures of stock mispricing involves trading small firms. Small stocks are illiquid and expensive to trade. It is difficult to acquire a significant number of shares without moving the price. The concentration of small stocks in extreme portfolios is an even bigger problem if decile portfolios are used.

Not only are small firm stocks more expensive to trade than large firm stocks, they are also more likely to be difficult to borrow for short sales. We obtain indicative lending stock lending fees for all stocks for each month in our sample period from Markit. We designate a stock as hard to borrow if the indicative fee for that stock is among the highest 20% of all stocks, not just stocks with options, during that month. Optionable stocks are larger and more liquid than those without options, so we would expect fewer than 20% of stocks with options to be hard-to-borrow. For each quintile portfolio of each mispricing measure, we calculate the proportion of stocks in the portfolio that are hard to borrow each month, and average the proportion over the months of our sample period. Table B.1.5 Panel B reports these averages.

For several measures, the quintile portfolio containing overvalued stocks is overweighted in stocks that are hard to borrow. The low IPD portfolio contains stocks in which the implied stock price is low relative to the actual stock price. On average, 33.88% of the stocks in that portfolio are hard to borrow. In comparison, only 10.74% of the high IPD stocks are hard to borrow. Similarly, 30.72% of the portfolio with high implied lending fees is composed of hard to borrow stocks. Only 5.56% of the stocks in the low borrowing fee portfolio are hard-to-borrow. For the low CW portfolio, which consists of stocks with high implied volatilities from puts relative to the implied volatilities from calls, 37.07% of stocks are hard to borrow. In contrast, only 12.34% of the stocks in the high CW portfolio are hard to borrow. The proportion of stocks that are hard to borrow increases monotonically from 6.69% in the low O/S portfolio to 21.13% in the high O/S portfolios.

Table B.1.5 Panel B demonstrates that option-based measures of stock mispricing pick out small and hard-to-borrow stocks. That is, they find stocks with market frictions that may prevent investors from profiting from mispricing. This may be because informed investors trade the options because frictions like high borrowing costs make it difficult to trade shares. Or, it may be that these stocks adjust more slowly to information in options.

In the appendix, we present returns and four factor alphas for equal-weighted quintile portfolios formed using each of the option-based measures of mispricing. These measures appear to produce large long-short returns, but the significant mispricing appears to be almost entirely on the short-side. As Table B.1.5 Panel B shows, the quintile portfolios that these measures indicate are overpriced are overweighted with small and hard-to-borrow stocks. Option-based measures of mispricing are very good at finding market frictions.

We next see if option-based measures of stock mispricing can be used to generate abnormal returns after minimizing the influence of small, hard-to-borrow stocks. We calculate returns and four-factor alphas for each of the three months following portfolio formation for value-weighted quintile portfolios. Value-weighting minimizes the contribution of small, illiquid stocks to the portfolio returns. Because hard-to-borrow stocks are often small, value-weighting also minimizes the impact on returns of stocks that are difficult to short. As we will see, some option-based measures of mispricing that work with equal-weighted portfolios have no ability to predict returns of value-weighted portfolios. Table B.1.6 presents time-series averages of monthly returns and Fama-French-Carhart four factor alphas (the original Fama and French [1993] augmented with the Cahart [1997] momentum factor) for portfolios. Panel A reports results for measures based on differences between implied and actual stock prices. Long-short portfolios formed both on IPD and implied borrowing fees produce significant abnormal returns in the three months following portfolio formation. For IPD, for example, the long-short portfolio produces alphas of 63.64 basis points, 79.36 basis points, and 83.92 basis points in the three months following portfolio formation. Both measures produce significant abnormal returns on the short side. For the high implied borrowing fee portfolio, in Panel B, alphas are -38.73 basis points, -30.61 basis points, and -38.29 basis points in the three months following portfolio formation. More interesting is that the high IPD portfolio, that is the one in which implied prices are highest relative to actual prices, earns positive abnormal returns of 26, 39, and 43 basis points over months t+1 through t+3. Each of these alphas is significantly different from zero at the 1% level. Short sales are not needed to earn the abnormal returns of the high IPD portfolios. Because these portfolios are value-weighted, trading costs and price impact should be minimal.

Panel C to F presents results for value-weighted quintile portfolios formed using measures based on implied volatility. In general, these results are weak. Portfolios based on CW, for example, do not provide significant long-short returns or significant long or short-side abnormal returns. In the appendix, where portfolios are equal-weighted rather than value-weighted, CW, and other measures produce large and significant returns for short portfolios. Nevertheless, in value-weighted portfolios, as shown in Panel D, skewness provides significant long-short returns in months two and three after portfolio formation. It also provides significant positive long-side alphas in month t+3, and, as we will show, in later months as well. $\Delta PVOL$ fares best of the measures, and provides significant long-short abnormal returns as well as significant negative abnormal returns for short portfolios.

Panel G, H, and I reports returns and four-factor alphas for value-weighted portfolios formed using measures based on option volume. Sorts based on PP do not produce any statistically significant four-factor alphas. Sorts on OOI yield a statistically significant negative four-factor alpha for the second month and the second month only. On the other hand, results for value-weighted portfolios based on O/S are quite strong. Alphas are statistically significant for the short side (high options to stock volume) each month. In addition, sorts on O/S, like IPD, produce statistically significant alphas each month on the long side. Alphas for the low O/S portfolio are 25.60 basis points, 26.75 basis points, and 26.88 basis points for the three months after portfolio formation.

The results in Table B.1.6 show that at least three option-based measures seem to predict positive alphas for value-weighted portfolios. IPD and O/S based portfolios produce positive alphas in each of the three months after portfolio formation. The low skewness portfolio has positive alphas in each of the three months following formation, but it is only in the third month that the alpha is significantly different from zero. We will show, however, that the low skewness portfolio continues to earn abnormal returns in succeeding months. These three measures then generate trading strategies that produce significant alphas but do not involve shorting stocks and do not rely on trading small firms. Cumulative positive alphas over the three months following portfolio formation are significant but modest. For high IPD portfolios they are just over 1% and for low O/S portfolios they are about 80 basis points.

Portfolios formed using four of the option-based measures of mispricing, IPD, implied fees, $\Delta PVOL$, and O/S, have negative and statistically significant abnormal returns after portfolio formation. Using value-weighted portfolios as we do in Table B.1.6 reduces the impact of short-sale constraints, but they may still be a factor. To see if the negative abnormal returns can be earned by shorting stocks, each month we sort stocks into value-weighted quintile portfolios based on each of the nine option-based measures of mispricing after taking out all stocks that are hard-to-borrow that month. The remaining stocks all have indicative borrowing fees below the 80th percentile that month. We then calculate returns and four-factor alphas for each of these portfolios over each of the next three months. We report the alphas in Table B.1.7 alongside the alphas from Table B.1.6 for portfolios that include the hard-to-borrow stocks. To save space, we report the alphas for just the high and low portfolios and the long-short portfolio.

Panel A provides results for portfolios formed using measures based on the difference between implied and actual stock prices. When we exclude hard-to-borrow stocks, the alphas of the high implied fee portfolios shrink dramatically. For example, the alpha for the high implied fee portfolio for month t+1 is -0.3873 with a t-statistic of -2.65 when all stocks are included, but just -0.0499 with a t-statistic of -0.33 when we exclude hard-to-borrow stocks. The alpha for month t+2 for portfolios with high implied fees is -0.3061 with a t-statistic of -2.27 when all stocks are included, but 0.0173 with a t-statistic of 0.13 when we exclude hard-to-borrow stocks. The implied fee variable is supposed to identify hard-to-borrow stocks, so it is not surprising that it loses power when we omit hard-to-borrow stocks. Results for IPD are also weaker when we omit hard-to-borrow stocks. When all stocks are included, the alpha of the low IPD portfolio is -0.3810 with a t-statistic of -2.33. When we exclude hard-to-borrow stocks, the alpha falls to -0.0597 with a t-statistic of -0.33.

Panels C to F and G to I of Table B.1.7 present results for measures based on implied volatilities and on option volume. In general, regardless of the measure used, alphas and t-statistics fall for quintile portfolios of overvalued stocks when we exclude hard-to-borrow stocks. For O/S in particular, alphas move much closer to zero when hard-to-borrow stocks are omitted. For the high O/S portfolio with all stocks included, alphas are -0.1854 with a t-statistic of -4.37, -0.2005 with a t-statistic of -2.41, and -0.2311 with a t-statistic of -3.94 for the three months after portfolio formation. When the hard-to-borrow stocks are omitted, the alphas shrink to -0.0662 with a t-statistic of -1.55 for the first month, -0.0714 with a t-statistic of -1.81 for the second month, and -0.1128 with a t-statistic of -2.37 for the third month. Most of the poor performance of high O/S stocks is explained by the fact that a high O/S measure is an indication that the stock is hard to borrow.

To summarize, when stocks are value-weighted, some option-based measures of mispricing are unable to produce portfolios with significant negative alphas. Using value-weighed portfolios but dropping hard-to-borrow stocks eliminates abnormal returns to short-selling based on IPD or implied fees. There still appears to be some returns to shorting based on $\Delta PVOL$ and O/S, but returns are greatly diminished when hard-to-borrow stocks are eliminated.

Results in Table B.1.6 showed that three measures, IPD, skewness, and O/S, produced portfolios that earned positive and significant abnormal returns in the three months after portfolio formation. These abnormal returns are particularly perplexing. The portfolios are value-weighted, so tiny illiquid stocks are not driving the results. A long only portfolio strategy appears to earn these abnormal returns, so no short-selling is required. The returns over the first three months are, however, relatively small. This raises the question of whether these portfolios continue to earn positive abnormal returns after the three months. In Table B.1.8, we report alphas for high and low IPD, skewness, and O/S portfolios for each of the 12 months following portfolio formation.

Table B.1.8 reports four-factor alphas for the 12 months after portfolio formation for portfolios based on IPD, skewness and O/S. In each case, high and low quintile alphas are reported along with the alphas of the long-short portfolio. What is of particular interest is that these measures appear to produce portfolios with positive and significant alphas for several months after portfolio formation. The first column of Table B.1.8 shows that the high IPD portfolio earns positive and statistically significant abnormal returns for each month from t+1 through t+6. They continue to earn mostly positive but mostly insignificant returns for months t+7 through t+12. This suggests that positive alphas can be earned for several months after portfolio formation through a long-only strategy. Long only portfolios provide cumulative alphas of 3.10% over months t+1 through t+12. The low skewness portfolio earns positive alphas for months t+1 through t+11, and the alphas are statistically significant for months t+3 through t+6. The cumulative return for the twelve months, obtained by summing the individual month alphas is 2.67% for the low skewness portfolio. Low O/S portfolios have positive four-factor alphas in each of the 12 months following portfolio formation. They are statistically significant for months t+1 through t+9. After 12 months, the cumulative alpha for low the O/S portfolio is 2.64%.

These positive alphas appear to be abnormal returns that investors could earn. There are no short selling costs or constraints to with which to contend. Institutional investors who are constrained to long-only strategies should be able to earn these returns. The portfolios are value-weighted, so a strategy to exploit these alphas does not require buying tiny stocks.

3.4.3 Arbitrage Risk

Option-based measures of mispricing may allow investors to earn abnormal returns, but it may be risky to earn those abnormal returns. Investing in stocks that are indicated to be underpriced by IPD, O/S, or skewness may lead investors to hold undiversified portfolios for long periods of time. Each of the quintile portfolios holds a large number of stocks, but it is possible that portfolios are overweighted in some industries or in stocks with common characteristics. A way to see if exploiting these measures leads investors to take extra risks is to compare the Sharpe ratios of these portfolios of stocks that are identified as underpriced with the Sharpe ratio of the market portfolio. The Sharpe ratio measures the return to total risk ratio for an investors holdings. We estimate it for the high IPD portfolio by calculating the excess return of the portfolio (return minus the riskfree rate) in the month after formation for each month over 2004-2013. The Sharpe ratio for IPD for the first month after portfolio formation is the ratio of the time-series average of the first month excess returns to the standard deviation of the first month excess returns. We also calculate Sharpe rations for the high IPD portfolio for each of months 2 through 12 after portfolio formation. Analogous Sharpe ratios are calculated for the low skewness and low O/S portfolios. Table B.1.10 reports the Sharpe ratios.

For IPD, the Sharpe ratio for the month after portfolio formation is 0.18, and it exceeds 0.2 for months t+2 through t+6[8]. It is somewhat lower over the next six months but exceeds 0.15 in all but one month. For comparison, the Sharpe ratio for the S&P 500 was 0.1219 over the sample period, while the Sharpe ratio for the CRSP value-weighted index was 0.1236. An investor who invested only in the in the high IPD portfolio had a larger ratio of return to risk than an investor who held the market portfolio. The last two columns of Table B.1.10 provide Sharpe ratios for the low O/S and low skewness portfolios. These Sharpe ratios are a little smaller than the Sharpe ratios for IPD, but still much higher than the Sharpe ratio of the CRSP value-weighted index. It appears that the mispricing indicated by O/S and skewness could also be exploited without taking extra risk.

3.4.4 Combining Measures

We have used several option-based measures of stock mispricing to form portfolios that produce positive or negative abnormal returns in the subsequent months. Even though all of the measures are derived from options, it is not clear whether the measures are based on the same information. If different option-based measures contain different information, they will produce stock sortings that have low correlations with each other. If this is true, combinations of option-based measures of stock mispricing can be used to produce larger abnormal returns. We are particularly interested in whether we can combine measures to produce portfolios that earn larger positive abnormal returns for long positions.

Table B.1.9 reports correlations of quintile sortings of stocks produced by the option-based measures. There are some high correlations. Quintile portfolio placements by $\Delta CVOL$ and $\Delta PVOL$ have a correlation of 0.570, while the correlation between placements by CW and implied lending fees is -0.500. For the most part though, correlations are low. For example, the correlation between IPD and O/S is -0.057. The low absolute values of correlations suggest that double sorts on two measures may produce portfolios that will earn larger abnormal returns.

There are 36 possible pairs of our nine option-based measures. We select three pairs of measures to use in double sorts: IPD and O/S, IPD and skewness, and implied lending fees and skewness. Each of these measures does a good job of identifying mispricings by itself,

[8]We do not claim that these ratios are independent.

and the correlations of quintile portfolio placements between pairs of measures are low. For each pair of measures, we first sort stocks into quintiles by the first measure (IPD, IPD, and implied fees) and then sort each quintile into five portfolios by the second measure (O/S, skewness, and skewness). Each double sort produce 25 portfolios with equal numbers of stocks. We then calculate the value-weighted return and Fama-French-Carhart four-factor abnormal return for each portfolio for the month after portfolio formation. Table B.1.11 reports the abnormal returns for these double-sorted portfolios.

Value-weighted portfolios formed from stocks with two bearish option measures, low IPD and high O/S, low IPD and high skewness, or high implied fees and high skewness, have large negative alphas. For example, the value-weighted portfolio formed from the combination of low IPD and high O/S earns a average abnormal return of -96.31 basis points in the month after portfolio formation. This is far larger in absolute value than the abnormal return earned by the low quintile IPD portfolio (-38.10 basis points) or the high quintile O/S portfolio (-18.54 basis points). Likewise, the portfolio formed from the combination of high skewness and high implied fee stocks earns an abnormal return of -65.80 basis points in the month after portfolio formation, which is larger than the abnormal return earned by the implied fee quintile portfolio or the high skewness portfolio. Double sorts can produce portfolios that underperform by more than portfolios produced by sorting on one measure at least before shorting costs. The double-sorts do not, however, seem to produce portfolios with larger positive alphas in the month after portfolio formation. Table B.1.11 shows that the alpha for the portfolio of low O/S and high IPD stocks is only 22 basis points. The alpha for the portfolio of stocks with low skewness and low implied fees is positive but less than one basis point. Double sorts do not improve on any positive alphas earned by value-weighted portfolios formed using just one measure.

Table B.1.12 reports the proportion of stocks that are hard-to-borrow in each portfolio produced by the double sorts in Table B.1.11. As might be expected, the double-sort portfolios with the largest negative alphas contain a large fraction of stocks that are hard-to-borrow. The portfolio formed from stocks in the low IPD high O/S quintiles earns an alpha of -96.31 basis points, but Table B.1.12 shows that 65.3% of the stocks in that portfolio are hard-to-borrow. Similarly, the high implied fee high skewness portfolio earns an abnormal return of -65.80 basis points in the month after portfolio formation, but 31% of the stocks in that portfolio are hard-to-borrow.

3.5 Concluding Comments

Several recent empirical papers show that information in options can predict stock returns. On the one hand, it is not surprising that informed investors would choose to trade options rather than shares. Options provide greater leverage than can be obtained with shares and can be used to make bearish bets when it is difficult to sell short. On the other hand, it is somewhat surprising that stock prices can lag options by several months. Information from options is readily available and, in the absence of market frictions, should be incorporated in stock prices quickly.

In this paper, we examine the profitability of stock trading using three categories of option-based measures of stock mis-pricing: measures based on stock prices implied by options, measures based on implied volatilities, and measures based on option trading volume. When we minimize the influence of tiny stocks is by using value-weighted portfolios, some of these measures appear to have no predictive ability. Some identify portfolios of overvalued stocks that earn negative alphas in the months after portfolio formation, but most of this predictive ability disappears when we exclude hard-to-borrow stocks from the portfolios. Option-based measures of stock mis-pricing do find stocks that adjust slowly to information because of market frictions.

Three option-based measures of mis-pricing, IPD, the difference between implied and actual stock prices, O/S, the logarithm of the ratio of option volume to stock volume, and skewness, the difference between the implied volatilities of out-of-the-money puts and at-the-money calls, identify undervalued stocks that produce positive abnormal returns in value-weighted portfolios. These appear to be returns that investors can actually earn. The portfolios are value-weighted, so the strategies do not involve trading tiny, illiquid stocks. In addition, short selling is not required. When IPD is used to identify undervalued stocks, cumulative four-factor alphas for value-weighted portfolios are about 3.1% after 12 months. When O/S is used to find underpriced stocks, four-factor alphas cumulate to about 2.6% in one year. The low skewness portfolio has a four-factor alpha of 2.7% for 12 months. Arbitrage risk does not appear to be a significant impediment to exploiting the mis-valuation implied by these option-based measures. Sharpe ratio measures for long portfolios formed on each of these measures exceed the Sharpe ratio of the market portfolio over the same time period.

The portfolio placements of stocks by different option-based measures have relatively low correlations. This suggests that double sorts on different option-based measures of mis-pricing may produce larger abnormal returns than are produced by single sorts. This is true, but only on the short side. Double sorts do not generate larger long-side alphas than do single sorts.

Option-based measures of stock mis-pricing are good at identifying frictions that slow the response of stock prices to information in options. The mis-priced stocks are often small or hard-to-borrow. Some of these measures, however, seem to predict realizable if modest positive abnormal returns.

Chapter 4

Accounting Transparency and the Implied Volatility Skew

4.1 Introduction

The volatility smile in the options market refers to the fact that the Black and Scholes [1973] implied volatility as a function of the strike price resembles the shape of a skew. A myriad of more complex models, characterized by jumps and time-varying volatilities, have since been developed to identify the *best* model that fits various aspects of the time-series and cross-sectional properties of option prices.[1] Through these studies we have learned that a sizeable stock price jump risk premium is necessary to fit the time-series and cross-sectional properties of option prices which highlights the importance of stock price jump risk in option prices.[2]

In this paper, we study an economically intuitive explanation for the jump risk embedded in the implied volatility skew in individual stock options. We model the equity option price where the firm has a capital structure of debt and equity where investors do not know

[1] The literature began with deterministic models of volatility, giving rise to the so-called local volatility models (see Cox and Ross [1976], Derman and Kani [1994], Dupire [1994], and Rubinstein [1994]), or stochastic, as in the case of Heston [1993] as well as including stock price jumps (Merton [1976]), stochastic jumps in volatility (Bates [1996]), and stochastic jumps in volatility and stochastic jumps in the stock price process (Duffie et al. [2000]).

[2] The magnitude of the variance risk premia, stock jump risk premia, and volatility jump risk premia has been well studied on the stock market index options as well as for individual firms options see for instance: Eraker et al. [2003], Broadie et al. [2007], Broadie et al. [2009], Christoffersen et al. [2012], Andersen et al. [2015] and others.

the precise value of the firm, however, the firm periodically issues noisy accounting reports. This effect of corporate disclosure quality on equity option pricing allows for the positive probability per unit time that the firm can instantaneously go bankrupt which introduces jump risk into the model, as equity value can suddenly go to zero from any positive level. We study the effect of this jump risk on the volatility skew by extending the Duffie and Lando [2001] (DL) model to the pricing of stock options in a model that is solved to closed form solution up to Bivariate Cumulative Standard Normal Distribution.

Since large jumps in stock returns are closely tied to news about future cash flows and discount rates, one might speculate that the volatility smile of stock options may have something to do with the quality of corporate disclosures (Maheu and McCurdy [2004]). Intuitively, a firm with timely, clear, and detailed disclosures will seldomly impose a *surprise* on the market. In contrast, a firm that makes only infrequent and sketchy disclosures is likely to catch investors off guard. In this respect, the quality of corporate disclosures is perhaps proxying for the overall level of jump risk (both likelihood and magnitude of jumps) in stock returns.

Our empirical analysis suggests that both the leverage ratio and the quality of corporate disclosure can explain the cross-sectional variation of the implied volatility skew. Specifically, we highlight an important interaction between the two effects. When a firm is perceived to be opaque, whether its leverage ratio is high or low has less of an impact for the pricing of stock options. Similarly, when the firm reports higher level of accounting transparency, the higher the leverage ratio the more skewed in the implied volatility smile is i.e. the closeness to default (in terms of higher leverage ratio) is better reflected in the options market with more expensive out-of the money options.

Our paper is the first, to the best of our knowledge, to suggest and document a relationship between the implied volatility skew, accounting transparency measure, and leverage. As such it offers an interesting support of the DL insight and also furthers our understanding of the economic sources of the volatility skew. However, our approach is not without its limitations. For example, to emphasize the role of firm-level accounting noise, we have avoided modeling economy-wide jump risks that may affect all of the stock returns. These risk factors are responsible for index option skew and also partly for the skew in individual stock options through their effect on the pricing kernel. Therefore, our model is not expected to fully account for the magnitude of the skew. Our empirical results can also be consistent with the heterogeneous beliefs model of Buraschi and Jiltsov [2006]. From a Bayesian point of view, the reason why different investors have different beliefs about the fundamentals of a firm is that they are updating with different information. In the absence of truthful reporting, any news about a company is necessarily noise, and the difference in beliefs will be high. The two theories therefore offer very similar empirical predictions.

We contribute to the research agenda that focuses more on the economic reasons for the departure from Black-Scholes. Our paper compliments existing models of option models, Chen et al. [2020] Gamba and Saretto [2020], that use firm investment and other firm structural components in the implied volatility skew.[3]

The rest of this paper is organized as follows. We conduct a brief literature review in Section 4.2 and then in Section 4.3 we present our theoretical model, and in Section 4.4 use numerical examples to illustrate the relationship between accounting transparency and the volatility smile. In Section 4.5 we document the construction of the volatility smile, the accounting transparency measures, and other aspects of the data. In Section 4.6 we conduct regression tests of the main predictions of the model with robustness tests in Section 4.7. We conclude with Section 4.8.

4.2 Literature Review

There has been ample recent research that studies the the risk management behavior of option market makers. For example Bollen and Whaley [2004] find that the supply and demand of options can account for the difference in volatility skew across index and stock option markets (for a theoretical model of option demand pressure on option prices see Garleanu et al. [2009] as well as Fournier and Jacobs [2020]) and Cetin et al. [2006] show empirically and theoretically the impact of stock illiquidity on option prices. Despite the importance of the microstructure of equity options on their prices, the capital structure of a firm is known to have important bearings on the pricing of stock options. Since equity is essentially a call option on the value of the firm, one cannot consistently model both firm value and stock price as geometric Brownian motions. This led to Geske [1979] (Geske et al. [2016]) treatment of stock options as compound options. Since equity value is equal to zero when bankruptcy occurs, the stock return attains a heavier left tail than in a Lognormal distribution, implying a downward-sloping volatility skew—the more levered the firm is, the more skewed is the implied volatility function. This relationship has been confirmed empirically by Toft and Pryck [1997] (TP), who apply the compound option approach to the Leland [1994] model.[4]

[3]Xiao and Vasquez [2020] study the impact of measures of default and credit risk on future option returns with a stylized model of capital structure that prices options (with embedded jumps in the firm asset value process) which cannot be solved in closed form. Additionally Tian and Wu [2020] also show empirically a principal risk source component of the cross-sectional variation of individual stock options implied volatility skew is related to a firm's default risk.

[4]Dennis and Mayhew [2002] find less of an impact of leverage on the implied volatility skew as well as a different sign using a different data set of options data.

Our modeling of the effect of corporate disclosure quality on equity option pricing is an amalgamation of TP and the incomplete accounting information model of Duffie and Lando [2001] (DL). In DL, the firm value is assumed to follow a Geometric Brownian motion. However, investors do not know the precise value of the firm. Instead, the firm periodically issues noisy accounting reports. DL show that there is a positive probability per unit time that the firm can go bankrupt within the next instant even if the firm is reported to be *safe*, because the true firm value may lie somewhere close to the default threshold. This introduces jump risk into the model, as equity value can suddenly go to zero from any positive level. The crucial parameter in the DL model, the variance of the accounting noise, governs the intensity of the jump to default.[5]

A recent literature has emerged that is studying the quality of information (accounting and other) disclosure on option prices. In particular Dubinsky et al. [2019] quantify the impact of earnings announcements on option prices using deterministic jump components.[6] Our research question is different from that done in Smith [2018] (and Smith [2019]) which constructs a theoretical option pricing model to specifically incorporate the impact of an accounting *disclosure event date* on the price of an option. The paper is modeling the impact of the event itself directly on prices which is different from modeling the *quality* of the information level itself. As well the paper has no direct empirical analysis or support mentioned whereas our model theoretical predictions are consistent with our empirical results as defined in our model.

This paper should also not be confused with the work of Vanden [2008] which constructs a multi-period rational expectations model in which investors have private asymmetric information and trade using a derivative contract. Our model framework does not make any predictions regarding investors with different types of information quality.

Note that our research question is different from that done in the thesis work of Chandia [2014] which analyzes whether the option volatility skew is impacted by the accounting quality through the channel of *accrual choice* and quality. The work contains no theoretical predictions and their definition of *accrual choice* as well as quality are ambiguous whereas our model theoretical predictions are consistent with our empirical results as defined in our model.

[5] Du et al. [2019] find that a sizeable asset jump risk premia is needed, in addition to priced stochastic asset volatility, to help reconcile short term observed credit spreads on investment grade firms.

[6] Additionally see Lee [2012], Jeon et al. [2021], and Baker et al. [2020] for additional components of economic jump sources.

4.3 The Model

Our goal is to examine equity option pricing with consideration for the firm's capital structure as well as the fact that public information about the firm provides only a noisy estimate of its true value. The natural framework for this is the DL structural credit risk model with incomplete accounting information. Our derivation of the stock option pricing formula takes three steps. Step 1 takes into account endogenous bankruptcy by shareholders and expresses equity value as the solution to an optimal stopping problem. Step 2 assumes perfectly observed firm value and prices stock options using the compound options approach. Step 3 incorporates noisy accounting information into the pricing formula. The results of the first two steps are directly from DL and TP, and the final step is a combination of the two.

4.3.1 Equities

To begin, assume that we have a firm whose asset level is described by V_t, a geometric Brownian motion with drift μ and volatility σ. The firm generates cash flow at a rate δV_t, which is distributed to shareholders as well as being used to service a par consol bond with a total coupon rate C. The tax rate is θ, generating tax benefits for the bond at a rate of θC. All agents are risk-neutral and discount cash flows at a constant rate of r.

In this setting, one can conjecture that shareholders will choose to liquidate the firm if the asset level V_t becomes sufficiently low. Namely, the stopping policy that maximize the discounted present value from operating the firm, including tax benefits,

$$E\left(\int_t^{\tau} e^{-r(s-t)}\left(\delta V_s - C + \theta C\right) ds | V_t\right), \tag{4.3.1}$$

takes the form $\tau\left(V_B\right) = \inf\left\{t : V_{tB}\right\}$. Indeed, DL verify the optimality of such a stopping rule and show the corresponding equity value as

$$w\left(v\right) = \frac{\delta v}{r-\mu} - \frac{v_B\left(C\right)\delta}{r-\mu}\left(\frac{v}{v_B\left(C\right)}\right)^{-\gamma} - \left(1-\theta\right)\frac{C}{r}\left(1-\left(\frac{v}{v_B\left(C\right)}\right)^{-\gamma}\right) \tag{4.3.2}$$

when $v > v_B\left(C\right)$, and $w\left(v\right) = 0$ for $v_B\left(C\right)$. The terms represent, respectively, the present value of future cash flows generated by the assets, the present value of cash flows lost to bankruptcy, and the cost of debt service minus the tax benefit. Here, $v_B\left(C\right)$ is the default threshold

$$v_B\left(C\right) = \frac{\left(1-\theta\right)C\gamma\left(r-\mu\right)}{r\left(1+\gamma\right)\delta}, \tag{4.3.3}$$

where $\gamma = \frac{m+\sqrt{m^2+2r\sigma^2}}{\sigma^2}$ and $m = \mu - \frac{1}{2}\sigma^2$.

Notice that the total coupon rate C can be determined by maximizing the initial value of equity plus the market value of the consol bond. This is the notion of optimal capital structure pursued in Leland [1994] and Leland and Toft [1996]. It is not essential to our subsequent results and is therefore omitted.

4.3.2 Stock Options

Given shareholders' optimal decision to liquidate the firm once the firm asset level drops below $V_B = v_B(C)$, we can price stock options as compound options on the firm's assets.

To facilitate the derivations, define a standard Brownian motion $Z_t \equiv \log V_t$ and let $\underline{v} = \log V_B$. The price at time t of a stock option with maturity T and strike price K is

$$h_t(u) = e^{-r(T-t)} E\left(\left(w\left(e^{Z_T}\right) - K\right)^+ 1_{\{\tau > T\}} | Z_t = u\right). \tag{4.3.4}$$

The indicator function means that the option payoff is zero if bankruptcy is declared prior to its maturity.

Using Bayes' rule, this can be rewritten as

$$\begin{aligned} & e^{-r(T-t)} \int_{\underline{v}}^{\infty} \left(w\left(e^x\right) - K\right)^+ P\left(Z_T \in dx, \tau > T | Z_t = u\right) \\ = \; & e^{-r(T-t)} \int_{\underline{v}}^{\infty} \left(w\left(e^x\right) - K\right)^+ P\left(\tau > T | Z_t = u, Z_T = x\right) P\left(Z_T \in dx | Z_t = u\right) \end{aligned} \tag{4.3.5}$$

This expression involves two probabilities. The first is the probability of survival through T given that the standard Brownian motion Z is *pinned* at the two end points, and can be written as $\psi\left(u - \underline{v}, x - \underline{v}, \sigma\sqrt{T-t}\right) = 1 - \exp\left(-\frac{2(u-\underline{v})(x-\underline{v})}{\sigma^2(T-t)}\right)$. The second is simply the density of a normal random variable with mean $u + m(T-t)$ and variance $\sigma^2(T-t)$. We therefore obtain

$$\begin{aligned} h_t(u) \;=\; & \frac{e^{-r(T-t)}}{\sqrt{2\pi\sigma^2(T-t)}} \int_{\underline{v}}^{\infty} \left(w\left(e^x\right) - K\right)^+ \exp\left(-\frac{\left(x - u - m(T-t)\right)^2}{2\sigma^2(T-t)}\right) \\ & \left(1 - \exp\left(-\frac{2\left(u - \underline{v}\right)\left(x - \underline{v}\right)}{\sigma^2(T-t)}\right)\right) dx. \end{aligned} \tag{4.3.6}$$

After substituting in the value of equity in equation (4.3.2) and some further manipulations, we arrive at the final expression:

$$\begin{aligned} h_t(u) &= Fe^{u+\left(m+\frac{1}{2}\sigma^2-r\right)l}\left(N\left(y^*+\sigma\sqrt{l}\right)-e^{\left(2m\sigma^{-2}+2\right)(\underline{v}-u)}N\left(y^*+\sigma\sqrt{l}+\frac{2\left(\underline{v}-u\right)}{\sigma\sqrt{l}}\right)\right) \\ &\quad +Be^{\gamma(\underline{v}-u)}\left(N\left(y^*-\gamma\sigma\sqrt{l}\right)-e^{\left(2m\sigma^{-2}-2\gamma\right)(\underline{v}-u)}N\left(y^*-\gamma\sigma\sqrt{l}+\frac{2\left(\underline{v}-u\right)}{\sigma\sqrt{l}}\right)\right) \\ &\quad -\left(A+K\right)e^{-rl}\left(N\left(y^*\right)-e^{2m\sigma^{-2}(\underline{v}-u)}N\left(y^*+\frac{2\left(\underline{v}-u\right)}{\sigma\sqrt{l}}\right)\right), \end{aligned} \tag{4.3.7}$$

where

$$\begin{aligned} l &= T-t, \\ F &= \frac{\delta}{r-\mu}, \\ A &= \frac{\left(1-\theta\right)C}{r}, \\ B &= \frac{\left(1-\theta\right)C}{r\left(1+\gamma\right)}, \\ y^* &= -\frac{x^*-u-m\left(T-t\right)}{\sigma\sqrt{T-t}}, \end{aligned}$$

and e^{x^*} is the firm asset level that corresponds to an equity value of K:

$$Fe^x-A+Be^{-\gamma(x-\underline{v})}=K. \tag{4.3.8}$$

With some differences in notations, this pricing formula is first derived by TP. They show that it leads to a downward-sloping volatility smile. Furthermore, the skewness of the pattern increases with firm leverage. Our comparative statics in Section 3 will re-examine these findings.

4.3.3 Incomplete Accounting Information

The pricing formula (4.3.7) assumes that the firm asset level V_t is known to investors, which may seem like a harmless assumption. However, recent accounting scandals suggest that this is quite far from reality. DL assume that firm assets are observed only periodically, and with noise. Therefore, at time t the value of the firm's assets is not known with perfect precision. Instead, it is governed by a conditional distribution that depends on the reported firm assets as well as the absence of bankruptcy. DL use this framework to investigate the pricing of defaultable bonds, producing strong predictions for short-maturity credit spreads. We intend to study its impact on the pricing of stock options.

First, we follow DL in assuming that Z_0 is observed without noise. Then, at time t, the firm reports the value of $Y_t = Z_t + U_t$, where U_t is independent of Z_t and normally distributed with mean $-\frac{a^2}{2}$ and variance a^2, so that the reported firm asset level $\widehat{V}_t = V_t e^{U_t}$ is a noisy but unbiased version of V_t.

Next, we show that the stock option value can be expressed as an integral of equation (4.3.7) over the distribution of Z_t given information available at t. This can be seen by writing the option price H_t as

$$\begin{aligned} H_t &= e^{-r(T-t)} E\left(\left(w\left(e^{Z_T}\right) - K\right)^+ 1_{\{\tau>T\}} | Y_t = y, Z_0 = z_0, \tau > t\right) \\ &= e^{-r(T-t)} \int_{\underline{v}}^{\infty} \left(w\left(e^x\right) - K\right)^+ P\left(Z_T \in dx, \tau > T | Y_t = y, Z_0 = z_0, \tau > t\right). \end{aligned} \tag{4.3.9}$$

The joint density of Z_T and no default until T given the current and lagged asset values, as well as survival to t, can be decomposed by repeated applications of Bayes' rule:

$$\begin{aligned} & P\left(Z_T \in dx, \tau > T | Y_t = y, Z_0 = z_0, \tau > t\right) \\ = & \int_{u=\underline{v}}^{\infty} P\left(Z_T \in dx, Z_t \in du, \tau > T | Y_t = y, Z_0 = z_0, \tau > t\right) \\ = & \int_{u=\underline{v}}^{\infty} P\left(\tau > T | Z_t = u, Z_T = x\right) P\left(Z_T \in dx, Z_t \in du | Y_t = y, Z_0 = z_0, \tau > t\right) \\ = & \int_{u=\underline{v}}^{\infty} P\left(\tau > T | Z_t = u, Z_T = x\right) P\left(Z_T \in dx | Z_t = u\right) P\left(Z_t \in du | Y_t = y, Z_0 = z_0, \tau > t\right) \end{aligned} \tag{4.3.10}$$

Of the three probabilities above, the first is simply the survival probability of the *pinned* Brownian motion, previously referred to as $\psi\left(u - \underline{v}, x - \underline{v}, \sigma\sqrt{T-t}\right)$. The second is the density of a normal random variable with mean $u + m\left(T - t\right)$ and variance $\sigma^2\left(T - t\right)$. The last is the density of Z_t given current and lagged asset reports and survival. We denote it by $g\left(u | y, z_0, t\right) du$.

Combining these two equations, we obtain

$$H_t = \frac{e^{-r(T-t)}}{\sqrt{2\pi\sigma^2\left(T-t\right)}} \int_{\underline{v}}^{\infty} \left(w\left(e^x\right) - K\right)^+ \int_{\underline{v}}^{\infty} \exp\left(-\frac{\left(x - u - m\left(T-t\right)\right)^2}{2\sigma^2\left(T-t\right)}\right)$$

$$\left(1-\exp\left(-\frac{2\left(u-\underline{v}\right)\left(x-\underline{v}\right)}{\sigma^2\left(T-t\right)}\right)\right)g\left(u|Y_t,Z_0,t\right)dudx. \tag{4.3.11}$$

Switching the order of integration, it is easy to see that

$$H_t=\int_{\underline{v}}^{\infty}h_t\left(u\right)g\left(u|y,z_0,t\right)du, \tag{4.3.12}$$

which is in some sense obvious given the assumed risk-neutrality.

Following DL, the last unknown quantity g can be written as

$$\begin{aligned}
g\left(u|y,z_0,t\right)du &= P\left(Z_t|Y_t=y,Z_0=z_0,\tau>t\right)\\
&= \frac{P\left(Z_t\in du,\tau>t|Y_t=y,Z_0=z_0\right)}{P\left(\tau>t|Y_t=y,Z_0=z_0\right)}\\
&= \frac{P\left(Z_t\in du,\tau>t|Y_t=y,Z_0=z_0\right)}{\int_{\underline{v}}^{\infty}P\left(Z_t\in du,\tau>t|Y_t=y,Z_0=z_0\right)}.
\end{aligned} \tag{4.3.13}$$

The numerator, denoted as $b\left(u|y,z_0,t\right)du$, is equal to

$$\begin{aligned}
b\left(u|y,z_0,t\right)du &= \frac{P\left(\tau>t|Z_t=u,Z_0=z_0\right)P\left(Z_t\in du,Y_t\in dy\right)}{P\left(Y_t\in dy\right)}\\
&= \frac{\psi\left(z_0-\underline{v},z-\underline{v},\sigma\sqrt{t}\right)\phi_U\left(y-u\right)\phi_Z\left(u\right)du}{\phi_Y\left(y\right)},
\end{aligned} \tag{4.3.14}$$

where ϕ_U, ϕ_Z and ϕ_Y are respectively the densities of U_t, Z_t and Y_t. We note that $U_t\sim N\left(-a^2/2,a^2\right)$, $Z_t\sim N\left(z_0+mt,\sigma^2t\right)$ and $Y_t\sim N\left(-a^2/2+z_0+mt,a^2+\sigma^2t\right)$. The last step above uses the independence between Z_t and U_t.

Putting everything together, we obtain

$$g\left(u|y,z_0,t\right)=\frac{\sqrt{\frac{\alpha}{2\pi}}\left(1-\exp\left(-\frac{2}{\sigma^2t}\widetilde{z}_0\widetilde{u}\right)\right)\exp\left(-\frac{1}{2a^2}\left(\widetilde{y}-\widetilde{u}\right)^2\right)\exp\left(-\frac{1}{2\sigma^2t}\left(\widetilde{u}-\widetilde{z}_0-mt\right)^2\right)}{N\left(\sqrt{\alpha}\beta\right)-N\left(\sqrt{\alpha}\left(\beta-\frac{2}{\sigma^2t\alpha}\widetilde{z}_0\right)\right)\exp\left(-\frac{\alpha\eta}{2}+\frac{2}{\sigma^4t^2\alpha}\widetilde{z}_0\left(\widetilde{z}_0-\alpha\beta\sigma^2t\right)\right)}, \tag{4.3.15}$$

where $\widetilde{z}_0=z_0-\underline{v}$, $\widetilde{u}=u-\underline{v}$, $\widetilde{y}=y-\underline{v}+a^2/2$, and

$$\begin{aligned}
\alpha &= \frac{\sigma^2t+a^2}{a^2\sigma^2t},\\
\beta &= \frac{\sigma^2t\widetilde{y}+a^2\left(\widetilde{z}_0+mt\right)}{\sigma^2t+a^2},
\end{aligned}$$

$$\eta = \frac{\sigma^2 t a^2 \left(\widetilde{y} - (\widetilde{z}_0 + mt)\right)^2}{\left(\sigma^2 t + a^2\right)^2}.$$

This expression is slightly different in form from, but equivalent to, the one in DL.

In order to solve for our option pricing model with incomplete accounting information, given by H_t, in closed form, we first note that the function $g(u|y, z_0, t)$ can be re-written as a difference of Normal distribution probability density functions with different means and variances in the general form (see appendix for details of derivation and notation):

$$g(u|z_0, y) = \frac{L_1}{L_0} \times e^{\frac{-(u-M_1)^2}{2\psi}} - \frac{L_2}{L_0} \times e^{\frac{-(u-M_2)^2}{2\psi}} \tag{4.3.16}$$

Our formula for H_t involves the integral of the product of the TP option pricing formula (denoted $h_t(u)$) and the density function $g(u|y, z_0, t)$. Expanding the product results in computing twelve integrals of the form in equation 4.3.17 (see Owen [1980]).

$$\int_R^{\infty} \Phi(A + Bx)\phi(x)dx = \Phi\left(\frac{A}{\sqrt{1+B^2}}\right) - \Phi_2\left(\frac{A}{\sqrt{1+B^2}}, R; \frac{-B}{\sqrt{1+B^2}}\right) \tag{4.3.17}$$

The resulting sum of the twelve integrals results in a closed for solution, up to Bivariate Normal probability function, of the expression H_t. The resulting expression and sketch of the proof for H_t is available in the appendix of this paper and an online mathematical appendix for full step by step derivation.

4.4 Numerical Examples

In this section we use numerical examples to illustrate the theoretical model outlined in the preceding section. We focus on the interaction between accounting transparency and leverage, asset volatility, as well as the maturity of the equity option in the determination of the implied volatility smile.

The model parameters are chosen to be as close as possible to those used by TP, so that comparisons can be made where appropriate. The basic parameters are $m = 0.01$, $\sigma = 0.2$, $r = 0.08$, and $\theta = 0.35$. The starting point of the firm value is set to $V_0 = 100$, and the reported firm value at $t = 1$ is $Y_1 = 100$. The leverage ratio of the firm is controlled by the coupon rate of the consol bond, C. Given the payout rate δ, the dividend yield on the stock is set to $\frac{\delta Y_1-(1-\theta)C}{E_1}$ where E_1 is the equity value corresponding to an asset level of Y_1. The dividend yield and the equity value are useful when inverting the Black-Scholes formula for the implied volatility.

In the examples below, the case of low leverage assumes $C = 1.72$ and $\delta = 3.70\%$; medium leverage, $C = 3.77$ and $\delta = 4.54\%$; high leverage, $C = 6.42$ and $\delta = 5.66\%$. These parameter values are taken directly from TP's Table I. We also assume that the accounting precision parameter a can take on three values: 0.04, 0.08, and 0.12. Lastly, we take the maturity of the stock option to be 1 month.

4.4.1 Leverage

We first examine the role of leverage where the accounting reports are assumed to be precise. We do this by setting $a = 0$, which reduces the model to that of TP. To isolate the shape of the volatility smile pattern, we normalize the strike price by the current value of equity, and divide all of the implied volatilities by the at-the-money implied volatility, resulting in a *standardized* plot.

Figure C.1.1 presents the volatility smile for cases with low, medium, or high leverage ratios. Panel 1 is where the accounting precision is set to zero, representing precisely observed asset values. It shows that the volatility skew increases with leverage, something that TP document both in theory and empirically. Notably, the magnitude of the skew is very close to what TP present in their Table II. The remaining three panels then illustrate what happens when we have imprecise accounting reports. Notice that for the case with $a = 0.00$ (Panel 1) we still have an implied volatility skew that increases with leverage, but this relationship reverses itself under the presence of somewhat higher accounting noise (Panels 2, 3 and 4). This observation suggests that there is a level of accounting precision under which the skewness of the smile is insensitive to leverage. Consequently, one has to be careful in designing regression tests for the relationship between skewness and leverage. It is possible, for example, that simply regressing the skewness on leverage would not be able to uncover any relation at all.

To explain this puzzling behavior, we examine the relationship between the volatility smile and accounting precision under different assumptions of firm leverage. First, Figure C.1.2 shows that this relationship is clearly monotonic, with higher skewness associated with higher accounting noise. Since option value is a convex function of equity value in TP's model, the accounting noise boosts option value due to Jensen's effect. With deep-in-the-money call options (with low strike prices) that are very close to the no-arbitrage lower bound, a small increase in the option value requires a large increase in equity volatility.[7] This results in the elevated volatility skew.

In Figure C.1.2, the impact of accounting imprecision can be ostensibly much larger than the effect of leverage, but one has to keep in mind that we do not yet know what a *reasonable* level of a is. A value of $a = 0.12$, for example, implies an accounting report that can regularly be 12 percent off target. We also see from Figure C.1.2 that accounting noise disproportionately affects the left side of the volatility smile. The intuition is of course that the imprecise observation of firm assets introduces downward jump risk due to the presence of the default boundary.

The most important observation from Figure C.1.2 is perhaps the decreasing effect of accounting noise with leverage—the variation of the skew is noticeably smaller in Panel 3 than in Panel 1. This observation, which helps to explain the relationship between skewness and leverage in Figure C.1.1, can be understood intuitively that in our model the volatility skew is generated by default through two channels. First, the firm asset level can diffuse down to the default boundary. Second, the firm value can suddenly *jump* to default due to incomplete accounting information. As the firm moves closer to the default boundary (represented by higher leverage ratios), it becomes easier to default through the normal diffusion channel, and the relative importance of the jump risk decreases.

Figure C.1.3 shows a bar graph of the change in the implied volatility skewness with respect to leverage for different accounting transparency ($a = 0.04$, $a = 0.08$, and $a = 0.12$ where a higher value of $'a'$ is associated with less transparency). We show bar charts for three sets of differences in the implied volatility surface where each different bar chart color (blue, orange, and grey) corresponds to a different point in change in the implied volatility skewness. For each set differences in the implied volatility surface we show bar charts for different accounting transparency levels $a = 0.04$, $a = 0.08$, and $a = 0.12$, hence this results in three bar charts for each set of differences in the implied volatility surface. For a particular differences in the implied volatility surface (say the set of blue colored bar charts), we see that the lower the value of $'a'$ accounting transparency the more negative the change in the implied volatility skewness with respect to leverage. The model prediction is the same for all three sets of differences in the implied volatility surface.

[7] Another way to understand it is that these options have very low vega according to the Black-Scholes model.

4.5 Data

The data used for our empirical tests is merged from the following standard sources: quarterly balance sheet data is from COMPUSTAT, daily and monthly stock data is from CRSP, equity options data is obtained from the OptionMetrics Database, and earnings information is from the I/B/E/S data. The sample period is from January 1997 to December 2017. OptionMetrics equity options data is first available as of January 1996, however, the merge between OptionMetrics and the earnings information is sparse during 1996 hence we begin our sample at January 1, 1997.

Our equity options data consists of closing end of day call and put option best bid and best offer quotes from the OptionMetrics Database which collects closing option quotes data from all U.S. equity option exchanges. Equity options for individual firms are American in nature in the sense that they can be exercised at any time. For each quoted option contract price the corresponding contract open interest, daily trade volume, and Black-Scholes delta and implied volatility are reported. The Black-Scholes delta and implied volatility are computed using the Cox et al. [1979] binomial lattice model in order to incorporate the early exercise features of American options. We apply filters to our options data set, specifically we remove contracts with missing implied volatility, open interest, trade volume, and delta. We remove contracts that have less than 10 days remaining to maturity in order to account for the rollover of option contracts. We remove option contracts with option best bid or best offer prices that are less than or equal to zero, and cases where the best offer is less than or equal to the best bid. We also require the absolute value of the option contract delta to be between 0.02 and 0.98 to avoid using very deep in-the-money and out-of-the-money contracts that are mis-priced and have low liquidity.

We take as at-the-money option the closest option contract that as long as the moneyness $((K/S_0))$ is between 0.99 and 1.03, with corresponding moneyness $((K/S_0)_t^{High})$ and implied volatility $(IVOL_t^{High})$ for every firm each day.

For the out-of-the-money put, we choose the closest option contract that as long as the moneyness below 0.97 and above 0.92 (and if not available below 0.97 then the lowest below 0.92), with corresponding moneyness $((K/S_0)_t^{Low})$ and implied volatility $(IVOL_t^{Low})$. As a metric for the volatility skew we consider the following variable:

$$SLOPE_t = \frac{IVOL_t^{High} - IVOL_t^{Low}}{(K/S_0)_t^{High} - (K/S_0)_t^{Low}} \tag{4.5.1}$$

So for a typical skew, this variable is negative and the more negative the more skewed the smirk/smile is. In the robustness section 4.7 we show that our results are robust to scaling the numerator of equation 4.5.1 by the ATM implied volatility ($IVOL_t^{High}$). Note that this variable only measures the left hand side of the skew, which is where our theory's predictions are the strongest. Our measure in equation 4.5.1 is computed each day for every firm and averaged across all options within the time interval of 15 to 45 days (and in robustness section 4.7 we average across all options of all times to maturity for each day for each firm) and then averaged over the quarter for each firm.

Our measure of leverage is computed using quarterly COMPUSTAT data as the ratio of book value of debt divided by sum of debt and market value of equity using quarterly data from COMPUSTAT. We assume that book values of debt and preferred stock are adequate proxies for the corresponding market values. In order to account for heterogeneity of firm characteristics, we use several control variables known to impact the implied volatility skew in the most recent empirical tests in Morellec and Zhdanov [2019]. Specifically we control for the: market-to-book ratio (M/B), market capitalization of the firm (Size), stock momentum (Momentum), stock beta (Beta), idiosyncratic stock return skewness (Idio Skew), and at-the-money option implied volatility (Atm Ivol). The market-to-book ratio (M/B) is the ratio of quarterly market equity divided by book equity using quarterly data from COMPUSTAT. We also control for the market capitalization of the firm (size) which is the log of the product of the stock price and shares outstanding from CRSP monthly stock files (firms with share codes 10 and 11 common shares). Momentum is the past 6 month cumulative monthly stock returns from CRSP which controls for the stock momentum over the previous 6 months of returns. Beta is the stock beta with the market estimated from 36 months rolling regressions adjusted by 3 months of lags for asynchronous trading as per the Dimson [1979] adjustment. Idio Skew is the idiosyncratic skewness of daily returns estimated quarterly using daily CRSP stock returns. Atm Ivol is average of call and put contract implied volatility with $|\Delta| = 0.5$ and 30 days to maturity, and using OptionMetrics Volatility surface computed daily on a firm level and then averaged over the quarter.

Our measures of accounting quality are based on earnings information from I/B/E/S database. We compute the dispersion in analyst forecast (Disp) and number of analyst covering stock (Nanalyst) from the data set of firm characteristics in Green et al. [2017] (GHZ, henceforth).[8] The list of firm characteristics in GHZ is constructed using all firms with common shares that are listed on the AMEX, NYSE, or NASDAQ, that have end of month value on CRSP, quarterly and annual balance sheet reporting on COMPUSTAT and earnings information reported to the I/B/E/S data.

[8]We thank Jeremiah Green for making the SAS code to construct the data set freely available on his website.

As the number of analysts (Nanalyst) increases, the quality of the accounting transparency of the firm increases since more analysts are paying attention to the earnings statement of the firm and will yield a more precise estimate of firm value. As the dispersion in analyst forecast (Disp), decreases from a high dispersion quantity to a low dispersion quantity the quality of the accounting transparency of the firm increases since more is associated with a more precise statement of the quality of accounting information.

INSERT TABLE C.1.1 **HERE**

In Section C.1.2, Table C.1.1 presents quarterly summary statistics for main variables from January 1997 to December 2017. Our implied volatility skewness measure equation 4.5.1 is negative on average and up to the 99-th percentile. A negative value indicates that the OTM implied volatility ($IVOL_t^{Low}$) is higher than the ATM implied volatility ($IVOL_t^{High}$), the more expensive OTM options are indicated by a higher implied volatility. The average firm leverage is 0.20 (or 20%)[9]. The Nanalyst ranges from zero to over 50 with a mean of five and a median number of analysts of three.

INSERT TABLE C.1.2 **HERE**

Correlations between variables are presented in Table C.1.2. Our implied volatility skewness measure equation 4.5.1 shows only moderate correlation with leverage, Nanalyst, and M/B of -0.13, 0.19 and 0.10 respectively. There is also very high correlation between Nanalyst and size of 0.76, however, this is not a concern in our empirical analysis since we do not directly use the level of Nanalyst and our analysis is robust to the exclusion of size.

4.6 Empirical Hypothesis and Supporting Evidence

We will focus on the empirical implication of the impact of leverage on the volatility smile based on the model prediction in Section 4.4. We noted that as the quality of accounting data declined, one should expect that the reported leverage becomes less influential on the pricing of options. More precisely, we will test the following hypothesis for different accounting quality measures and the impact of leverage on the volatility smile.

[9]Note that we remove financial and utilities firms (sector= 9 (SIC codes 60 to 67) and SIC= 49).

Main Hypothesis: a higher value of number of Nanalysts (lower value of Disp) is associated with a higher level of transparency (lower a in our model) which our model predicts would mean a more negative impact of leverage on skewness.[10]

Recall that TP predict that firms with higher leverage should exhibit more steeply negatively sloped skews. Hence when regressing our skew metric on leverage, the coefficient should be negative as long as the accounting transparency is not too low. Under our main empirical hypothesis, we expect to see that the coefficient is larger in absolute value for firms with greater transparency.

To measure the differential impact of leverage, based on the quality of accounting transparency on the volatility smile, we separate cross-section of firms based on the 50th percentile of each accounting transparency measure each quarter of each year. We define those firms that have *low accounting transparency* as those with the number of Nanalysts below the 50th percentile (Disp above the 50th percentile) each quarterly cross section. Correspondingly, We define those firms that have *high accounting transparency* as those with the number of Nanalysts above the 50th percentile (and Disp below the 50th percentile) each quarterly cross section.

We then estimate the impact of leverage on the volatility smile using quarterly Fama and Macbeth [1973] regressions within in each subset. The resulting estimates for each of the measures of accounting quality are reported in Tables C.1.3 and C.1.4 (for Nanalyst and Disp respectively). Panel A (B) is estimated using the firms with *low accounting transparency* (*high accounting transparency*).

INSERT TABLES C.1.3 and C.1.4 **HERE**

The variable (and sign) of interest is leverage. In each of the tables C.1.3 and C.1.4, for panels A and B, column 1 reports the univariate quarterly Fama and Macbeth [1973] regressions within in each subset regressions using leverage. In columns 2 to 7, we add individual control variables (Momentum, Beta, Size, M/B, Idio Skew, and Atm Ivol) one at a time to the univariate regression for leverage. Additionally in each of the columns, in each panel A and B, in the Tables C.1.3 and C.1.4 we report the R-squared (R^2) and average number of firms in each quarterly cross-sectional regressions (N obs). R^2, within each subset of data in Panels A(B), range from 1.6% up to 17.23% (when including leverage and all control variables). The average number of firms in each quarterly cross-sectional

[10] A more negative number is a negative number further from zero.

regressions total between both subset of data in Panels A(B) is roughly 1600 in column 1 to roughly 1200 in column 7.

We first look at Table C.1.3 column 1 in panel A low transparency firms and find that the coefficient is negative (-0.22) and statistically significant at the 99% level (t-stat of -6.09). In the corresponding columns 1 in panel B high transparency firms, the leverage coefficient is again negative (-0.34) and significant at the 99% level (t-stat of -17.36) and more negative (or larger in absolute magnitude) than the coefficient of leverage in panel A. This provides preliminary evidence that low accounting transparency implying a less marked impact of leverage than high accounting transparency (our main empirical hypothesis).

As we add control variables in Table C.1.3 columns 2 to 7 in panel A low transparency firms and find that the leverage coefficient is negative and statistically significant at the 99% level in each case. In the corresponding columns 2 to 7 in panel B high transparency firms, the coefficient is more negative (or larger in absolute magnitude) and significant at the 99% level. In column 7 in panel A (B), the leverage coefficient is -0.21 (-0.29) and both significant at the 99% level with a t-stat of -8.16 (t-stat of -17.37) and the leverage coefficient is more negative than the coefficient of leverage in panel A (controlling for firm characteristics). This provides evidence that low accounting transparency implying a less marked impact of leverage than high accounting transparency (our main empirical hypothesis) as measured by Nanalyst. In Table C.1.4 (Disp) for each of columns 1 to 7 in panel A low transparency firms and find that the leverage coefficient is negative and statistically significant at the 99% level in each case. In the corresponding columns 1 to 7 in panel B high transparency firms, the coefficient is more negative than the coefficient of leverage in panel A (controlling for firm characteristics).

So our findings appear consistent with low disclosure quality implying a less marked impact of leverage. This is consistent with the situation depicted in Figure C.1.3 and confirms our main empirical hypothesis.

4.7 Robustness Tests

In this section we provide several robustness tests for our main empirical findings in Section 4.6. We present our results for all four measures of accounting quality using (*i*) the average times to maturity across all options of all for each day for each firm and then averaged over the quarter, (*ii*) an alternative measure of implied volatility skewness averaged across all options within the time interval of 15 to 45 days (*iii*) using our alternative measure of implied volatility skewness estimated across the average times to maturity across all options of all for each day for each firm and then averaged over the quarter for each firm.

Tables C.1.5 and C.1.6 present our results using equation 4.5.1 estimated using the average times to maturity across all options of all for each day for each firm and then averaged over the quarter (for Nanalyst and Disp respectively). In Tables C.1.5 and C.1.6 for each of columns 1 to 7 in panel A low transparency firms and find that the leverage coefficient is negative and statistically significant at the 99% level in each case. For Table C.1.5, in the corresponding columns 1 to 7 in panel B high transparency firms, the coefficient is more negative than the coefficient of leverage in panel A (controlling for firm characteristics) which confirms our main empirical hypothesis. For Table C.1.6, in the corresponding columns 1 to 6 in panel B high transparency firms, the coefficient is more negative than the coefficient of leverage in panel A (controlling for firm characteristics) which confirms our main empirical hypothesis (except for column 7 of C.1.6).

When comparing the size of the estimates in Tables C.1.3 and C.1.4 using options with times to maturity of 15 to 45 days to maturity compared to those of Tables C.1.5 and C.1.6 with the average times to maturity across all options, we find that the difference between the leverage coefficients in panel A and B for each of the columns in Tables C.1.3 and C.1.4 is larger than that in Tables C.1.5 and C.1.6 (except for column 7 of C.1.6). This observation highlights that the impact of leverage is more pronounced for options with shorter average times to maturity than those with longer average times to maturity.

INSERT TABLES C.1.5 and C.1.6 **HERE**

We consider an alternative measure of implied volatility skewness to show our results are robust to different measures of skewness. We define a modified measure of equation 4.5.1 scaling the numerator by the ATM implied volatility ($IVOL_t^{High}$).

$$SLOPE_t(v2) = \frac{\frac{IVOL_t^{High} - IVOL_t^{Low}}{IVOL_t^{Low}}}{(K/S_0)_t^{High} - (K/S_0)_t^{Low}} \tag{4.7.1}$$

As implied volatility skew in equation 4.5.1, our skew equation 4.7.1 is negative and is computed each day for every firm and averaged across all options within the time interval of 15 to 45 days and then averaged over the quarter. Table C.1.1 shows that the skew equation 4.7.1 is more negative than equation 4.5.1 on average (-1.01 versus -0.58) and in median (-0.94 versus -0.55). Table C.1.2 shows that the correlation between the two measures is quite high at 0.63. The implied volatility skew in equation 4.7.1 is highly correlated with Atm Ivol at a level of 0.53, however, adding this variable or removing it in our regression analysis does not change our result.

Tables C.1.7 and C.1.8 present our results using equation 4.7.1 estimated using the time interval of 15 to 45 days of all options. In Tables C.1.7 and C.1.8 for each of columns 1 to 7 in panel A low transparency firms and find that the leverage coefficient is negative and statistically significant at the 99% level in each case. In the corresponding columns 1 to 7 in panel B high transparency firms, the coefficient is more negative than the coefficient of leverage in panel A (controlling for firm characteristics) which confirms our main empirical hypothesis using a different measure of implied volatility skewness. When comparing the estimates in Tables C.1.7 and C.1.8 to those in Tables C.1.3 and C.1.4 using equation 4.5.1, we find that the difference between the leverage coefficients in panel A and B for each of the columns in Tables C.1.7 and C.1.8 is larger than those in Tables C.1.3 and C.1.4 for each accounting transparency variable respectively.

INSERT TABLES C.1.7 and C.1.8 **HERE**

We consider an alternative version of our implied volatility skew equation 4.7.1 which is computed as the average across all options of all times to maturity for each day for each firm then averaged over the quarter for each firm.

INSERT TABLES C.1.9 and C.1.10 **HERE**

In Tables C.1.9 and C.1.10 present our results using equation 4.7.1 estimated using the average times to maturity across all options of all for each day for each firm and then averaged over the quarter. In Tables C.1.9 and C.1.10 for each of columns 1 to 7 in panel A low transparency firms and find that the leverage coefficient is negative and statistically significant at the 99% level in each case. In the corresponding columns 1 to 7 in panel B high transparency firms, the coefficient is more negative than the coefficient of leverage in panel A (controlling for firm characteristics) which confirms our main empirical hypothesis using a different measure of implied volatility skewness and using all average option times to maturity. When comparing the estimates in Tables C.1.9 and C.1.10 to those in Tables C.1.7 and C.1.8 using implied volatility skew equation 4.7.1, we find that the difference between the leverage coefficients in panel A and B for each of the columns in Tables C.1.7 and C.1.8 is larger than those in Tables C.1.9 and C.1.10 for each accounting transparency variable respectively. This observation shows the robustness of the same finding using implied volatility skew equation 4.5.1 (now using equation 4.7.1) that the impact of leverage is more pronounced for options with shorter average times to maturity than those with longer average times to maturity.

So our findings confirms our main empirical hypothesis and show that the main conclusion is robust to using options of different times to maturity, different specification of implied volatility skewness, and both.

4.8 Conclusion

We have developed a model of the interaction between financial leverage, the quality of corporate disclosure and the implied volatility skew for individual firms. Our model implies that firms with lower accounting quality have more pronounced skews. In addition, this relationship is weaker the higher the leverage. From a different angle, we show that the impact of leverage on the skew is stronger for firms with greater transparency and are able to provide empirical support for this prediction. We also find that our data is consistent with the prediction that the negative relationship between the skew and option maturity is stronger for lower transparency firms.

Our model provides an economically intuitive and empirically testable explanation for differences in deviations from the Black & Scholes option pricing framework across firms. Taken together with findings of a relationship between accounting transparency and the valuation of corporate debt, this provides support for the modeling framework of Duffie and Lando [2001] who link the precision of accounting information to the valuation of corporate securities.

Chapter 5

Conclusion

Chapter 2 of this dissertation is based on my job market paper Szaura [2020]. In cross-sectional regressions and portfolio sorts of over a hundred characteristics and factors, on average only 2.4% predict the cross-section of corporate bond returns when adjusting for higher statistical benchmarks to take into account all characteristics and factors being tested in a large unbiased data set. A horse-race of all characteristics and factors in cross-sectional regressions finds a higher number of corporate bond, rather than stock, characteristics and factors that predict the cross-section of corporate bond returns when adjusting for higher benchmarks. In addition to the lower number of corporate bond characteristics and factors that predict the cross-section of stock returns, my results suggest that the stock and corporate bond markets are more segmented than previously documented. Some avenues for potential future research currently being explored are whether corporate bond signals help predict corporate bond index returns (and whether stock signals help) as well as incorporating transaction costs into a co-entropy modeling approach.

Chapter 3 of this dissertation is based on the working paper Cremers et al. [2019]. Many measures of option implied stock mis-pricing have been proposed in the academic literature. However, we show that the majority of existing measures rely on shorting stocks that have high shorting fees. Once we remove stocks that are hard to short, many of the existing measures' statistical and economic significance drop or are not significant. We contribute to the academic literature by proposing a novel measure of option implied stock mis-pricing whose abnormal trading profits are driven by taking a long position in larger firms stocks where more option trading is done and whose abnormal trading profits are robust to removing stocks that are hard to short. Our measure, IPD, is constructed from a novel intra-day options trades data-set, hence is only calculated during periods of options trading and periods of price discovery. IPD captures informed trading for value-weighted portfolios whose abnormal trading profits are driven by taking a long position in larger firms and does not require taking a short position in firms that have high shorting fees. Some avenues for potential future research are to refine the construction of our IPD measure to consider are using an estimation technique for our IPD measure using a sequential estimation technique that takes into account the time between trades (say using Kalman-Filter estimation of the IPD measure).

Chapter 4 of this dissertation is based on the working paper Doshi et al. [2020]. In this paper we show theoretically and empirically that firms with higher accounting transparency have an implied volatility smirk that is more sensitive to leverage (vice versa). The more clear the accounting information the more skewed the implied volatility smirk. Our theoretical predictions rely on extending the Duffie and Lando [2001] credit risk model to stock option pricing whereby incomplete accounting information and the risk of bankruptcy together act as an economic source of jump risk for stocks. Our model can be solved in closed form up to Bivariate Normal Cumulative Distribution Function. Some avenues for potential future research are whether there is a differential impact of the asset volatility and firm debt maturity on the implied volatility skew when taking into account accounting transparency. Our model, as in the Duffie and Lando [2001], is based on the Leland [1994] model, however, if there is a differential impact of the asset volatility and firm debt maturity on the implied volatility skew when taking into account accounting transparency we would extend the underlying model by using the Leland and Toft [1996] (or in the case of stochastic asset volatility Du et al. [2018]).

APPENDICES

Appendix A

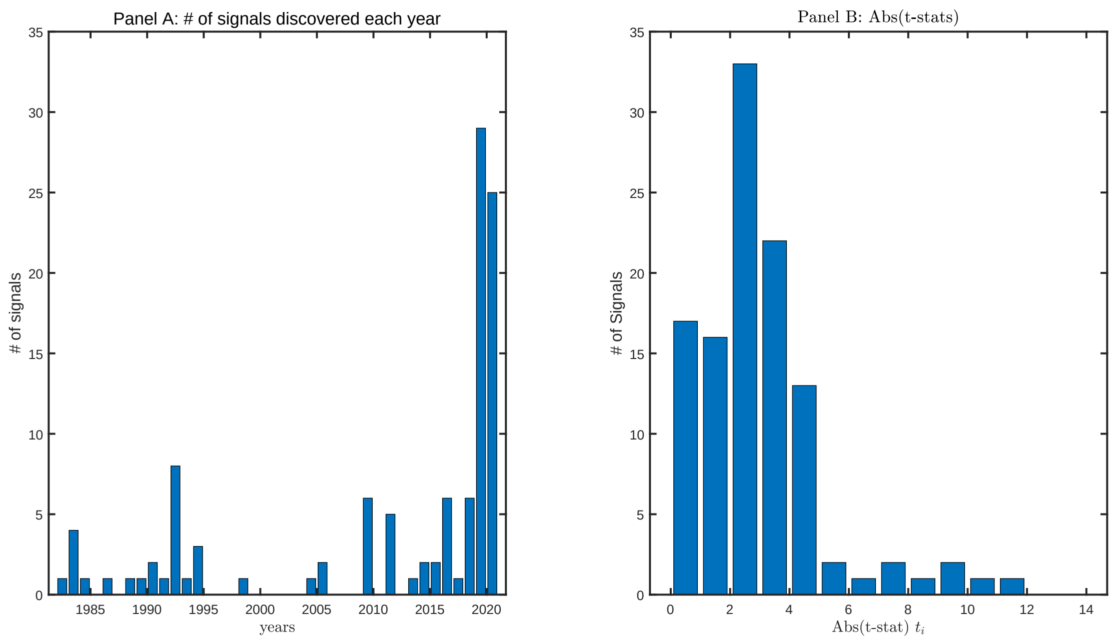
Appendix for Chapter 2

This appendix contains definitions and specific construction of variables used in Chapter 2.

A.1 Variable Construction, Figures, and Tables

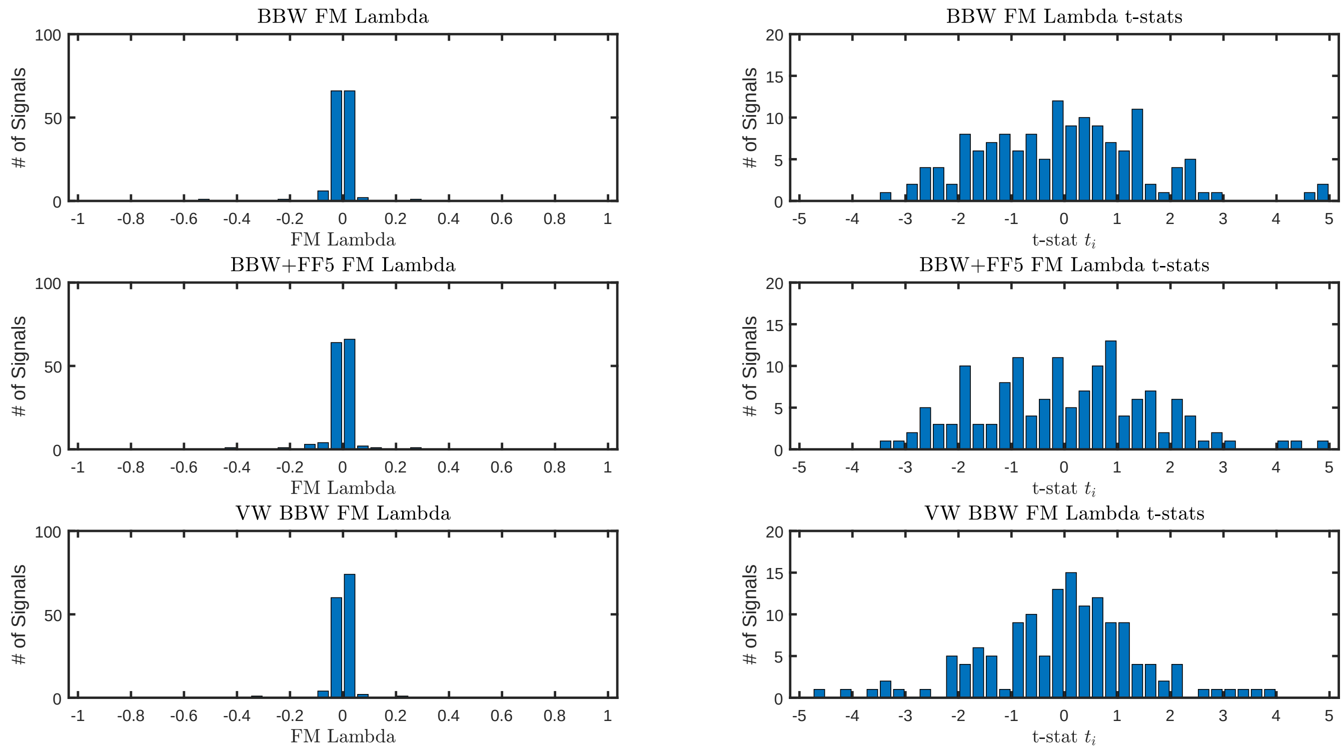
A.1.1 Figures

Figure A.1.1: : Number of Discovered Signals in the Cross-Section of Corporate Bond Returns



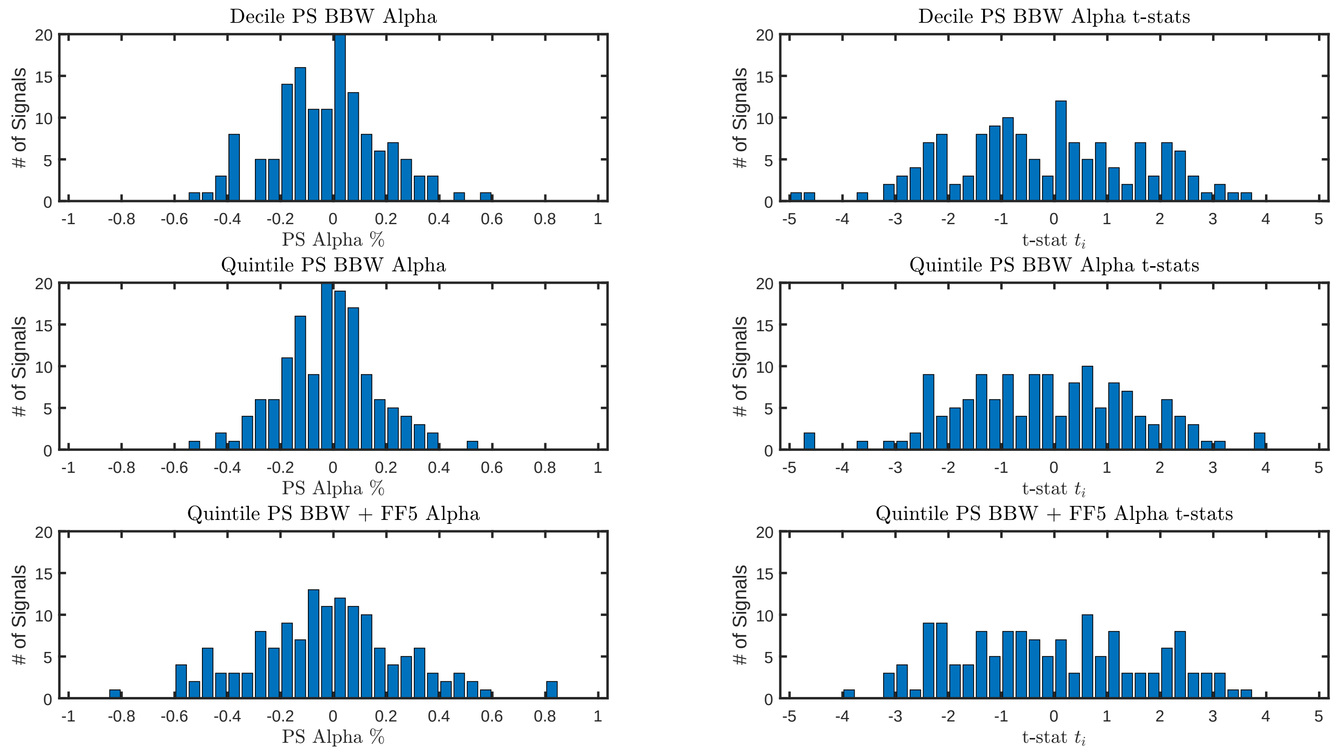
Panel A is the yearly number of discovered signals that predict the CS of Corporate Bond Returns. Panel B presents their Absolute t-statistics.

Figure A.1.2: : Fama-Macbeth λ_s and t_{λ_s}



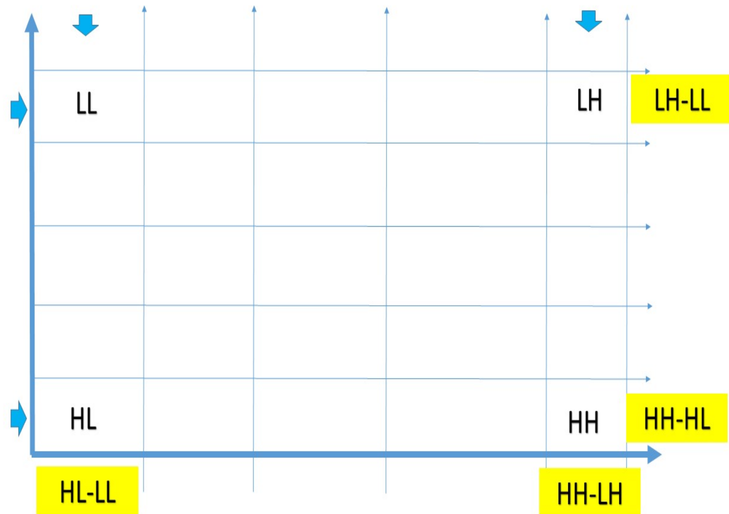
Each of the three rows of graphs shows the distribution of the average slope coefficient (λ_s) on the left-hand-side graph and corresponding t-statistic (t_{λ_s}) distribution on the right-hand-side for the 3 different univariate FM CS regressions of one-month-ahead risk-adjusted corporate bond excess returns of each of 143 signals. The first row are risk adjusted using the Bai et al. [2019b] (BBW) four factor corporate bond return model, second row are risk adjusted using the BBW and the Fama and French [2015] (FF5) five factor stock model, and third are risk adjusted using the BBW and value weighted least squares weighting bonds using one month lagged market capitalization.

Figure A.1.3: : Portfolio Sort α_s and t_{α_s}



Each of the three rows of graphs shows the distribution of the high minus low portfolio alpha (α_s) on the left-hand-side graph and corresponding t-statistic (t_{α_s}) distribution on the right-hand-side for the 3 different set of portfolios sorts of each of 103 signals. The first row high minus low portfolio alpha (α_s) and corresponding t-statistic (t_{α_s}) using the BBW corporate bond factor model (Decile PS (BBW)). The second row report the quintile high minus low portfolio α_s and t_{α_s} using the BBW (Quintile PS (BBW)). The third report the quintile high minus low portfolio alpha α_s and t_{α_s} using the BBW and FF5 (Quintile PS (BBW+FF5)).

Figure A.1.4: : Diagram of 5×5 conditional double sorts



5×5 conditional double sorts where corporate bonds returns are sorted into quintile portfolios based on a signal S_1 and then within each quintile sort into quintiles on signal S_2. I focus on the four corner portfolios of the 5×5 matrix of conditionally sorted portfolios. I denote the corner portfolios as HH, HL, LH, and LL depending on whether the corporate bond return is assigned to the high or low quintile based on the ranking of the first signal and on the second (conditionally on it's ranking within the first signal).

A.1.2 Tables

Table A.1.1: : Corporate Bond Return Summary Statistics

Variable	N	Mean	Median	Std. Dev.	Percentiles 10th	25th	75th	90th
Bond Returns (in %)	1,035,628	0.58	0.31	5.35	−2.17	−0.61	1.58	3.41
Rating (numerical scores)	1,661,839	9.51	8	5.32	4	6	11	21
Time to Maturity (in years)	1,729,992	9.49	6.39	9.06	1.88	3.37	12.72	23.17
Age (years)	1,729,929	4.74	3.4	4.6	0.65	1.61	6.36	10.2
Amount Outstanding (millions of USD)	1,729,992	375.71	250	549.34	3.26	11.9	500	1000
Yield (in %)	1,575,156	5.73	4.78	138.3	2.1	3.31	6.1	8.09
Coupon (in %)	1,729,766	5.3	5.5	2.2	2.6	4.15	6.63	7.75

Notes: This table reports the number of bond-month observations: the cross-sectional mean, median, standard deviation and various monthly percentiles of corporate bond excess returns (in percentage) and bond characteristics including credit rating (numerical scores), time to maturity (in years), age of the bond since issuance (in years), amount outstanding (Size, million), yield to maturity (in percentage), and coupon payment rate (in percentage). Ratings are numerical scores (1 refers to a AAA rating and 21 refers to a C rating) and a higher numerical score implies higher credit risk. The data ranges from July 1, 2002 to October 31, 2019.

Table A.1.2: : Stock based characteristics: Variable Definitions

No.	Acronym	Description/Name
1	spi	special items
2	mve f	market value of equity fiscal year
3	bm	Book to market
4	ep	earnings to price
5	cashpr	cash productivity
6	dy	dividend to price
7	lev	leverage
8	sp	sales to price
9	roic	return on invested capital
10	rd sale	R& D to sales
11	rd mve	R& D to market capiitalization
12	agr	asset growth
13	gma	gross profitability
14	chcsho	change in shares outstanding
15	lgr	growth in long term debt
16	acc	Working Capital accruals
17	pctacc	percent accruals
18	cfp	cash flow to price ratio
19	absacc	Absolute Accruals
20	chinv	change in inventory
21	cf	cash flow
22	hire	employee growth rate
23	sgr	sales growth
24	chpm	ib divided by sales
25	chato	change in sales /average total assets
26	pchsale pchinvt	% change sales less % change in inventory
27	pchsale pchrect	% change in sales less % change in A/R
28	pchgm pchsale	% change in gross margin less % change in sales
29	pchsale pchxsga	% change in sales less % change in SG& A
30	depr	deprecation/PPE
31	pchdepr	% change in depreciation
32	chadv	Change in Advertising expenses
33	invest	capital expenditures and inventory
34	egr	growth in common shareholder equity

35	pchcapx	% change in CAPX
36	grcapx	growth in capital expenditure
37	tang	debt capacity/firm tangibility
38	currat	current ratio
39	pchcurrat	% change in current ratio
40	quick	quick ratio
41	pchquick	% change in quick ratio
42	salecash	sales to cash
43	salerec	sales to receivables
44	saleinv	sales to inventory
45	pchsaleinv	percentage change in sales to inventory
46	cashdebt	cash flow to debt
47	realestate	real estate holdings
48	grltnoa	growth in long-term net operating assets
49	rdbias	% change in R& D less % change in income over equity
50	roe	return on equity
51	operprof	operating profitability
52	chpmia	Industry adjusted change in profit margin
53	chatoia	industry adjusted change in asset turnover
54	chempia	industry adjusted change in employees
55	bm ia	Industry adjusted book to market
56	pchcapx ia	industry adjusted % change in capital expenditures
57	tb	tax income to book income
58	cfp ia	industry adjusted cash flow to price ratio
59	herf	industry sales concentration
60	orgcap	organizational capital
61	mve m	market value of equity monthly
62	pps	lag log price per share
63	rdq	research and developpment quarterly
64	prccq	price of common shares quarterly
65	chtx	change in tax expense
66	roaq	return on assets
67	roeq	return on equity quarterly
68	rsup	revenue surprise
69	stdacc	accrual volatility
70	sgrvol	sales growth volatility
71	roavol	earnings volatility
72	stdcf	cash flow volatility

73	cash	casj holdings
74	cinvest	corporate investment
75	sue	unexpected quarterly earnings
76	aeavol	Abnormal earnings announcement volume
77	ear	earnings announcement
78	disp	dispersion in forecasted EPS
79	chfeps	change in forecasted EPS
80	fgr5yr	forecasted growth in 5yr EPS
81	MEANREC	Mean Analyst Recom.
82	chrec	Change in Mean Analyst Recom.
83	nanalyst	number of analyst covering stock
84	sfe	scaled earnings forecast
85	MEANEST	Mean Analyst Estimate
86	mom6m	6 month momentum
87	mom12m	12 month momentum
88	mom36m	36 month momentum
89	mom1m	1 month momentum
90	dolvol	dollar trading volume
91	chmom	change in 6 month momentum
92	turn	share turnover
93	indmom	industry momentum
94	maxret	maximum daily return
95	retvol	return volatility
96	baspread	bid ask spread
97	std dolvol	std dev dollar trading volume
98	std turn	volatility of liquidity share turnover
99	beta	beta
100	betasq	beta squared
101	rsq1	adjusted rsquared
102	pricedelay	Price Delay
103	idiovol	idiosyncratic return volatility

Notes: This table presents definitions from Green et al. [2017] (GHZ). The detailed description of each variable is outlined in Green et al. [2017].

Table A.1.3: : Variable Construction and Paper References

Variable Name	Construction	Paper Reference
Bond Market Risk Factor	Estimated 48 monthly rolling returns on BBW factors	Bai et al. [2019b]
Bond Credit Risk Factor	Estimated 48 monthly rolling returns on BBW factors	Bai et al. [2019b]
Bond Liquidity Risk Factor	Estimated 48 monthly rolling returns on BBW factors	Bai et al. [2019b]
Bond Downside Risk Factor	Estimated 48 monthly rolling returns on BBW factors	Bai et al. [2019b]
Bond VIX Factor	Estimated 48 monthly rolling returns on FF5 factors, bond term, bond default, and VIX	Chung et al. [2019]
Bond idiosyncratic volatility	RMSE from estimated BBW factors	Chung et al. [2019]
Bond Total volatility	Monthly average squared bond log-prices	Chung et al. [2019]
Bond Term Risk Factor	Estimated 48 monthly rolling returns on FF5 factors, bond term, bond default,	Gebhardt et al. [2005a]
Bond Default Risk Factor	Estimated 48 monthly rolling returns on FF5 factors, bond term, bond default,	Gebhardt et al. [2005a]
Bond Return Momentum	Cumulative Product of lagged 6 months of bond returns	Jostova et al. [2013]
Bond Illiquidity (Amihud [2002])	Monthly average of the ratio of absolute daily bond return over dollar volume of trade	Lin et al. [2011]
Bond Illiquidity (Pastor and Staumbaugh [2003])	Estimated 48 monthly rolling returns on bond term, default, and liquidity factors	Lin et al. [2011]
Stock Illiquidity (Amihud [2002])	Monthly average of the ratio of absolute daily stock return over dollar volume of trade	Lin et al. [2011]
Stock Illiquidity (Pastor and Staumbaugh [2003])	Estimated 48 monthly rolling returns on FF5 and liquidity factors	Lin et al. [2011]
Intermediary Asset Risk	Estimated 48 monthly rolling returns on BBW factors and intermediary factor	He et al. [2017]
Expected Default Frequency	Joint firm asset value and volatility estimation using method in Bharath and Shumway [2008]	Bharath and Shumway [2008]
Change Option implied volatility	Change in end of month ATM implied volatilities	Cao et al. [2020]
Size	log of market capitalization (product of share price and number of common shares outstanding)	Chordia et al. [2017]
Stock Return Reversal	1 month lagged stock return	Chordia et al. [2017]
Net New Investment Accruals	Percentage change in net new investments	Bhojraj and Swaminathan [2009]
Net Working Capital Accruals	Net working capital investment divided by total assets	Bhojraj and Swaminathan [2009]
Total Accruals	Total Accruals	Chichernea et al. [2019]
Change in WC	Working Capital accruals	Chichernea et al. [2019]
Change in NCO	non-current operating accruals	Chichernea et al. [2019]
Change in NOA	net operating assets	Chichernea et al. [2019]
Change in FIN	financial accruals	Chichernea et al. [2019]
Options Illiquidity (ORES)	Monthly averaged dollar-volume weighted option price relative effective spreads	Christoffersen et al. [2018]
Stock Turnover	Daily ratio of trading volume over shares outstanding averaged over the month	Akbas et al. [2017]
Options Order Flow	Delta weighted Option Buy minus Sell Trade Volume over total volume (signed CBOE/ISE volumes)	Bollen and Whaley [2004] and Hu [2014]
O/S Ratio	Monthly average of the daily log option volume to stock volume ratio	Ge et al. [2016]
Option Implied Price Deviation	Joint estimation of option trade price implied stock prices using Manaster and Rendleman [1982]	Cremers et al. [2019]
Option Implied Borrowing Fee	Equation 15 of Muravyev et al. [2020]	Muravyev et al. [2020]
Put to Call Volume Ratio	Customer Open Buy put-to-call volume ratio (signed CBOE/ISE volumes)	Pan and Poteshman [2006]
Risk Neutral Skewness	Volatility surface OTM put less ATM average put and call (with d2mat=30)	Xing et al. [2010]
Deviations from Put-Call Parity	Monthly average Put-call parity deviations weighted by option open interest	Cremers and Weinbaum [2010]

Notes: Description of additional variable levels used in Section ??. Filters and description of data used is detailed in Section 2.3.2.

Table A.1.4: : Data Summary Statistics

Variable Name	N	Mean	Median	Std. Dev.	Percentiles 10th	25th	75th	90th
special items (spi)	1208300	−0.006	−0.001	0.022	−0.018	−0.006	0	0.001
market value of equity fiscal year (mve f)	1286886	71702.63	33260.02	91488.82	3545.5	10657.46	99076.16	190147.2
Book to market (bm)	1286886	0.635	0.542	0.594	0.177	0.312	0.789	1.183
earnings to price (ep)	1286886	0.022	0.061	0.383	0.002	0.042	0.083	0.108
cash productivity (cashpr)	1285921	−3.012	−3.048	29.704	−17.834	−6.518	1.398	12.109
dividend to price (dy)	1282447	0.027	0.024	0.027	0.002	0.013	0.035	0.05
leverage (lev)	1286866	5.254	1.881	7.438	0.414	0.878	7.78	13.37
sales to price (sp)	1286886	1.208	0.706	1.918	0.326	0.457	1.147	2.238
return on invested capital (roic)	1286438	0.05	0.048	0.521	0.017	0.03	0.091	0.148
R& D to sales (rd sale)	465511	0.048	0.026	0.437	0	0.01	0.043	0.086
R& D to market capiitalization (rd mve)	466840	0.048	0.02	0.116	0	0.006	0.039	0.097
asset growth (agr)	1279329	0.068	0.04	0.202	−0.066	−0.013	0.103	0.205
gross profitability (gma)	1279329	0.168	0.108	0.176	0.021	0.042	0.244	0.4
change in shares outstanding (chcsho)	1278524	0.039	0	0.21	−0.054	−0.027	0.015	0.111
growth in long term debt (lgr)	1279273	0.078	0.039	0.261	−0.075	−0.024	0.112	0.229
Working Capital accruals (acc)	1263635	−0.034	−0.03	0.054	−0.085	−0.055	−0.006	0.012
percent accruals (pctacc)	1263635	−2.649	−0.949	7.377	−5.746	−2.077	−0.243	0.619
cash flow to price ratio (cfp)	1271204	0.152	0.116	0.26	−0.001	0.068	0.188	0.364
Absolute Accruals (absacc)	1263635	0.044	0.033	0.047	0.005	0.016	0.059	0.088
change in inventory (chinv)	1257787	0.004	0	0.024	−0.008	−0.001	0.006	0.021
cash flow (cf)	1263635	0.064	0.058	0.064	0	0.02	0.098	0.144
employee growth rate (hire)	1278829	0.023	0.006	0.145	−0.078	−0.03	0.054	0.134
sales growth (sgr)	1278985	0.056	0.036	0.233	−0.12	−0.029	0.118	0.238
ib divided by sales (chpm)	1278947	0.001	0.002	5.534	−0.068	−0.02	0.021	0.066
change in sales /average total assets (chato)	1271212	−0.007	−0.001	0.105	−0.064	−0.019	0.013	0.055
% change sales less % change in inventory (pchsale pchinvt)	1109203	−0.127	0.003	1.045	−0.35	−0.11	0.104	0.313

% change in sales less % change in A/R (pchsale pchrect)	1268558	−0.012	−0.006	0.271	−0.214	−0.084	0.078	0.189
% change in gross margin less % change in sales (pchgm pchsale)	1278985	−0.001	0.004	0.513	−0.141	−0.042	0.063	0.175
% change in sales less % change in SG& A (pchsale pchxsga)	827216	0.001	0.002	0.149	−0.144	−0.051	0.052	0.126
deprecation/PPE (depr)	1194522	0.21	0.163	0.236	0.051	0.098	0.251	0.377
% change in depreciation (pchdepr)	1183894	0.02	0.001	0.223	−0.166	−0.061	0.076	0.185
Change in Advertising expenses (chadv)	487296	0.042	0.015	0.228	−0.167	−0.052	0.115	0.304
capital expenditures and inventory (invest)	1196508	0.03	0.013	0.084	−0.022	0	0.048	0.092
growth in common shareholder equity (egr)	1279302	0.066	0.043	0.482	−0.179	−0.021	0.136	0.308
% change in CAPX (pchcapx)	1163557	0.098	0.042	11.949	−0.375	−0.122	0.22	0.524
growth in capital expenditure (grcapx)	1161921	0.27	0.064	1.694	−0.476	−0.199	0.379	0.992
debt capacity/firm tangibility (tang)	1206314	0.446	0.451	0.146	0.249	0.36	0.522	0.665
current ratio (currat)	1239452	3.463	1.33	6.839	0.776	1.008	2.538	8.733
% change in current ratio (pchcurrat)	1226611	0.045	0	0.356	−0.19	−0.079	0.086	0.254
quick ratio (quick)	1232426	2.925	1.145	6.033	0.581	0.841	1.767	8.345
% change in quick ratio (pchquick)	1218649	0.052	0.003	0.394	−0.214	−0.092	0.098	0.288
sales to cash (salecash)	1286043	23.188	4.722	70.482	0.259	0.825	17.042	55.129
sales to receivables (salerec)	1275785	6.898	3.731	14.994	0.125	0.194	7.859	12.065
sales to inventory (saleinv)	1115067	44.554	9.649	126.501	0.397	5.354	21.613	68.556
percentage change in sales to inventory (pchsaleinv)	1101577	0.08	0.002	0.61	−0.261	−0.101	0.11	0.354
cash flow to debt (cashdebt)	1254023	0.087	0.062	0.14	0.007	0.014	0.134	0.217
real estate holdings (realestate)	221408	0.305	0.269	0.168	0.109	0.191	0.402	0.563
growth in long-term net operating assets (grltnoa)	946232	0.054	0.036	0.115	−0.019	0.004	0.076	0.14
% change in R& D less % change in income over equity (rdbias)	402055	−0.192	−0.144	17.574	−0.439	−0.253	−0.009	0.113

return on equity (roe)	1279302	0.129	0.126	0.491	−0.012	0.071	0.202	0.326
operating profitability (operprof)	1279302	0.662	0.472	1.211	0.144	0.261	0.789	1.428
Industry adjusted change in profit margin (chpmia)	1278947	0.298	−0.003	18.736	−4.295	−0.328	0.411	4.285
industry adjusted change in asset turnover (chatoia)	1271212	0.007	0.003	0.124	−0.084	−0.021	0.044	0.103
industry adjusted change in employees (chempia)	1278829	−0.233	−0.063	1.04	−0.45	−0.158	−0.006	0.064
Industry adjusted book to market (bm ia)	1286886	6.559	0.125	79.608	−3.967	−0.372	0.985	7.29
industry adjusted % change in capital expenditures (pchcapx ia)	1163557	15.175	−0.505	132.642	−3.676	−1.43	−0.068	2.918
tax income to book income (tb)	1105392	−0.047	−0.048	1.732	−1.074	−0.558	0.37	0.886
industry adjusted cash flow to price ratio (cfp ia)	1271204	−2.03	0.06	17.157	−3.036	−0.075	0.266	0.912
industry sales concentration (herf)	1286886	0.074	0.041	0.072	0.025	0.03	0.086	0.189
organizational capital (orgcap)	609700	0.003	0.002	0.004	0	0	0.005	0.008
market value of equity monthly (mve m)	1286886	72328865.8	34605005.77	92978675.88	3443634.34	10951661.08	94645948.73	201889909
lag log price per share (pps)	1286886	3.66	3.723	0.991	2.485	3.156	4.195	4.709
research and developpment quarterly (rdq)	1286613	18921.37	19012	1663.77	16561	17478	20377	21129
price of common shares quarterly (prccq)	1284405	115.361	41.04	2503.57	12.5	23.78	65.58	105.3
change in tax expense (chtx)	1283207	0	0	0.009	−0.003	−0.001	0.001	0.004
return on assets (roaq)	1286154	0.007	0.005	0.019	−0.001	0.002	0.013	0.023
return on equity quarterly (roeq)	1286144	0.029	0.03	0.134	−0.008	0.016	0.048	0.082
revenue surprise (rsup)	1284296	−0.002	0.006	0.138	−0.041	−0.007	0.022	0.051
accrual volatility (stdacc)	674986	1.532	0.09	23.062	0.042	0.063	0.142	0.236
sales growth volatility (sgrvol)	1269965	0.078	0.023	0.743	0.006	0.012	0.051	0.111
earnings volatility (roavol)	1271940	0.008	0.004	0.015	0.001	0.001	0.009	0.018
cash flow volatility (stdcf)	674986	3.204	0.104	51.81	0.048	0.07	0.168	0.322
casj holdings (cash)	1286423	0.107	0.078	0.102	0.011	0.028	0.158	0.25
corporate investment (cinvest)	1201633	−0.004	0	0.626	−0.082	−0.022	0.021	0.083

unexpected quarterly earnings (sue)	1284405	−0.008	0	0.205	−0.006	−0.001	0.002	0.006
Abnormal earnings announcement volume (aeavol)	1286569	0.696	0.514	0.801	−0.01	0.213	0.964	1.541
earnings announcement (ear)	1286572	0.001	0	0.063	−0.061	−0.027	0.029	0.062
dispersion inforecasted EPS (disp)	1263387	0.096	0.026	0.328	0.007	0.013	0.056	0.154
change in forecasted EPS (chfeps)	1272473	0.002	0	0.564	−0.13	−0.02	0.02	0.17
forecasted growth in 5yr EPS (fgr5yr)	1213878	10.327	9.9	7.428	4.31	7	12.6	16.18
Mean Analyst Recom. (MEANREC)	1273363	2.365	2.32	0.412	1.88	2.07	2.63	2.91
Change in Mean Analyst Recom. (chrec)	1273193	−2.359	−2.315	0.434	−2.933	−2.632	−2.052	−1.85
number of analyst covering stock (nanalyst)	1286886	17.776	18	7.495	8	13	23	27
scaled earnings forecast (sfe)	1271208	0.044	0.072	0.929	0.032	0.054	0.093	0.119
Mean Analyst Estimate (MEANEST)	1273647	25.705	2.84	435.834	0.73	1.55	4.88	8.54
6 month momentum (mom6m)	1283840	0.044	0.042	0.244	−0.186	−0.062	0.143	0.258
12 month momentum (mom12m)	1279142	0.092	0.086	0.361	−0.279	−0.078	0.239	0.422
36 month momentum (mom36m)	1259470	0.227	0.201	0.55	−0.381	−0.073	0.47	0.776
1 month momentum (mom1m)	1286886	0.008	0.009	0.102	−0.092	−0.037	0.053	0.101
dollar trading volume (dolvol)	1286692	17.411	17.753	1.343	15.696	16.766	18.427	18.678
change in 6 month momentum (chmom)	1279142	0.003	−0.014	0.408	−0.342	−0.169	0.149	0.343
share turnover (turn)	1286174	1.959	1.476	1.672	0.721	1.003	2.202	3.728
industry momentum (indmom)	1286886	0.095	0.098	0.252	−0.199	−0.046	0.215	0.353
maximum daily return (maxret)	1286886	0.04	0.028	0.045	0.014	0.02	0.043	0.07
return volatility (retvol)	1286886	0.019	0.014	0.017	0.008	0.01	0.021	0.032
bid ask spread (baspread)	1286886	0.025	0.02	0.021	0.012	0.015	0.028	0.042
dollar trading volume (std dolvol)	1286875	0.348	0.324	0.124	0.225	0.266	0.405	0.493
volatility of liquidity share turnover (std turn)	1286886	4.13	2.427	6.134	0.954	1.474	4.297	8.021

beta (beta)	1286286	1.101	1.074	0.547	0.441	0.692	1.416	1.829
beta squared (betasq)	1286286	1.512	1.153	1.468	0.195	0.479	2.004	3.345
adjusted rsquared (rsq1)	1286286	0.322	0.328	0.163	0.102	0.192	0.455	0.526
Price Delay (pricedelay)	1286286	0.042	0.024	0.308	−0.077	−0.017	0.079	0.191
idiosyncratic return volatility (idiovol)	1286286	0.035	0.028	0.022	0.019	0.023	0.04	0.059
change in WC accurals (d WCA)	473953	−10.766	0	376.334	−498	−165	156	431
change in non-current operating accruals (d NCO)	506669	185.497	62	614.241	−308	−58.2	382.9	876
change in financial accruals (d FIN)	478842	−66.813	−1.169	572.343	−755	−273	163	541.183
EDF (EDF)	1173808	0.066	0	0.185	0	0	0.003	0.22
opt implied fee (implied fee)	1103363	−0.001	−0.001	0.014	−0.011	−0.005	0.004	0.009
CW (cw)	1132525	−0.004	−0.003	0.023	−0.016	−0.008	0.003	0.01
OS ratio (os)	1088493	0.15	0.125	0.124	0.018	0.052	0.208	0.327
IPD (ipd)	1095765	0	0	0.005	−0.003	−0.001	0.001	0.003
OOI (hu oi)	1092563	0	0	0	0	0	0	0
PP (potesh)	1036607	0.463	0.428	0.204	0.234	0.318	0.565	0.751
change in ivol (d ivol)	1051163	−0.001	−0.002	0.151	−0.132	−0.06	0.055	0.127
CVOL (cvol)	1052260	−0.001	−0.001	0.079	−0.068	−0.032	0.029	0.065
PVOL (pvol)	1052346	−0.001	−0.001	0.083	−0.072	−0.032	0.03	0.068
RN Skewness (skewness)	1054619	0.056	0.042	0.047	0.022	0.03	0.064	0.106
IVOL Level (avg ivol)	1016960	0.308	0.25	0.191	0.165	0.198	0.347	0.499
Bond short term reversal (bond st rev)	871705	0.006	0.004	0.04	−0.019	−0.005	0.016	0.033
bond momentum (bond mom)	247778	0.036	0.022	0.12	−0.026	0.003	0.055	0.105
bond downside risk (bond Var5)	438178	0.041	0.029	0.042	0.011	0.017	0.047	0.076
bond Amihud (bond Amihud)	1120061	0.015	0.003	0.172	0	0.001	0.01	0.026
Std Dev. Log Price (Std. log Price)	1208997	0.012	0.008	0.019	0.002	0.004	0.013	0.022
bond VIX Beta (d VIX beta)	703401	−0.001	−0.001	0.005	−0.006	−0.003	0.001	0.004
bond Int Beta (int cap rf beta)	692589	0	−0.002	0.168	−0.164	−0.066	0.066	0.164
bond idio vol (bond RMSE)	706716	0.023	0.017	0.021	0.007	0.011	0.028	0.044
bond UNC Beta (UNC 12 beta)	697681	−0.127	−0.002	1.262	−1.453	−0.539	0.442	1.064
bond PS beta (PS VWF beta)	704417	−0.017	−0.015	0.269	−0.287	−0.124	0.102	0.265
Mac Duration (Mac Dur)	1225654	12.425	9.613	6.788	4.915	7.7	17.974	23.674

Skewness Log Price (Bond Skewness)	1107368	−0.124	−0.092	0.926	−1.354	−0.676	0.448	1.027
Kurtosis Log Price (Bond Kurtosis)	1019831	0.302	−0.186	2.144	−1.54	−0.971	1.125	2.912
Default Beta (Default Beta)	717440	0.556	0.394	0.991	−0.18	0.103	0.87	1.492
Term Beta (Term Beta)	717369	0.414	0.369	0.411	0.004	0.175	0.621	0.957
Liquidity Beta (Liquidity Beta)	703763	−0.054	0.014	1.453	−1.283	−0.4	0.454	1.272
bond idio vol (bond RMSE)	713931	0.023	0.018	0.019	0.008	0.011	0.027	0.042
Bond Beta to Stock Factor (BBeta SMrkt)	717565	0.094	0.068	0.321	−0.195	−0.045	0.222	0.439
Bond Beta to RMW (BBeta RMW)	716820	−0.054	−0.031	0.58	−0.613	−0.252	0.161	0.486
Bond Beta to CMA (BBeta CMA)	717766	−0.095	−0.044	0.738	−0.737	−0.318	0.187	0.532
Bond Beta to HML (BBeta HML)	717718	0.067	0.024	0.454	−0.342	−0.125	0.218	0.529
Bond Beta to SMB (BBeta SMB)	716825	0.014	−0.007	0.36	−0.323	−0.134	0.136	0.371
bond idio kurt (bond idio kurt)	719248	0.024	0.019	0.021	0.008	0.012	0.029	0.046
bond idio skew (bond idio skew)	719464	0.025	0.019	0.022	0.008	0.012	0.029	0.047

Notes: We compute 143 stock, balance sheet, and earnings signals from the data set used in Green et al. [2017] (GHZ) augmented with those signals in Table A.1.3 from July 2002 to October 2019. We thank Jeremiah Green for making his SAS code freely available on his website. The list of firm characteristics from GHZ is constructed using all firms with common shares that are listed on the AMEX, NYSE, or NASDAQ, that have end of month value on CRSP, quarterly and annual balance sheet reporting on COMPUSTAT and earnings information reported to the $I/B/E/S$ data. The GHZ data set is available beginning from January 1980, however, we only require data beginning from July 2002 when TRACE corporate bond reporting began. The summary statistics of the GHZ data is only of those firms that have corporate bonds outstanding.

Table A.1.5: : Corporate Bond Factor Classification

Risk Type	Risk Category	Description	Paper Reference
Common	Financial	Bond Market Risk Factor	Bai et al. [2019b]
Common	Financial	Bond Credit Risk Factor	Bai et al. [2019b]
Common	Microstructure	Bond Liquidity Risk Factor (Roll [1984])	Bai et al. [2019b]
Common	Financial	Bond Downside Risk Factor	Bai et al. [2019b]
Common	Other	Bond Reversal Risk Factor	Bai et al. [2019b]
Common	Financial	Bond Term Risk Factor	Gebhardt et al. [2005a]
Common	Financial	Bond Default Risk Factor	Gebhardt et al. [2005a]
Common	Financial	Bond VIX Factor (Ang et al. [2006])	Chung et al. [2019]
Common	Microstructure	Intermediary Factor	He et al. [2017]
Common	Microstructure	Bond Illiquidity Factor (Pastor and Staumbaugh [2003])	Lin et al. [2011]
Common	Microstructure	Stock Illiquidity Factor (Pastor and Staumbaugh [2003])	Lin et al. [2011]
Common	Microstructure	Bond Permanent Component Factor (Sadka [2006])	Lin et al. [2011]
Common	Microstructure	Bond Temporary Component Factor (Sadka [2006])	Lin et al. [2011]
Common	Financial	Systematic bond risk	Bai et al. [2020c]
Common	Financial	Idiosyncratic bond risk	Bai et al. [2020c]
Common	Macroeconomic	Bond Uncertainty Risk (Jurado et al. [2015])	Bai et al. [2020a]
Common	Behavioural	Bond Mispricing Factor	Avramov et al. [2019]
Characteristic	Other	Short Term Bond Reversal	Bai et al. [2019a]
Characteristic	Other	Long Term Bond Reversal	Bai et al. [2019a]
Characteristic	Other	Bond Momentum	Jostova et al. [2013]
Characteristic	Financial	Bond Total Volatility	Chung et al. [2019]
Characteristic	Financial	Bond Idiosyncratic Volatility	Chung et al. [2019]
Characteristic	Financial	Stock Total Volatility	Chung et al. [2019]
Characteristic	Financial	Stock Idiosyncratic Volatility	Chung et al. [2019]
Characteristic	Microstructure	Bond Illiquidity (Amihud [2002])	Lin et al. [2011]
Characteristic	Microstructure	Stock Illiquidity (Amihud [2002])	Lin et al. [2011]
Characteristic	Financial	Change in Option implied volatility (An et al. [2010])	Cao et al. [2020]
Characteristic	Financial	Change in Call Option implied volatility (An et al. [2010])	Cao et al. [2020]
Characteristic	Financial	Change in Put Option implied volatility (An et al. [2010])	Cao et al. [2020]
Characteristic	Financial	Level of Option implied volatility	Cao et al. [2020]
Characteristic	Financial	Exponential GARCH (Glosten et al. [1993])	Cao et al. [2020]
Characteristic	Accounting	Default Risk (Bharath and Shumway [2008])	Chordia et al. [2017]
Characteristic	Accounting	Size	Chordia et al. [2017]
Characteristic	Other	Stock Momentum	Chordia et al. [2017]

Characteristic	Other	Stock Return Reversal	Chordia et al. [2017]
Characteristic	Accounting	Profitability	Chordia et al. [2017]
Characteristic	Accounting	Accruals	Chordia et al. [2017]
Characteristic	Behavioural	Earnings Surprise	Chordia et al. [2017]
Characteristic	Accounting	Asset Growth	Choi and Kim [2018]
Characteristic	Accounting	Investment	Choi and Kim [2018]
Characteristic	Accounting	Gross Profitability	Choi and Kim [2018]
Characteristic	Accounting	Net Issuance	Choi and Kim [2018]
Characteristic	Accounting	Net Book Equity to Market Equity	Choi and Kim [2018]
Characteristic	Accounting	Net New Investment Accruals	Bhojraj and Swaminathan [2009]
Characteristic	Accounting	Net Working Capital Accruals	Bhojraj and Swaminathan [2009]
Characteristic	Accounting	Total Accruals	Chichernea et al. [2019]
Characteristic	Accounting	Change in WC	Chichernea et al. [2019]
Characteristic	Accounting	Change in NCO	Chichernea et al. [2019]
Characteristic	Accounting	Change in NOA	Chichernea et al. [2019]
Characteristic	Accounting	Change in FIN	Chichernea et al. [2019]
Characteristic	Microstructure	Bond Effective Spreads	Anderson et al. [2018]
Characteristic	Microstructure	Bond Lending Fee	Anderson et al. [2018]
Characteristic	Microstructure	Bond Utilization Fee	Anderson et al. [2018]
Characteristic	Microstructure	Unconstrained Insurance Bond Holding	Murray and Nikolova [2019]
Characteristic	Microstructure	Constrained Insurance Bond Holding	Murray and Nikolova [2019]
Characteristic	Microstructure	High Insurance Bond Holding	Becker and Ivanisha [2015]
Characteristic	Microstructure	Low Insurance Bond Holding	Becker and Ivanisha [2015]
Characteristic	Microstructure	Bond Supply	Goldberg and Nozawa [2019]
Characteristic	Microstructure	Bond Demand	Goldberg and Nozawa [2019]
Characteristic	Behavioural	Good News (earnings greater than forecast)	Defond and Zhang [2014]
Characteristic	Behavioural	Bad News (earnings less than forecast)	Defond and Zhang [2014]
Characteristic	Behavioural	Change in Earnings	Easton et al. [2009]
Characteristic	Behavioural	Negative News Dummy	Easton et al. [2009]
Characteristic	Behavioural	Negative News and Change in Earnings	Easton et al. [2009]
Characteristic	Financial	OLS Moving Average Bond Return	Lin et al. [2020]
Characteristic	Financial	Elastic Net Moving Average Bond Return	Lin et al. [2020]
Characteristic	Financial	Trend Moving Average Bond Return	Lin et al. [2019]
Characteristic	Financial	OLS Predicted Bond Return	Bali et al. [2020]
Characteristic	Financial	PCA Predicted Bond Return	Bali et al. [2020]
Characteristic	Financial	PLS Predicted Bond Return	Bali et al. [2020]
Characteristic	Financial	LASSO Predicted Bond Return	Bali et al. [2020]
Characteristic	Financial	RIDGE Predicted Bond Return	Bali et al. [2020]

Characteristic	Financial	ENET Predicted Bond Return	Bali et al. [2020]
Characteristic	Financial	RF Predicted Bond Return	Bali et al. [2020]
Characteristic	Financial	FFN Predicted Bond Return	Bali et al. [2020]
Characteristic	Financial	LSTM Predicted Bond Return	Bali et al. [2020]
Characteristic	Financial	Combination Predicted Bond Return	Bali et al. [2020]
Characteristic	Financial	Bond Book to Market Ratio	Bartram et al. [2020]
Characteristic	Financial	Bond Return Total Skewness	Bai et al. [2020b]
Characteristic	Financial	Bond Return Total Kurtosis	Bai et al. [2020b]
Characteristic	Financial	Bond Return Coskewness	Bai et al. [2020b]
Characteristic	Financial	Bond Return Cokurtosis	Bai et al. [2020b]
Characteristic	Financial	Bond Return Idiosyncratic Skewness	Bai et al. [2020b]
Characteristic	Financial	Bond Return Idiosyncratic Kurtosis	Bai et al. [2020b]
Characteristic	Behavioural	Merger Bids	Asquith and Kim [1982]
Characteristic	Behavioural	Leveraged Buyouts	Asquith and Wizman [1990]
Characteristic	Behavioural	Merger Bids	Billet et al. [2004]
Characteristic	Behavioural	Poison Puts	Cook and Easterwood [1994]
Characteristic	Behavioural	Super Poison Puts	Crabbe [1991]
Characteristic	Behavioural	Common Share Repurchases	Defusco et al. [1990]
Characteristic	Behavioural	Executive Option SEC Stamp Date	Dennis and McConnell [1986]
Characteristic	Behavioural	Dividend Increases	Dhillon and Johnson [1994]
Characteristic	Behavioural	Dividend Decreases	Dhillon and Johnson [1994]
Characteristic	Behavioural	Pure Stock Exchange Merger	Eger [1983]
Characteristic	Behavioural	Expected Downgrade	Hand et al. [1992]
Characteristic	Behavioural	Unexpected Downgrade	Hand et al. [1992]
Characteristic	Behavioural	Expected Upgrade	Hand et al. [1992]
Characteristic	Behavioural	Unexpected Upgrade	Hand et al. [1992]
Characteristic	Behavioural	Rating Downgrade IG	Hand et al. [1992]
Characteristic	Behavioural	Rating Downgrade SG	Hand et al. [1992]
Characteristic	Behavioural	Rating Upgrade IG	Hand et al. [1992]
Characteristic	Behavioural	Rating Upgrade SG	Hand et al. [1992]
Characteristic	Behavioural	Dividend Announcement	Handjinicolaou and Kalay [1984]
Characteristic	Behavioural	Corporate Spin off Announcements	Hite and Owers [1983]
Characteristic	Behavioural	Going Private	Marais et al. [2003]
Characteristic	Behavioural	Special Dividend Announcement	Narayanan and Shastri [1988]
Characteristic	Behavioural	Leveraged Buyout	Warga and Welch [1993]
Characteristic	Behavioural	Unexpected Dividend Increase	Woolridge [1983]
Characteristic	Behavioural	Unexpected Dividend Decrease	Woolridge [1983]

Notes: This table presents a list of all signals (and which paper they were discovered in) that have been discovered in the cross-section of corporate bond returns. As in Harvey et al. [2016], we count common risk factors as well as characteristics even though as HLZ note, characteristics are not risk factors. If firm's characteristic is correlated with the CS of corporate bond returns a long-short portfolio can typically be formed in order to proxy for the unknown risk factor and hence in this form the characteristic can be considered as a risk factor. Note that the following papers only report p-values or asterixes indicating levels of significance and not t-statistics: Dann [1981], Eberhart and Siddique [2002], Kim and McConnell [1977], Mansi and Reeb [2002], Maxwell and Stephens [2003], Maxwell and Rao [2003] and Parrino [1997].

Table A.1.6: : Empirically Simulated Benchmark t-statistics when Number of Discovered Signals is Estimated

	M		Bonf.		Holm		$\text{BHY}_{\alpha=1\%}$		BHY	
S. R. /C.I.	10%	90%	10%	90%	10%	90%	10%	90%	10%	90%
$r = 1$	**288**		**3.76**		**3.74**		**3.81**		**3.13**	
	207	388.5	3.68	3.86	3.68	3.86	3.68	3.95	2.97	3.33
$r = 3/2$	**421**		**3.83**		**3.82**		**3.81**		**3.12**	
	322	537	3.76	3.92	3.76	3.92	3.76	3.95	3.01	3.27
$r = 2$	**555**		**3.90**		**3.87**		**3.82**		**3.12**	
	441	689	3.80	3.97	3.80	3.97	3.61	3.97	3.03	3.30

Notes: Simulated t-statistic benchmarks from empirical distribution of discovered t-statistics with level of significance of 5% and under of the Bonferroni, Holm, and BHY multiple hypothesis testing methods. M is the estimated number of discovered signals.

Table A.1.7: : Univariate FM regressions and PS

	FM (BBW)		FM (BBW+FF5)		VW FM (BBW)		Decile PS (BBW)		Quintiles (BBW)		Quintiles (BBW+FF5)	
Variable Name	λ_s	t_{λ_s}	λ_s	t_{λ_s}	λ_s	t_{λ_s}	α_s	t_{α_s}	α_s	t_{α_s}	α_s	t_{α_s}
spi	−0.0018	−0.26	0.0036	0.46	0.0051	0.63	0.2643	1.99	0.1444	1.65	0.1393	1.96
mve f	0	1.26	0	0.76	0	1.37	0.3934	2.42	0.1927	2.08	0.2003	2.36
bm	−0.0008	−0.85	−0.0008	−0.8	0	0.01	0.1644	1.1	0.0645	0.59	0.0754	0.73
ep	0.0017	0.98	0.0031	1.43	0.0027	0.52	0.3017	3.23	0.133	2.09	0.067	1.06
cashpr	0	−0.86	0	−0.97	0	0.03	−0.0642	−0.63	0.0267	0.48	0.0409	0.72
dy	0.008	0.4	−0.0038	−0.18	0.0133	0.5	0.2592	2.08	0.2045	2.18	0.1584	2.01
lev	0	−0.21	−0.0001	−0.98	0	0.4	−0.2299	−1.38	−0.1411	−1.12	−0.0788	−0.61
sp	0	−0.08	−0.0002	−0.77	−0.0004	−1.31	−0.4872	−2.84	−0.1655	−2.55	−0.141	−2.09
roic	0.0042	2.46	0.0065	2.82	0.003	0.53	0.0862	1.1	0.0959	0.93	0.0739	0.72
rd sale	0.0027	0.69	0.0023	0.6	0.0025	1.11	0.1511	3.35	−0.034	−0.48	0.03	0.33
rd mve	−0.0055	−1.01	−0.003	−0.46	−0.0019	−0.27	−0.5894	−2.12	−0.3864	−2.57	−0.3034	−1.85
agr	0.0011	1.26	0.0013	1.2	0.0012	0.97	−0.1165	−1.79	0.0057	0.07	−0.0251	−0.32
gma	0.0016	1.29	0.0026	1.84	0.0002	0.12	0.1226	0.85	0.0231	0.17	−0.0208	−0.17
chcsho	−0.0002	−0.21	0.0004	0.46	0.0009	0.74	−0.2127	−2.2	−0.108	−1.83	−0.0458	−0.76
lgr	0.0008	1.44	0.0009	1.05	0.0011	1.02	−0.0733	−1.16	−0.0013	−0.02	−0.0225	−0.35
acc	0.0049	1.25	0.006	1.36	0.0067	1.15	0.3091	1.87	0.2635	1.88	0.2156	1.85
pctacc	0	0.62	0	0.38	0	0.19	0.4393	2.63	0.2965	2.69	0.2426	2.68
cfp	−0.0006	−0.34	−0.0012	−0.73	0.0008	0.5	−0.5911	−2.16	−0.3851	−2.26	−0.2968	−2.26
absacc	−0.0063	−1.49	−0.0081	−1.73	−0.008	−0.95	−0.2648	−1.64	−0.1351	−1.47	−0.1166	−1.34
chinv	−0.0004	−0.09	0.0019	0.33	0.0132	1.72	−0.2568	−2.46	0.0171	0.32	−0.0133	−0.22
cf	0.0024	0.52	0.0071	1.34	−0.0003	−0.05	−0.134	−0.85	−0.0942	−0.72	−0.0957	−0.77
hire	0.0014	1.42	0.0021	2.03	0.0008	0.66	0.164	1.89	0.0676	1.12	0.0453	0.73
sgr	0.0016	1.63	0.0024	2.24	0.0012	0.92	−0.1171	−1.66	0.0109	0.14	−0.0041	−0.06
chpm	0.0003	0.48	0.0003	0.33	0.0009	0.7	−0.1537	−1.13	−0.0801	−0.6	−0.0326	−0.28
chato	0.0004	0.32	−0.0002	−0.14	0.0012	0.77	−0.0461	−0.88	−0.0816	−1.37	−0.033	−0.59
pchsale pchinvt	−0.0003	−0.8	−0.0001	−0.12	−0.0003	−0.71	0.0274	0.29	−0.036	−0.59	−0.0044	−0.06
pchsale pchrect	0.0001	0.1	−0.0002	−0.24	−0.0001	−0.08	−0.1314	−2.93	−0.0494	−0.78	−0.0533	−0.84
pchgm pchsale	0.0014	1.48	0.0009	0.95	0.0009	0.57	−0.0911	−0.61	−0.1062	−0.85	−0.028	−0.27
pchsale pchxsga	−0.0023	−1.46	−0.0005	−0.31	−0.0015	−0.64	−0.1584	−1.3	−0.1096	−1.36	−0.138	−1.69
depr	0	0.04	−0.0001	−0.17	0.0005	0.29	−0.0365	−0.84	0.046	0.97	0.0645	1.24
pchdepr	−0.0014	−1.55	−0.0008	−0.96	−0.002	−1.81	−0.0593	−0.59	−0.05	−0.82	−0.0884	−1.32
chadv	0.002	2.11	0.0029	3	0.0009	0.84	−0.1634	−1.51	−0.024	−0.27	0.0216	0.28
invest	0	0.01	0.0014	0.69	0.003	1.1	−0.1307	−1.87	−0.0508	−0.64	−0.0904	−1.01

egr	−0.0003	−0.45	−0.0001	−0.18	0.0004	0.5	0.0836	0.77	−0.0029	−0.04	0.0181	0.25
pchcapx	0.0003	1.03	0.0006	2.19	0.0001	0.12	−0.0539	−0.36	0.0349	0.37	0.009	0.09
grcapx	0.0002	1.06	0.0001	0.4	0.0006	1.05	0.2337	1.73	0.1605	1.88	0.098	1.33
tang	0.0009	0.39	0.0003	0.13	0.0006	0.29	0.1071	0.57	0.016	0.14	0.0211	0.21
currat	−0.0001	−1.73	−0.0001	−2.27	−0.0001	−2.18	−0.2589	−1.33	−0.151	−1.39	−0.0932	−0.83
pchcurrat	−0.0008	−2.4	−0.001	−2.1	−0.0003	−0.5	0.0348	0.66	0.0084	0.14	0.0463	1.01
quick	−0.0001	−1.53	−0.0001	−1.95	−0.0001	−2.24	−0.3765	−2.24	−0.1593	−1.39	−0.1038	−0.88
pchquick	−0.0008	−2.65	−0.0008	−1.96	−0.0004	−0.73	0.0438	1.03	0.0025	0.05	0.0363	0.86
salecash	0	−0.68	0	0.55	0	−0.76	−0.1657	−1.07	−0.1544	−1.12	−0.1929	−1.48
salerec	0	0.35	0	0.77	0	0.11	−0.2405	−1.17	−0.1948	−1.24	−0.2038	−1.46
saleinv	0	−0.26	0	−0.1	0	−0.69	−0.0783	−0.62	−0.1114	−1.04	−0.1171	−1.3
pchsaleinv	−0.0006	−1.24	−0.0005	−0.97	−0.0007	−1.42	0.012	0.13	−0.0635	−1.06	−0.0292	−0.43
cashdebt	0.0017	1.1	0.0027	1.5	0.0009	0.28	−0.025	−0.17	0.0801	0.52	0.0403	0.27
realestate	0.0008	0.66	0	0	0.0019	2.15	0.3157	2.25	0.0933	1.49	0.0784	1.51
grltnoa	−0.0008	−0.51	0.0015	0.87	0.0001	0.1	0.072	0.57	−0.0552	−0.52	−0.0663	−0.66
rdbias	−0.0002	−0.59	0.0002	0.5	−0.0006	−1.46	0.8306	2.48	0.3322	1.74	0.1754	1.07
roe	0.0007	1.32	0.0004	0.58	0.0016	1.97	−0.1791	−0.98	−0.0069	−0.07	0.027	0.3
operprof	−0.0001	−0.19	0.0002	0.65	−0.0003	−0.79	0.198	1.41	0.0938	0.71	0.0315	0.26
chpmia	0	−0.19	0	−0.09	0	−0.24	−0.0183	−0.36	0.0441	0.67	−0.0111	−0.13
chatoia	0	−0.02	−0.0006	−0.39	−0.0004	−0.17	0.0046	0.06	0.0143	0.17	0.0781	1.21
chempia	0.0001	0.13	0.0003	0.53	0.0006	1.07	0.0324	0.51	0.0972	1.49	0.0728	1.23
bm ia	0	−1.98	0	−1.18	0	−0.04	−0.0768	−0.63	−0.0365	−0.51	−0.0126	−0.18
pchcapx ia	0	0.7	0	0.43	0.0001	2.75	0.0186	0.21	−0.0664	−0.91	−0.04	−0.6
tb	0	0.15	−0.0001	−0.92	0.0002	1.29	−0.0252	−0.49	0.0096	0.14	0.03	0.51
cfp ia	0.0001	0.85	0.0001	1.52	−0.0001	−1.54	0.0883	0.73	0.0149	0.2	0.046	0.53
herf	−0.0007	−0.39	−0.0011	−0.52	−0.0026	−0.99	0.0278	0.59	−0.0342	−0.57	−0.0229	−0.4
orgcap	0.0711	1.12	0.12	1.69	−0.0201	−0.18	0.2747	2.83	0.1601	1.59	0.1029	1.08
mve m	0	2.09	0	1.3	0	1.98	0.3683	3.06	0.2912	2.57	0.2861	2.74
pps	0.0004	1.14	0.0006	1.71	0.0002	0.78	0.403	2.86	0.3302	2.76	0.2745	2.48
rdq	0	−1.77	0	−1.14	−0.0001	−2.25	−0.3371	−2.1	−0.2047	−1.79	−0.2014	−1.94
prccq	0	1.36	0	1.66	0	0.49	0.3335	2.62	0.25	2.39	0.1986	2.03
chtx	0.0552	1.81	0.0772	2.39	0.0516	1.32	0.1521	1.18	0.0599	0.77	0.052	0.64
roaq	0.0029	0.21	0.013	0.9	0.0041	0.18	0.4647	2.84	0.2453	2.21	0.1979	1.75
roeq	−0.0016	−0.6	−0.0015	−0.58	0.0011	0.27	0.1251	0.83	0.1698	1.52	0.1685	1.43
rsup	−0.0002	−0.07	0.0018	0.52	0	−0.01	0.0079	0.05	0.0424	0.36	0.0099	0.08
stdacc	−0.0001	−0.64	0.0001	0.88	−0.0002	−1.46	−0.2147	−2.47	−0.161	−2.04	−0.1663	−2.15
sgrvol	−0.0023	−0.81	−0.0031	−0.92	−0.0047	−0.76	−0.412	−3.21	−0.2765	−3.22	−0.2139	−2.99
roavol	−0.0255	−1.54	−0.0345	−2.03	−0.0349	−1.42	−0.26	−1.87	−0.1314	−0.99	−0.1308	−1.03

stdcf	−0.0001	−0.93	0.0001	1.17	−0.0002	−1.62	−0.3906	−2.15	−0.2521	−2.26	−0.246	−2.37
cash	0.0009	0.33	−0.0001	−0.04	0.001	0.47	0.117	1	0.0029	0.03	0.0566	0.59
cinvest	0.0001	0.42	0.0003	0.77	0	0.01	0.0906	0.6	0.0389	0.44	−0.0228	−0.25
sue	0.0057	1.15	0.0054	0.96	−0.0235	−1.84	−0.13	−0.76	−0.1199	−0.85	−0.1278	−0.98
aeavol	0	−0.14	−0.0001	−0.46	−0.0003	−0.72	−0.0823	−1.3	−0.1714	−1.58	−0.1305	−1.42
ear	−0.0055	−1.27	−0.0044	−1.15	−0.0129	−1.94	0.1208	1.54	0.0551	0.85	0.0601	0.89
disp	−0.0023	−2	−0.0028	−2.3	−0.0049	−1.68	−0.5894	−2.64	−0.4265	−3.23	−0.3753	−3.12
chfeps	0.0015	2.09	0.0014	1.79	0.0013	1.5	0.5442	2.44	0.3677	3.58	0.3389	3.75
fgr5yr	0	−0.63	0	0.15	−0.0001	−0.77	0.0125	0.14	−0.0611	−0.75	−0.0366	−0.5
MEANREC	0.0007	0.87	0	0	−0.0002	−0.22	−0.2572	−2.26	−0.1555	−2.86	−0.1219	−1.8
chrec	−0.0009	−1.63	−0.0001	−0.15	0.0008	1.02	0.0559	0.61	0.1093	2.06	0.0791	1.53
nanalyst	0	0.54	0	0.81	0	1.27	−0.0805	−1.24	−0.0214	−0.43	0.0144	0.3
sfe	0.0016	2.48	0.0025	2.62	0.0022	1.52	0.3381	2.37	0.1642	1.92	0.114	1.43
MEANEST	0	2.05	0	2	0	2.61	0.2524	2.01	0.2377	2.26	0.1892	2.02
mom6m	0.0022	0.98	0.0029	1.31	0.0008	0.3	0.5037	2.51	0.3763	3.16	0.3542	3.03
mom12m	0.0009	0.48	0.0023	1.26	0.0001	0.07	0.4596	2.25	0.3772	3.12	0.3467	2.78
mom36m	0.0005	0.72	0.0021	2.49	0.0001	0.11	0.1252	1.21	0.0701	0.9	0.0677	0.91
mom1m	0.0202	4.74	0.0205	4.84	0.0232	3.98	0.5676	3.11	0.5761	3.38	0.5442	3.96
dolvol	0	−0.19	−0.0001	−0.35	0.0002	0.48	−0.0771	−0.47	0.1196	1.22	0.1359	1.46
chmom	0.0022	2.3	0.0023	2.17	0.0011	0.81	0.4653	3.69	0.2289	2.55	0.2517	2.49
turn	−0.0004	−1.9	−0.0006	−2.4	−0.0003	−0.8	−0.2556	−1.77	−0.1889	−2.16	−0.1503	−1.55
indmom	0.0015	0.78	0.0042	2.25	0.0002	0.08	0.1463	1.04	0.2509	2.45	0.2198	2.19
maxret	−0.0135	−1.2	−0.0176	−1.51	−0.0168	−1.04	−0.0634	−0.35	−0.1027	−0.84	−0.0925	−0.76
retvol	−0.0932	−2.65	−0.1142	−3.25	−0.0747	−1.6	−0.2102	−0.94	−0.1552	−1.13	−0.1607	−1.19
baspread	−0.0922	−2.82	−0.1161	−3.49	−0.0733	−1.59	−0.0426	−0.22	−0.184	−1.25	−0.1746	−1.19
std dolvol	0.0019	0.69	0.0022	0.87	−0.0004	−0.09	0.2531	2.08	0.1624	1.56	0.1383	1.55
std turn	−0.0001	−1.07	−0.0001	−1.41	−0.0002	−1.64	−0.0097	−0.06	−0.1168	−1.17	−0.1109	−1.09
beta	−0.0011	−1.46	−0.0017	−2.17	−0.0002	−0.19	−0.5372	−2.47	−0.2519	−2.11	−0.1765	−1.51
betasq	−0.0004	−1.4	−0.0006	−1.91	−0.0001	−0.26	−0.5378	−2.47	−0.2522	−2.12	−0.1767	−1.52
rsq1	0.0004	0.22	−0.0018	−1	−0.0008	−0.25	0.0127	0.16	−0.0605	−0.95	−0.0203	−0.29
pricedelay	−0.0009	−0.92	−0.0001	−0.07	0.0031	0.97	0.0989	1.38	0.2501	2.21	0.2644	2.61
idiovol	−0.0553	−2.41	−0.0671	−2.65	−0.0167	−0.53	−0.4632	−2.19	−0.358	−2.37	−0.2903	−2.33
d WCA	0	1.34	0	1.7	0	1.01	0.1366	0.8	0.0693	0.66	0.0361	0.45
d NCO	0	2.33	0	2.19	0	2.21	0.3301	2.06	0.2334	2.29	0.2031	2.22
d FIN	0	0.85	0	−1.15	0	1.22	0.0593	0.4	0.0338	0.4	0.0512	0.71
EDF	−0.0084	−2.58	−0.0104	−2.8	−0.0081	−2.08	−0.4469	−2.04	−0.3559	−2.2	−0.2963	−1.98
implied fee	−0.0201	−1.15	−0.0272	−1.39	−0.0181	−0.68	−0.1747	−1.46	−0.135	−2.22	−0.1242	−1.97
cw	0.0121	0.96	0.0113	0.83	0.0343	1.51	0.2007	1.69	0.1115	1.16	0.1363	1.53

os	−0.0028	−1.42	−0.0024	−1.11	−0.0006	−0.22	0.0839	0.84	0.1318	1.17	0.0912	0.94
ipd	0.023	0.34	0.0617	0.89	−0.0585	−0.65	0.2342	2.38	0.1316	1.61	0.1031	1.32
hu oi	−0.5278	−0.74	−0.4152	−0.54	−1.3327	−0.89	−0.0236	−0.24	−0.1498	−2.24	−0.1254	−2.22
potesh	0.0001	0.06	0.0012	0.9	−0.002	−0.58	0.0836	0.71	0.2169	1.55	0.1192	1.25
d ivol	−0.0073	−2.6	−0.0053	−1.86	−0.0121	−3.34	−0.5593	−3.24	−0.5383	−4.85	−0.5493	−4.58
cvol	−0.0113	−2.04	−0.0063	−1.06	−0.0233	−3.37	−0.3999	−2.37	−0.412	−4.57	−0.423	−4.52
pvol	−0.0111	−1.82	−0.0096	−1.7	−0.0235	−3.55	−0.1537	−0.75	−0.2353	−2.04	−0.2802	−2.23
skewness	−0.0064	−0.75	−0.0131	−1.36	−0.0222	−1.84	−0.0283	−0.18	0.0098	0.07	−0.0016	−0.01
avg ivol	−0.0082	−2.35	−0.0091	−2.7	−0.0044	−0.92	−0.1453	−0.62	−0.2333	−1.65	−0.1834	−1.36
NI monthly	0.0061	0.16	0.0198	0.53	0.0182	0.36	0.0095	0.06	−0.1614	−1.48	−0.1399	−1.4
bond st rev	−0.2196	−12.77	−0.207	−9.85	−0.3085	−16.57	−0.8423	−3.01	−0.4687	−2.89	−0.4188	−2.73
bond mom	−0.0206	−1.26	−0.0178	−1	−0.0878	−4.6	0.8425	2.47	0.4797	2.03	0.3658	1.43
bond Var5	−0.0418	−1.99	−0.0455	−1.94	−0.0092	−0.6	−0.4903	−2.27	−0.3699	−2.46	−0.2808	−2.29
bond Amihud	0.0441	1.76	0.0489	7.18	0.0591	2.64	0.0015	1.44	0.0006	0.92	0.0004	0.62
Std. log Price	0.0461	1.34	0.0401	1.04	0.0847	3.01	−0.1558	−0.77	−0.1625	−1.44	−0.1814	−1.6
d VIX beta	0.2847	4.8	0.263	4.26	0.2081	3.37	−0.0885	−0.72	0.005	0.07	0.0652	0.84
int cap rf beta	0.0075	2.28	0.006	1.67	0.0068	2.01	0.132	1.04	0.0735	0.87	0.0994	1.19
bond RMSE	−0.0498	−1.21	−0.0822	−1.8	0.0158	0.44	−0.4679	−2.85	−0.4022	−2.81	−0.3351	−2.41
UNC 12 beta	0.0013	2.79	0.0014	2.8	0.0007	2	0.3955	2.23	0.3288	2.27	0.3183	2.44
PS VWF beta	−0.0013	−1.1	−0.0019	−1.15	0	0.03	−0.26	−1.41	−0.1898	−1.55	−0.1682	−1.53
Mac Dur	−0.0005	−2.48	−0.0004	−1.99	−0.0004	−2.65	−0.2551	−1.44	−0.1095	−0.97	−0.1166	−1.08
Bond Kurtosis	0.0001	2.55	0.0001	2.33	0.0001	3.67	0.2301	2.23	0.1159	2.35	0.1064	2.17
Bond Skewness	−0.0006	−2.96	−0.0005	−2.58	−0.0007	−4.05	−0.3419	−3.88	−0.2227	−3.55	−0.2006	−3.59
Default Beta	−0.0028	−3.41	−0.0023	−2.75	−0.0022	−3.09	−0.2214	−1.74	−0.2533	−2.49	−0.2142	−2.69
Term Beta	−0.003	−2.21	−0.0016	−0.97	−0.003	−2.07	0.1704	1.39	0.0599	0.78	0.0428	0.54
Liquidity Beta	0.0002	0.39	0.0001	0.11	−0.0001	−0.36	−0.0497	−0.28	−0.1413	−0.9	−0.134	−0.94
bond RMSE	−0.0682	−1.83	−0.112	−2.8	−0.0065	−0.21	−0.4221	−2.4	−0.3508	−2.59	−0.2867	−2.3
BBeta SMrkt	−0.0013	−1.12	−0.0041	−2.59	0.0002	0.15	−0.3448	−2.96	−0.2175	−2.65	−0.1874	−2.33
BBeta RMW	0.0008	1.58	−0.0018	−1.82	0.0005	0.91	−0.0885	−0.88	−0.1112	−1.21	−0.0612	−0.86
BBeta CMA	0.0004	0.74	−0.0013	−1.03	0.0002	0.56	−0.1711	−1.28	−0.1128	−1.4	−0.1042	−1.37
BBeta HML	−0.0001	−0.14	0.0012	0.71	0.0007	0.84	−0.0422	−0.36	−0.0295	−0.4	−0.0157	−0.19
BBeta SMB	−0.0001	−0.08	−0.0007	−0.35	−0.0003	−0.35	0.0547	0.37	−0.0528	−0.44	−0.0399	−0.38
bond idio kurt	−0.0679	−1.7	−0.0854	−1.9	0.0184	0.5	−0.474	−2.19	−0.3893	−2.4	−0.3313	−2.45
bond idio skew	−0.0727	−1.89	−0.0884	−2	0.0078	0.23	−0.478	−2.26	−0.3953	−2.44	−0.3379	−2.48

Notes: Columns 2 to 7 report the average slope coefficient (λ_s) and corresponding t-statistic (t_{λ_s}) for 3 different univariate FM CS regressions (equation 2.4.2) of the risk-adjusted corporate bond excess returns of each of the one month lagged signals listed in Table A.1.4. Columns 2 and 3 are risk adjusted using the Bai et al. [2019b] (BBW) four factor corporate bond return model, columns 4 and 5 are risk adjusted using the BBW and the Fama and French [2015] (FF5) five factor stock model, and columns 6 and 7 are risk adjusted using the BBW and value weighted least squares weighting bonds using one month lagged market capitalization. The risk-adjusted corporate bond excess returns are adjusted as per Brennan et al. [1998]. The control variables are: coupon amount, credit rating, log bond age, and log time to maturity. FM regressions are run each month for the period from July 1, 2005 to October 31, 2019. Newey and West [1987] t-statistics are reported to determine the statistical significance of the average slope coefficients. All right-hand-side variables in equation 2.4.2 are lagged by one month to the left-hand-side corporate bond excess return (or risk adjusted corporate bond excess return) variable. Columns 8 to 13 are 3 sets of univariate portfolio sorts (high minus low portfolio alpha (α_s) and t-statistic (t_{α_s})) of corporate bond returns where portfolios are formed by each of the signals. Portfolio returns are value weighted using amount outstanding as weights. Columns 8 and 9 report the decile high minus low portfolio alpha (α_s) and corresponding t-statistic (t_{α_s}) using the BBW corporate bond factor model (Decile PS (BBW)). Columns 10 and 11 report the quintile high minus low portfolio α_s and t_{α_s} using the BBW (Quintile PS (BBW)). Columns 12 and 13 report the quintile high minus low portfolio alpha α_s and t_{α_s} using the BBW and FF5 (Quintile PS (BBW+FF5)). Alphas are reported in monthly percentages. Newey and West [1987] t-statistics are reported to determine the statistical significance of the high minus low portfolio alpha coefficient. All portfolios are formed by one month lagged signals (listed in Table A.1.4) relative to the corporate bond excess returns and the factors used in the estimation of equation 2.4.1 in each of the specifications Quintile PS (BBW), Decile PS (BBW), and Quintile PS (BBW+FF5). Numbers are up to four decimal places.

Table A.1.8: : Distributional Characteristics of FM lambdas (λ_s) and PS alphas (α_s) and t-statistics

Panel A: FM lambdas (λ_s) and PS alphas (α_s)

					Percentiles							
	N	Mean	Median	Std. Dev.	*Min.*	1*st*	10*th*	25*th*	75*th*	90*th*	99*th*	*Max.*
FM (BBW)	143	−0.006	0	0.0567	−0.5278	−0.2196	−0.0113	−0.0009	0.0008	0.0029	0.0711	0.2847
FM (BBW+FF5)	143	−0.0055	0	0.0517	−0.4152	−0.207	−0.0131	−0.0013	0.0014	0.0054	0.12	0.263
VW FM (BBW)	143	−0.0114	0	0.1168	−1.3327	−0.3085	−0.0167	−0.0006	0.0009	0.0051	0.0847	0.2081
Decile PS (BBW)	143	−0.0247	−0.0252	0.2886	−0.8423	−0.5911	−0.4221	−0.2102	0.1366	0.3335	0.8306	0.8425
Quin. PS (BBW)	143	−0.0235	−0.0214	0.2017	−0.5383	−0.4687	−0.2533	−0.1544	0.0959	0.2453	0.4797	0.5761
Quin. PS (BBW + FF5)	143	−0.0194	−0.0133	0.1773	−0.5493	−0.423	−0.246	−0.1307	0.0781	0.2003	0.3658	0.5442

Panel B: FM lambda t-statistics (t_{λ_s}) and PS alphas t-statistics (t_{α_s})

					Percentiles							
	N	Mean	Median	Std. Dev.	*Min.*	1*st*	10*th*	25*th*	75*th*	90*th*	99*th*	*Max.*
FM (BBW)	143	−0.1426	−0.0768	1.8777	−12.7738	−3.4092	−2.0027	−1.2418	0.9758	1.8068	4.7897	4.8034
FM (BBW+FF5)	143	−0.0551	−0.0433	1.8148	−9.8526	−3.488	−2.103	−1.1492	1.0437	2.169	4.2607	4.8409
VW FM (BBW)	143	−0.0999	0.0772	2.0957	−16.5705	−4.6013	−1.8421	−0.7654	0.8354	1.5155	3.9755	7.5111
Decile PS (BBW)	143	−0.0998	−0.2358	1.7729	−3.8835	−3.2427	−2.3746	−1.4566	1.1822	2.4217	3.3508	3.685
Quin. PS (BBW)	143	−0.1926	−0.2695	1.7554	−4.8544	−4.5693	−2.4408	−1.4018	1.1619	2.2061	3.378	3.5829
Quin. PS (BBW + FF5)	143	−0.129	−0.1801	1.6622	−4.5804	−4.5161	−2.2877	−1.3669	1.0792	2.031	3.7491	3.9554

Notes: The 3 different FM lambdas and 3 different portfolio sorted high-minus low decile alphas are computed on a monthly basis. The number of observations (N), mean, median, standard deviation (Std. Dev.), and corresponding percentiles are measured using all 143 signals. Distributional summary statistics (mean, median, std. dev. and percentiles) are reported in percentages.

Table A.1.9: : Benchmark t-statistics, percentage of H_0 and SnM rejected

	Benchmark t-statistics			Percentage of H_0 rejected				Percentage of SnM rejected		
	Bonf.	Holm	BHY	SHT	Bonf.	Holm	BHY	Bonf.	Holm	BHY
FM (BBW)	3.41	3.41	3.41	21.68	2.8	2.8	2.8	87.1	87.1	87.1
FM (BBW+FF5)	3.49	3.49	3.49	24.48	2.8	2.8	2.8	88.57	88.57	88.57
VW FM (BBW)	3.55	3.55	3.55	17.48	4.2	4.2	4.2	76	76	76
Decile PS (BBW)	3.35	3.35	3.92	37.06	1.4	1.4	0	96.23	96.23	100
Quin. PS (BBW)	3.58	3.58	3.58	33.57	1.4	1.4	1.4	95.83	95.83	95.83
Quin. PS (BBW + FF5)	3.59	3.59	3.59	26.57	2.8	2.8	2.8	89.47	89.47	89.47

Notes: Benchmark t-statistics, percentage of H_0 and SnM rejected from FM regressions and PS in Table A.1.7. Results are presented under single hypothesis test with level of level of significance of 5% and under of the Bonferroni, Holm, and BHY multiple hypothesis tests.

Table A.1.10: : Multivariate FM Regressions of Corporate Bond Returns

Variable Name	FM		VW FM		FM (BBW)		VW FM (BBW)		FM (BBW+FF5)		VW FM (BBW+FF5)	
	λ_s	t_{λ_s}	λ_s	t_{λ_s}	λ_s	t_{λ_s}	λ_s	t_{λ_s}	λ_s	t_{λ_s}	λ_s	t_{λ_s}
Intercept	0.2229	1.88	0.085	0.59	0.2689	1.99	0.2402	1.51	0.1361	1.15	0.0124	0.07
spi	−0.0022	−0.49	0.004	0.8	−0.0018	−0.32	0.0001	0.02	−0.0021	−0.3	0.0009	0.17
mve f	0	0.88	0	−0.28	0	−0.34	0	0.28	0	1.15	0	1.28
bm	−0.0005	−0.73	−0.0001	−0.19	0	−0.03	0.0007	0.92	0	0.06	0.0004	0.44
ep	0.0009	0.52	−0.0004	−0.18	0.0011	0.4	−0.0003	−0.12	0.0006	0.25	−0.0005	−0.17
cashpr	0	−1.44	0	−2.17	0	−1.22	0	−1.79	0	−2.05	0	−1.81
lev	0	−0.16	0	0.08	−0.0001	−1.02	−0.0001	−0.34	0	−0.12	0	−0.1
sp	−0.0001	−0.43	−0.0001	−0.25	−0.0001	−0.32	−0.0001	−0.49	−0.0002	−1.24	0	−0.11
roic	−0.0004	−0.21	0.0014	0.6	−0.0024	−0.81	−0.0001	−0.04	−0.0022	−0.75	0.0044	1.26
agr	−0.0002	−0.23	−0.0003	−0.22	−0.001	−0.47	0.0001	0.03	0.0008	0.33	0.002	0.79
gma	−0.0017	−1.73	−0.001	−0.99	−0.0014	−1.11	−0.0004	−0.29	−0.0019	−1.77	−0.0019	−1.43
chcsho	−0.0004	−0.94	−0.0001	−0.27	−0.0003	−0.62	0.0005	1.02	−0.0009	−1.49	−0.0001	−0.19
lgr	0.0003	0.67	0.0003	0.48	0.0007	0.69	−0.0006	−0.58	−0.0006	−0.51	−0.002	−1.45
acc	−0.0003	−0.06	−0.0037	−0.49	0.0093	0.98	0.0026	0.27	0.0056	0.65	0.0001	0.01
pctacc	0	−0.46	0	0.23	0	−0.6	0	−0.64	0	0.12	0	0.89
cfp	−0.0019	−1.07	−0.0021	−1.63	−0.0028	−1.65	−0.0031	−2	−0.001	−0.63	−0.002	−1.39
absacc	0.0087	1.62	0.0019	0.48	0.0103	1.67	0.0018	0.4	0.0078	1.27	0.0001	0.03
chinv	−0.0068	−1.1	−0.0008	−0.14	−0.0072	−0.92	0.0012	0.17	−0.015	−1.62	−0.008	−1.07
cf	−0.0014	−0.34	−0.0026	−0.31	0.012	1.36	0.0065	0.7	0.0082	1.09	0.0064	0.59
hire	−0.0006	−0.68	−0.0008	−0.98	0.0004	0.33	−0.0003	−0.26	0.001	0.89	−0.0002	−0.16
sgr	0.001	0.63	0.0004	0.47	0.0012	0.64	0.0015	1.28	0.0008	0.5	0.0016	1.17
chpm	−0.0002	−0.84	0.0003	1.3	−0.0005	−0.69	0.0001	0.09	0.0005	1.12	0.0013	1.71
chato	−0.0014	−0.86	−0.0007	−0.66	0.0005	0.26	−0.0002	−0.11	0.0006	0.27	0.0003	0.12
pchsale pchinvt	0.0001	0.5	−0.0001	−0.41	−0.0002	−0.77	−0.0001	−0.45	0.0001	0.51	−0.0002	−0.51
pchsale pchrect	−0.0003	−0.73	−0.0003	−0.67	−0.0006	−1	−0.0009	−1.19	−0.0003	−0.6	−0.0005	−0.69
pchgm pchsale	−0.0003	−0.58	0.0001	0.14	0.0001	0.13	0.0003	0.52	−0.0006	−0.81	−0.0006	−0.91
pchsale pchxsga	−0.0007	−0.78	0	0.05	−0.0015	−1.41	−0.0018	−1.48	0.0002	0.19	0.0011	0.82
depr	0.0008	1.29	0.0008	1.77	0.0013	1.56	0.0006	1.08	0.0009	1.15	0.0003	0.4
pchdepr	0	−0.04	0.0001	0.12	−0.0008	−0.82	−0.0006	−0.96	−0.0005	−0.61	−0.0007	−1
chadv	0.0005	0.71	0.0005	0.82	0.0008	1.24	0.0003	0.4	0.0001	0.15	0.0003	0.3
invest	0.0003	0.14	−0.0001	−0.07	0.0008	0.4	0.0004	0.19	0.0031	1.31	0.0035	1.63
egr	−0.0003	−0.6	−0.0004	−0.84	0.0001	0.12	−0.0002	−0.26	−0.0004	−0.54	−0.0005	−0.67
pchcapx	0.0001	0.45	−0.0001	−0.44	0.0003	1.31	−0.0001	−0.17	0.0003	1.39	0.0002	0.36

grcapx	−0.0001	−0.59	−0.0001	−0.5	−0.0001	−0.37	−0.0001	−0.51	0	0.26	0.0001	0.29
tang	0.0007	0.54	0.0015	1.31	0.0006	0.42	0.0008	0.66	−0.0001	−0.1	0	0.02
currat	0	0.19	0.0001	0.93	−0.0001	−0.39	0.0001	0.72	0.0001	0.96	0.0003	1.4
pchcurrat	−0.0012	−1.19	−0.0001	−0.15	−0.0006	−0.45	0.0001	0.09	−0.0013	−1.2	0	0.01
quick	0	−0.4	−0.0001	−1.03	0	0.15	−0.0002	−0.74	−0.0002	−1.5	−0.0004	−1.74
pchquick	0.0007	0.89	0.0003	0.33	0.0002	0.15	−0.0003	−0.27	0.0005	0.51	−0.0001	−0.1
salecash	0	−1.04	0	0.68	0	−0.13	0	0.71	0	−0.36	0	0.6
salerec	0	−0.36	0	0.04	0	−0.26	0	0.58	0	0.63	0	0.01
saleinv	0	−0.91	0	−0.52	0	−1.08	0	−0.64	0	−1.06	0	−0.67
pchsaleinv	−0.0001	−0.32	0	−0.06	0.0003	0.84	0.0001	0.14	−0.0002	−0.43	−0.0002	−0.36
cashdebt	0.0016	1.04	0.001	0.41	−0.0002	−0.07	0.0016	0.53	0.0014	0.65	−0.0004	−0.13
realestate	0.0015	1.4	0.0007	1.01	0.0022	1.61	0.0015	1.24	0.0024	1.71	0.0032	2.16
grltnoa	−0.0006	−0.43	−0.0004	−0.26	−0.0032	−1.56	−0.0021	−1.01	−0.0038	−1.76	−0.0039	−1.52
rdbias	0	−0.17	−0.0002	−1.75	−0.0002	−0.79	−0.0002	−0.91	−0.0001	−0.65	−0.0001	−0.78
roe	0.0002	0.43	0.0002	0.33	−0.0001	−0.15	−0.0002	−0.29	0	0.07	0.0001	0.09
operprof	0	−0.08	0.0001	0.59	0	−0.26	0.0001	0.36	0.0001	0.47	0.0001	0.41
chpmia	0	−0.01	0	−0.19	0	−0.44	0	−1.01	0	0.47	0	−0.05
chatoia	−0.0002	−0.22	0.0005	0.42	−0.002	−1.36	−0.0004	−0.28	−0.0014	−0.69	−0.0007	−0.28
chempia	−0.0002	−0.49	0.0004	1.1	−0.0003	−0.62	0.0003	0.66	−0.0004	−0.72	0.0007	1.27
bm ia	0	−1.13	0	−1.51	0	−1.06	0	−1.39	0	−0.57	0	−0.74
pchcapx ia	0	1.62	0	1.42	0	0.7	0	−0.08	0	0.17	0	1.05
tb	0.0001	0.81	0.0001	1.61	0.0002	0.94	0.0002	1.37	0.0002	0.99	0.0002	1.18
cfp ia	0	0.52	0	1.08	0.0001	1.04	0	0.16	0.0001	0.95	0.0001	0.97
herf	0.0017	1.2	0.0001	0.11	0.0005	0.27	−0.0013	−1.14	0.0001	0.02	−0.0006	−0.26
orgcap	0.0371	0.92	0.0064	0.19	0.0395	0.86	0.0014	0.04	0.0082	0.17	−0.0092	−0.2
mve m	0	−0.86	0	0.83	0	0.4	0	−0.19	0	−0.99	0	−0.79
pps	−0.0003	−1.37	−0.0002	−1.28	−0.0003	−1.16	0	−0.06	−0.0004	−1.52	−0.0005	−1.2
rdq	0	−1.79	0	−0.55	0	−1.95	0	−1.51	0	−1.16	0	−0.09
prccq	0	0.19	0	0.12	0	0.52	0	0.25	0	1.4	0	0.66
chtx	0.0099	0.65	−0.021	−1.82	0.0158	0.87	−0.0218	−1.52	0.0334	1.46	−0.009	−0.48
roaq	0.0063	0.69	0.0058	0.62	0.0054	0.58	0.004	0.31	0.0008	0.06	0.0183	1.03
roeq	−0.0007	−0.5	−0.0006	−0.4	−0.0018	−0.89	0	0.02	0.0001	0.05	−0.0011	−0.74
rsup	0.0011	0.49	0.0013	0.61	0.0009	0.28	0.0021	0.6	0.0025	0.93	0.0019	0.59
stdacc	0.0002	1.07	−0.0001	−0.3	0.0002	0.41	0.0006	0.87	0.0001	0.12	0.0003	0.46
sgrvol	0.0014	1.24	0.0006	0.54	−0.001	−0.68	−0.0017	−1.04	−0.0001	−0.06	0.0007	0.39
roavol	−0.0043	−0.57	−0.0041	−0.51	−0.0158	−1.64	−0.0141	−1.54	−0.0133	−1.23	−0.0116	−1.05
stdcf	−0.0001	−1.12	0	0.43	−0.0001	−0.34	−0.0002	−0.67	0	0.08	0	−0.16
cash	−0.0004	−0.19	−0.0018	−0.98	−0.0013	−0.42	−0.003	−1.11	−0.0004	−0.1	−0.0035	−1.25

cinvest	−0.0001	−0.34	−0.0002	−0.79	−0.0004	−0.97	−0.0005	−1.37	−0.0002	−0.4	−0.0001	−0.19
sue	−0.0009	−0.22	−0.0088	−1.37	0.0005	0.08	−0.0142	−1.45	−0.0037	−0.7	−0.0102	−1.19
aeavol	0	0.03	−0.0001	−1.13	0.0002	1.2	−0.0001	−0.58	0.0004	2.17	0.0001	1.03
ear	−0.0021	−1.25	−0.0009	−0.54	−0.0008	−0.38	−0.0013	−0.67	−0.002	−1.06	−0.0026	−1.3
disp	−0.0004	−0.55	−0.001	−1.24	−0.001	−1.37	−0.0019	−1.81	−0.0003	−0.4	−0.002	−1.66
chfeps	0.0008	1.7	0.0004	0.99	0.0003	0.72	−0.0002	−0.24	0.0007	1.48	−0.0002	−0.22
fgr5yr	0	−1.78	0	−0.36	0	−1.26	0	0.05	0	−1.34	0.0001	1.85
MEANREC	−0.0002	−0.47	0.0002	0.59	0.0005	1.02	0.0002	0.47	0.0004	0.71	0.0001	0.12
chrec	−0.0003	−0.94	0	0.03	0	−0.07	0	−0.07	0.0002	0.31	−0.0004	−0.75
nanalyst	0	0.74	0	0.13	0	0.74	0	1.2	0	0.48	0	1.2
sfe	0.0005	1.27	0.0013	2.23	0.0003	0.57	0.001	1.43	0.0008	1.14	0.0006	0.71
MEANEST	0	0.04	0	−1.61	0	−0.49	0	−1.94	0	0.11	0	−1.4
mom6m	0.0076	1.29	0.0042	0.71	0.0175	2.05	0.0148	1.56	0.0138	1.94	0.0051	0.7
mom12m	−0.0034	−1.05	−0.0012	−0.38	−0.0086	−1.94	−0.0076	−1.57	−0.0065	−1.52	−0.003	−0.78
mom36m	0	0.03	−0.0003	−0.93	−0.0007	−1.51	−0.001	−1.73	0.0001	0.32	−0.0006	−1.07
mom1m	0.0189	4.95	0.0096	2.67	0.0177	3.55	0.0126	2.42	0.0174	4.7	0.01	2.2
dolvol	0.0002	1.12	0.0001	0.38	0	0.19	0.0002	0.73	0	0.16	−0.0001	−0.22
chmom	−0.0028	−0.98	−0.0007	−0.26	−0.0074	−1.8	−0.0056	−1.25	−0.0056	−1.79	−0.0006	−0.18
turn	−0.0007	−3.64	−0.0005	−2.14	−0.0006	−3.3	−0.0005	−2.32	−0.0006	−3.57	−0.0005	−1.98
indmom	0.0002	0.14	0.0001	0.08	0.0016	1.21	0.0019	1.38	−0.0001	−0.04	0.0015	0.82
maxret	−0.0081	−1	−0.0011	−0.11	−0.0055	−0.57	0.0003	0.03	−0.0068	−0.61	−0.0043	−0.35
retvol	−0.068	−2.16	−0.0489	−1.25	−0.0067	−0.18	−0.052	−1.1	−0.0395	−0.98	−0.0547	−0.96
baspread	0.0643	2.25	0.0373	1.12	0.011	0.35	0.0072	0.17	0.0292	0.84	0.0218	0.48
std dolvol	0.0009	0.85	−0.0008	−0.37	−0.0004	−0.24	−0.0012	−0.57	0.0007	0.34	0.0007	0.39
std turn	0.0001	1.64	0.0001	1.13	0.0001	1.17	0.0001	1.31	0.0001	1.35	0.0001	1.41
beta	−0.0012	−0.52	−0.0021	−1.15	−0.0019	−0.82	−0.0023	−1.28	−0.0034	−1.41	−0.0013	−0.75
betasq	0.0003	0.49	0.0005	0.81	0.0006	0.98	0.0002	0.35	0.001	1.28	0.0001	0.15
rsq1	0.0011	0.42	0.0009	0.34	0.0019	0.63	0.0034	1.09	0.0027	0.84	0.0001	0.04
pricedelay	0.0004	0.43	0.0004	0.63	−0.0007	−0.67	−0.0007	−0.81	0.0003	0.28	−0.0005	−0.42
idiovol	0.018	0.66	0.0522	1.49	0.027	0.79	0.0641	1.68	0.0225	0.64	0.0132	0.36
d WCA	0	0.56	0	0.76	0	0.68	0	−0.61	0	1.83	0	1.02
d NCO	0	1.85	0	0.68	0	1.73	0	−0.08	0	1.66	0	−0.16
d FIN	0	0.98	0	0.65	0	1.14	0	0.62	0	−0.54	0	−0.94
EDF	−0.002	−1	−0.0046	−1.9	−0.0025	−1.13	−0.0063	−2.09	−0.0017	−0.64	−0.0071	−2.04
avg med h Q	0.0007	0.07	0.007	0.85	−0.0132	−1.2	−0.0024	−0.22	−0.0034	−0.27	0.0061	0.48
cw	−0.0094	−1.28	−0.0024	−0.29	−0.0086	−0.95	−0.007	−0.6	−0.0108	−0.96	−0.0032	−0.24
os	−0.0024	−1.77	0.0001	0.1	−0.0007	−0.57	0	0	−0.0001	−0.07	0.0015	0.85
d ivol	−0.0494	−1.88	−0.0283	−1.19	−1.3441	−1.5	−1.3465	−1.19	−0.0636	−0.19	−1.0058	−1.21

cvol	0.0472	1.77	0.0192	0.8	1.3408	1.5	1.3348	1.18	0.0632	0.19	0.9991	1.2
pvol	0.0448	1.72	0.0254	1.08	1.3439	1.5	1.349	1.19	0.0602	0.18	1.0042	1.21
skewness	−0.0067	−2.11	0.0014	0.39	−0.0057	−1.49	−0.0008	−0.17	−0.005	−1.37	−0.0002	−0.03
avg ivol	−0.0022	−1.54	−0.0012	−1.14	−0.0029	−1.73	−0.0006	−0.47	−0.001	−0.53	−0.001	−0.69
bond st rev	−0.0104	−1.53	−0.0102	−1.35	−0.015	−1.37	−0.0218	−2	−0.033	−2.06	−0.0406	−2.64
bond Amihud	0	0.09	0	0.19	0	0	0	0	0	0	0	0
Std. log Price	0.0417	2.68	0.0557	3.24	0.0587	3.85	0.0751	4.44	0.04	2.7	0.0684	3.66
Mac Dur	−0.0005	−3.21	−0.0004	−3.61	−0.0005	−2.85	−0.0005	−3.48	−0.0005	−2.81	−0.0004	−2.96
Bond Kurtosis	0	0.87	0.0001	1.8	0	0.2	0.0001	1.32	0.0001	1.32	0.0001	1.91
Bond Skewness	−0.0008	−7.09	−0.0006	−5.59	−0.0008	−4.53	−0.0006	−4.38	−0.0009	−4.81	−0.0006	−3.51
Default Beta	0	−0.21	0.0003	1.39	−0.0034	−3.43	−0.0028	−2.99	−0.0029	−2.83	−0.0025	−2.22
Term Beta	0.0007	1.33	0.0002	0.36	−0.0006	−0.45	−0.001	−0.56	0.0003	0.19	−0.0006	−0.25
Liquidity Beta	0.0001	0.63	0.0001	1.07	−0.0003	−1.23	−0.0003	−1.06	−0.0006	−1.87	−0.0009	−1.96
bond RMSE	0.0323	1.66	−0.0091	−0.38	−0.0296	−1.6	−0.0501	−2.3	−0.0945	−3.18	−0.0689	−2.07
BBeta SMrkt	−0.0001	−0.13	0.0005	1.16	−0.0039	−3.72	−0.0027	−2.45	−0.0059	−2.99	−0.005	−2.09
BBeta RMW	0	0.13	−0.0002	−0.56	0.0002	0.46	−0.0002	−0.54	−0.0015	−1.36	−0.0015	−1.25
BBeta CMA	−0.0003	−1.52	−0.0001	−0.23	0.0001	0.28	0.0001	0.42	−0.0007	−0.69	−0.0006	−0.57
BBeta HML	0.0003	0.83	0.0004	1.15	0.002	2.39	0.0021	2.67	0.0018	1	0.002	1.1
BBeta SMB	0.0005	1.07	0.0004	1.03	−0.0002	−0.24	−0.0001	−0.12	−0.002	−1.4	−0.0026	−1.66
bond idio kurt	−0.0381	−0.88	0.0146	0.53	−0.0228	−0.39	0.1074	3.12	−0.0571	−0.89	0.0575	1.05
bond idio skew	0.0325	0.74	0.0086	0.3	0.0571	0.98	−0.0478	−1.59	0.1168	1.72	−0.0048	−0.08
Bond Credit Rating	0.0002	2.48	0.0001	2.36	0.0001	1.92	0.0001	2.5	0.0001	1.54	0	−0.17
Log(offering amount)	−0.0006	−3.78	−0.0004	−3.18	−0.0002	−1.34	−0.0001	−0.57	−0.0004	−2.13	−0.0001	−0.79
log(age)	0.0015	3.7	0.0011	4.22	0.0017	3.39	0.0014	3.67	0.002	3.33	0.0017	3.53
log(TTM)	0.0032	3.58	0.0031	3.83	0.0026	3.32	0.0027	3.82	0.0024	3.25	0.0024	3.46
coupon	0.0002	2.55	0.0003	2.5	0.0002	1.89	0.0002	2.22	0.0002	2.22	0.0002	1.69
R-squared	0.2837		0.2311		0.2564		0.2102		0.3238		0.3026	
Avg. Num. CS Obs.	4067.25		4064.85		2826.72		2825.02		2761.11		2759.48	

Notes: Columns 2 to 13 report the average slope coefficient (λ_s) and corresponding t-statistic (t_{λ_s}) for 6 different multivariate FM CS regressions of one-month-ahead corporate bond excess returns of 135 of the 143 signals. Columns 2 and 3 are raw excess corporate bond returns (columns 4 and 5 represent the corresponding value-weighted (VW), by one month lagged firm market capitalization, least squares estimated MV FM regression), columns 6 and 7 are risk adjusted using the Bai et al. [2019b] (BBW) four factor corporate bond return model (columns 8 and 9 represent the corresponding VW, by one month lagged firm market capitalization, least squares estimated MV FM regression), columns 10 and 11 are risk adjusted using the BBW and the Fama and French [2015] (FF5) five factor stock model (columns 12 and 13 represent the corresponding VW, by one month lagged firm market capitalization, least squares estimated MV FM regression). The risk-adjusted corporate bond excess returns are calculated as per Brennan et al. [1998]. All right-hand-side variables are lagged by one month to the left-hand-side return variable. Newey and West [1987] t-statistics are reported to determine the statistical significance of the average slope coefficients. Numbers are reported up to four decimal places for ease of display.

Table A.1.11: : Benchmark t-statistics, percentage of H_0 and SnM rejected: Multivariate FM Corporate Bond Return Regression

	Benchmark t-statistics			Percentage of H_0 rejected				Percentage of SnM rejected		
	Bonf.	Holm	BHY	SHT	Bonf.	Holm	BHY	Bonf.	Holm	BHY
FM	3.59	3.59	3.67	10.85	3.88	3.88	2.33	64.29	64.29	78.57
VW FM	3.61	3.61	3.61	8.53	2.33	2.33	2.33	72.73	72.73	72.73
FM (BBW)	3.6	3.6	3.86	9.3	2.33	2.33	0.78	75	75	91.67
VW FM (BBW)	3.48	3.48	3.66	12.4	3.1	3.1	2.33	75	75	81.25
FM (BBW+FF5)	3.48	3.48	3.48	11.63	1.55	1.55	1.55	86.67	86.67	86.67
VW FM (BBW+FF5)	3.51	3.51	4.09	9.3	0.78	0.78	0	91.67	91.67	100

Notes: Benchmark t-statistics, percentage of H_0 and SnM rejected from MV FM regressions in Table A.1.10. Results are presented under single hypothesis test with level of level of significance of 5% and under of the Bonferroni, Holm, and BHY multiple hypothesis tests.

Table A.1.12: : Multivariate FM Regressions of Stock Returns

	FM		VW FM		FM (Stock Price>$5)		VW FM (Stock Price>$5)	
Variable Name	λ_s	t_{λ_s}	λ_s	t_{λ_s}	λ_s	t_{λ_s}	λ_s	t_{λ_s}
Intercept	0.3162	0.77	0.859	2.23	0.3762	1.01	0.7965	2.16
spi	0.009	0.41	−0.0029	−0.16	−0.0169	−0.89	−0.0022	−0.12
mve f	0	1.74	0	1.39	0	1.38	0	1.16
bm	−0.0004	−0.2	−0.0004	−0.21	−0.0035	−2.07	−0.0008	−0.42
ep	0.0021	0.3	−0.0032	−0.42	−0.001	−0.15	−0.0009	−0.11
cashpr	0	−0.05	0	−0.62	0	0.03	0	−0.41
lev	−0.0004	−1.33	−0.0005	−1.57	−0.0001	−0.4	−0.0003	−1.12
sp	0.0001	0.19	0.0005	1.12	0.0009	1.89	0.0006	1.28
roic	−0.0075	−1.15	0.0023	0.37	−0.0075	−1.21	0.0015	0.21
agr	0.0015	0.31	−0.0055	−1.15	−0.0034	−0.79	−0.0068	−1.39
gma	−0.0029	−0.71	−0.0027	−0.72	−0.0049	−1.28	−0.0018	−0.48
chcsho	0.0007	0.33	−0.0033	−2.42	−0.0006	−0.3	−0.0024	−1.59
lgr	0	0.02	0.0022	0.97	0.0019	0.85	0.0026	1.12
acc	0.0099	0.52	−0.0093	−0.43	0.0094	0.54	−0.0112	−0.51
pctacc	0	0.12	−0.0001	−0.53	0	0.18	0	−0.01
cfp	0.0066	1.03	−0.0007	−0.14	0.0078	1.41	−0.0013	−0.23
absacc	0.0133	0.99	0.0116	0.94	0.0055	0.49	0.0099	0.79
chinv	0.0249	1.11	−0.0008	−0.04	0.0292	1.46	0.0031	0.14
cf	0.01	0.58	−0.0246	−1.42	0.0059	0.33	−0.0238	−1.25
hire	0.002	0.66	−0.0005	−0.2	0.0006	0.23	−0.0009	−0.32
sgr	−0.0037	−1.25	−0.0015	−0.4	0.0006	0.18	−0.0017	−0.44
chpm	0.0008	0.7	0.0011	1.36	0.0021	1.84	0.002	1.7
chato	−0.0035	−0.5	−0.0014	−0.27	−0.0117	−1.66	−0.0037	−0.71
pchsale pchinvt	−0.0002	−0.21	0.0001	0.06	−0.0002	−0.19	−0.0001	−0.05
pchsale pchrect	0.0003	0.19	0.0016	1.55	−0.0004	−0.29	0.0014	1.26
pchgm pchsale	0.0013	0.84	0.0014	1	0.0009	0.69	0.0013	0.88
pchsale pchxsga	−0.0018	−0.57	−0.0019	−0.65	−0.003	−0.93	−0.0013	−0.39
depr	−0.0011	−0.69	0.0017	1.05	0.0002	0.13	0.0019	1.07
pchdepr	0.0044	2.33	−0.0018	−0.92	0.0015	0.72	−0.0014	−0.65
chadv	0.003	1.16	−0.0008	−0.35	0.0022	1.05	−0.0006	−0.28
invest	−0.0051	−0.69	0.0033	0.54	−0.0066	−1.02	0.0036	0.61
egr	−0.0025	−2.15	−0.0023	−2.12	−0.002	−1.91	−0.002	−2.12
pchcapx	0.0004	0.79	0.001	1.69	0.001	2.13	0.001	1.62

grcapx	−0.0001	−0.29	0	0.12	0.0002	0.38	0.0001	0.2
tang	−0.0024	−0.58	−0.0008	−0.22	−0.0013	−0.36	−0.0001	−0.03
currat	0	−0.17	−0.0003	−0.89	−0.0002	−0.89	−0.0004	−1.31
pchcurrat	−0.0011	−0.33	−0.0025	−0.73	−0.0009	−0.29	−0.0034	−0.98
quick	0	−0.09	0.0003	0.84	0.0002	0.74	0.0004	1.3
pchquick	0.0016	0.52	0.003	1	0.0021	0.74	0.0038	1.25
salecash	0	−0.17	0	0.17	0	0.6	0	0.35
salerec	0	0.5	0	1.24	0	0.46	0	1.08
saleinv	0	−0.76	0	−0.78	0	0.15	0	−0.86
pchsaleinv	−0.0002	−0.13	−0.0005	−0.41	0.0008	0.67	−0.0004	−0.35
cashdebt	−0.0013	−0.24	0.0053	1.25	0.0087	1.76	0.0046	1.01
realestate	−0.0021	−0.55	0.003	0.83	0.0007	0.18	0.002	0.57
grltnoa	−0.0044	−0.79	0.0034	0.59	−0.0014	−0.26	0.0029	0.5
rdbias	0.0009	1.8	0	−0.05	0.0005	1.03	0	0.08
roe	0.0019	1.53	0.0024	1.6	0.0023	1.5	0.0028	1.79
operprof	−0.0004	−0.89	−0.0004	−0.89	−0.0008	−1.59	−0.0007	−1.51
chpmia	0	0.06	0	0.39	0	0.81	0	0.55
chatoia	−0.0011	−0.21	−0.0032	−0.68	0.0027	0.56	−0.0014	−0.29
chempia	−0.0014	−0.86	0.0021	1.46	−0.0001	−0.09	0.0024	1.7
bm ia	0	−0.04	0	0.03	0	−0.46	0	0.07
pchcapx ia	0.0001	1.05	0	0.28	0.0001	1.31	0	0.24
tb	−0.0001	−0.16	0	0.12	0	−0.02	0	0.07
cfp ia	0.0002	0.54	0.0001	0.44	0.0003	1.08	0.0001	0.35
herf	−0.0067	−1	−0.0028	−0.46	−0.0091	−1.5	−0.0017	−0.27
orgcap	0.243	1.44	0.046	0.37	0.2545	1.76	0.068	0.55
mve m	0	−1.58	0	−1.5	0	−1.4	0	−1.28
pps	−0.0005	−0.6	−0.0001	−0.16	−0.0008	−0.89	0	−0.05
rdq	0	−0.8	−0.0001	−2.25	0	−1.04	0	−2.2
prccq	0	1.55	0	0.52	0	1.26	0	0.38
chtx	0.0863	1.84	0.0596	1.19	0.0339	0.74	0.0607	1.14
roaq	−0.0032	−0.11	0.0287	0.87	−0.0093	−0.32	0.034	1.05
roeq	0.0032	0.96	0.0031	0.92	0.004	1.57	0.0008	0.28
rsup	0.01	1.42	0.0071	0.91	0.0096	1.45	0.0075	0.87
stdacc	−0.0003	−0.27	0.0005	0.65	0.0021	0.95	0.0032	1.7
sgrvol	−0.0043	−1	−0.0072	−1.84	−0.0082	−1.51	−0.01	−1.95
roavol	0.0028	0.1	0.0137	0.42	−0.0035	−0.12	0.0124	0.38
stdcf	0	−0.06	−0.0002	−0.74	−0.001	−1.21	−0.0012	−1.78
cash	0.0057	1.03	0.0016	0.26	0.0001	0.02	0.0008	0.13

cinvest	−0.0001	−0.15	−0.001	−1.19	−0.0009	−0.99	−0.0013	−1.52
sue	0.021	1.5	−0.0007	−0.05	−0.0113	−0.65	−0.0021	−0.1
aeavol	0.0003	0.53	−0.0001	−0.18	0.0001	0.26	−0.0001	−0.26
ear	0.0156	2.85	0.0083	1.45	0.0164	2.97	0.0071	1.22
disp	−0.0021	−1.34	−0.0014	−0.55	−0.0026	−1.61	−0.0029	−1.06
chfeps	0.0054	2.79	0.0029	2.07	0.0039	2.54	0.0021	1.45
fgr5yr	0.0001	0.93	0	0.38	0	0.24	0	0.52
MEANREC	−0.0005	−0.42	0.0005	0.39	−0.0005	−0.5	0.0008	0.62
chrec	−0.001	−1.03	0.0002	0.21	−0.0008	−0.76	0.0005	0.41
nanalyst	0.0001	1.39	0.0001	1.83	0.0001	1.93	0.0002	2.1
sfe	−0.0003	−0.24	−0.0043	−2.3	−0.0092	−2.4	−0.0117	−3.37
MEANEST	0	−1.48	0	1.52	0	0.68	0	1.44
mom6m	−0.0151	−1.05	0.0181	1.09	0.0213	1.6	0.0331	1.92
mom12m	0.0068	1.1	−0.0076	−1.05	−0.0098	−1.39	−0.0161	−1.99
mom36m	0.0023	1.78	0.0015	1.06	0.0022	1.79	0.0006	0.46
mom1m	0	0	−0.0084	−0.58	0.0095	0.82	−0.0022	−0.15
dolvol	−0.0006	−1.13	−0.0009	−1.55	−0.0007	−1.33	−0.0011	−1.9
chmom	0.0033	0.51	−0.0107	−1.33	−0.0133	−2.03	−0.0177	−2.11
turn	−0.0011	−1.95	0.0001	0.25	−0.001	−1.97	0.0001	0.29
indmom	0.0031	0.56	0.0052	1.15	0.0022	0.44	0.0041	0.88
maxret	0.0111	0.35	0.0552	1.78	0.0109	0.38	0.0553	1.74
retvol	−0.3239	−2.47	−0.3137	−1.86	−0.1696	−1.46	−0.3194	−1.85
baspread	0.3286	2.79	0.0316	0.26	0.1156	1.18	0.0268	0.22
std dolvol	−0.0056	−1.47	0.001	0.25	−0.0045	−1.39	0.0017	0.41
std turn	0.0002	1.45	−0.0001	−1.16	0.0001	1.2	−0.0001	−0.98
beta	0.0017	0.26	−0.0081	−1.32	−0.0058	−0.96	−0.0123	−1.74
betasq	0.0001	0.08	0.0018	1.18	0.0022	1.59	0.0031	1.87
rsq1	−0.0055	−0.59	0.0123	1.35	0	0	0.0163	1.61
pricedelay	−0.0018	−0.55	−0.0008	−0.28	−0.0032	−1.3	−0.0008	−0.28
idiovol	−0.0785	−0.93	0.2126	2.05	0.0349	0.41	0.2611	2.32
d WCA	0	0.62	0	1.74	0	1	0	1.7
d NCO	0	1.05	0	0.22	0	0.09	0	0.15
d FIN	0	2.35	0	1.82	0	2.09	0	1.57
EDF	−0.0092	−1.11	−0.0056	−0.77	0.0025	0.27	−0.0005	−0.06
implied fee	−0.0336	−1.05	−0.1097	−3.13	−0.0278	−1.02	−0.1047	−2.93
cw	−0.0129	−0.49	−0.0356	−1.1	−0.0045	−0.22	−0.0213	−0.66
os	−0.0181	−3	−0.0088	−1.94	−0.0129	−2.51	−0.0081	−1.77
d ivol	−0.0852	−0.8	0.0399	1.41	0.1961	1.29	−0.1884	−0.57

cvol	0.1046	0.98	−0.0487	−1.56	−0.1873	−1.23	0.182	0.54
pvol	0.0664	0.62	−0.0339	−1.16	−0.2025	−1.33	0.1927	0.58
skewness	−0.0288	−2.6	−0.0175	−1.63	−0.0105	−0.82	−0.0142	−1.23
avg ivol	−0.0012	−0.19	−0.0023	−0.51	−0.0048	−1.12	−0.003	−0.68
bond st rev	0.0758	2.24	0.0611	1.75	0.0757	2.65	0.0615	1.79
bond Amihud	−0.0337	−1.11	0.0163	0.68	−0.0254	−1.19	0.0157	0.65
Std. log Price	−0.1015	−1.14	0.0271	0.42	0.0028	0.05	0.012	0.19
Mac Dur	0.0003	1.27	0.0002	1.27	0.0002	0.77	0.0002	1.08
Bond Kurtosis	0.0004	1.43	0.0001	0.37	0.0002	0.74	0.0001	0.28
Bond Skewness	0.001	1.66	0.0009	1.53	0.0004	0.8	0.0008	1.42
Default Beta	0.0018	1.42	−0.0014	−1.22	0.0003	0.24	−0.0017	−1.44
Term Beta	−0.0013	−0.54	0.0007	0.37	0.0011	0.57	0.0013	0.69
Liquidity Beta	−0.0005	−0.54	0.0006	0.89	−0.0002	−0.22	0.0006	0.86
bond RMSE	−0.0474	−0.42	0.0375	0.35	−0.0688	−0.85	0.0661	0.61
BBeta SMrkt	0.002	0.69	−0.0032	−1.25	−0.0005	−0.2	−0.0041	−1.61
BBeta RMW	0.001	0.77	0.0019	1.36	0.0004	0.34	0.0021	1.46
BBeta CMA	−0.0026	−1.6	−0.0023	−1.55	−0.0022	−1.79	−0.0018	−1.28
BBeta HML	−0.0007	−0.28	−0.0026	−1.11	−0.0003	−0.16	−0.0018	−0.75
BBeta SMB	0.0019	0.6	0.0004	0.13	0.0012	0.49	−0.0001	−0.03
bond idio kurt	0.4017	0.83	0.0199	0.04	−0.4489	−1.17	−0.2548	−0.49
bond idio skew	−0.2736	−0.53	−0.0479	−0.1	0.5505	1.46	0.1795	0.35
Credit Rating	0	0.02	−0.0001	−0.46	0	0.24	0	−0.33
Log(off. amount)	0.0003	0.56	0.0008	1.86	0.0005	1.04	0.0008	2
log(age)	0.0001	0.18	0.0005	1.06	0.0004	0.66	0.0007	1.24
log(TTM)	−0.0009	−0.67	−0.0009	−0.74	−0.0012	−0.96	−0.0006	−0.44
coupon	−0.0004	−1.05	−0.0005	−1.55	−0.0001	−0.2	−0.0003	−0.99
R-squared	0.2963		0.4731		0.2746		0.4782	
Avg. Num. CS Obs.	730.5988		730.5926		690.2901		690.2901	

Notes: Columns 2 to 9 report the average slope coefficient (λ_s) and corresponding t-statistic (t_{λ_s}) for 6 different multivariate FM CS regressions of one-month-ahead corporate bond excess returns of 135 of the 143 signals. Columns 2 and 3 are raw excess corporate bond returns (columns 4 and 5 represent the corresponding value-weighted (VW), by one month lagged firm market capitalization, least squares estimated MV FM regression), All right-hand-side variables are lagged by one month to the left-hand-side return variable. FM regressions are run each month for the period from July 1, 2005 to October 31, 2019. Newey and West [1987] t-statistics are reported to determine the statistical significance of the average slope coefficients. Numbers are reported up to four decimal places for ease of display.

Table A.1.13: : Benchmark t-statistics, percentage of H_0 and SnM rejected: Multivariate FM Stock Return Regression

	Benchmark t-statistics			Percentage of H_0 rejected				Percentage of SnM rejected		
	Bonf.	Holm	BHY	SHT	Bonf.	Holm	BHY	Bonf.	Holm	BHY
FM	3.43	3.43	3.9	7.41	0	0	0	100	100	100
VW FM	3.43	3.43	3.9	5.93	0	0	0	100	100	100
FM (Stock Price > \$5)	3.43	3.43	3.9	6.67	0	0	0	100	100	100
VW FM (Stock Price > \$5)	3.43	3.43	3.9	7.41	0	0	0	100	100	100

Notes: Benchmark t-statistics, percentage of H_0 and SnM rejected from Multivariate FM regressions in Table A.1.12. Results are presented under single hypothesis test with level of level of significance of 5% and under of the Bonferroni, Holm, and BHY multiple hypothesis tests.

Table A.1.14: : Top Double Sorted Portfolios by t-statistics

Signals			Significance		
S_1	S_2	PF	SR	α_{S_1,S_2}	$t_{\alpha_{S_1,S_2}}$
MEANEST	log bond age	$HH - HL$	0.12	0.23	7.43
Bond Skewness	Bond Kurtosis	$HH - HL$	−0.22	−0.44	−7.28
sgr	log bond age	$HH - HL$	0.24	0.38	7.23
quick	log bond age	$LH - LL$	0.11	0.32	7.02
cashdebt	log bond age	$HH - HL$	0.07	0.21	6.98
roaq	log bond age	$HH - HL$	0.04	0.25	6.85
prccq	log bond age	$HH - HL$	0.19	0.22	6.6
salecash	log bond age	$HH - HL$	0.16	0.36	6.47
stdcf	log bond age	$LH - LL$	0.21	0.22	6.47
pps	log bond age	$HH - HL$	0.14	0.18	6.4
ipd	log bond age	$HH - HL$	0.18	0.28	6.33
currat	log bond age	$LH - LL$	0.16	0.33	6.25
invest	log bond age	$HH - HL$	0.12	0.3	6.22
lev	log bond age	$LH - LL$	0.11	0.19	6.18
roic	log bond age	$HH - HL$	0.1	0.25	6.01
cash	log bond age	$LH - LL$	0.13	0.36	6.01
pricedelay	log bond age	$HH - HL$	0.24	0.28	5.85
cashpr	log bond age	$HH - HL$	0.1	0.24	5.74
Bond Skewness	log bond age	$LH - LL$	0.22	0.23	5.72
std dolvol	log bond age	$LH - LL$	0.23	0.27	5.71
chadv	log bond age	$HH - HL$	0.17	0.33	5.46
d ivol	Mac Dur	$HH - LH$	−0.28	−0.9	−5.43
roeq	Bond Kurtosis	$HH - HL$	−0.09	−0.28	−5.39
mom12m	log bond age	$HH - HL$	0.24	0.28	5.37
sgrvol	log bond age	$LH - LL$	0.11	0.22	5.34
sgrvol	d ivol	$LH - LL$	−0.28	−0.24	−5.36
grcapx	log bond age	$HH - HL$	0.23	0.21	5.31
mom1m	baspread	$HH - LH$	0.31	0.87	5.3
depr	log bond age	$LH - LL$	0.05	0.27	5.27
lev	pvol	$LH - LL$	−0.19	−0.22	−5.28
roaq	d ivol	$HH - HL$	−0.08	−0.23	−5.32
roic	pvol	$HH - HL$	−0.24	−0.28	−5.29

depr	cvol	$LH-LL$	−0.24	−0.52	−5.23
pchcapx ia	log bond age	$LH-LL$	0.19	0.25	5.2
cashpr	log bond age	$LH-LL$	0.22	0.47	5.2
tb	Bond Kurtosis	$HH-HL$	−0.27	−0.33	−5.18
cvol	Mac Dur	$HH-LH$	−0.23	−0.8	−5.19
pchcapx	log bond age	$HH-HL$	0.2	0.25	5.17
turn	log bond age	$LH-LL$	0.11	0.14	5.17
herf	log bond age	$LH-LL$	0.11	0.23	5.17
cf	log bond age	$HH-HL$	0.1	0.24	5.17
bm	cvol	$LH-LL$	−0.23	−0.42	−5.12
pchquick	d ivol	$LH-LL$	−0.15	−0.65	−5.09
maxret	log bond age	$LH-LL$	0.22	0.17	5.07
roic	d ivol	$HH-HL$	−0.28	−0.27	−5.13
roe	mom1m	$LH-LL$	0.27	0.92	5.02
sgrvol	pvol	$LH-LL$	−0.27	−0.25	−5.04
rdbias	mve m	$HH-LH$	0.07	−0.63	−12.27
sp	rdbias	$HH-LH$	0.27	0.66	6.11
pchdepr	rdbias	$HH-LH$	0.01	0.12	5.36
rdbias	potesh	$HH-LH$	0.75	−0.18	−10.93
rdbias	tb	$HH-LH$	−0.06	−0.54	−5.77

Notes: Conditional 5×5 double sorts of sorting on signal S_1 and then within each quintile of signal S_1 sort on the value of signal S_2. Column 3 (PF) reports the portfolio formation combination. Column 4 reports the monthly portfolio Sharpe Ratio (SR). Column 5 and column 6 report the quintile high minus low portfolio alpha (α_{S_1,S_2}) and corresponding t-statistics ($t_{\alpha_{S_1,S_2}}$) using the BBW four factor corporate bond return factor model. Alphas are reported in monthly percentages. t-statistics are reported using the Newey and West [1987] method adjusting for 3 lags in order to determine the statistical significance of the high minus low portfolio alpha coefficient. All portfolios are formed by one month lagged signals (listed in Table A.1.4) relative to the corporate bond excess returns. Numbers are reported up to four decimal places for ease of display.

Table A.1.15: : Corporate Bond Return Machine Learning Results

Panel A: All Corporate Bonds

Method	R^2_{OS}	Decile PS (BBW)			Quintiles (BBW)			Quintiles (BBW+FF5)		
		α_S	t_{α_S}	$S.R.$	α_S	t_{α_S}	$S.R.$	α_S	t_{α_S}	$S.R.$
OLS	−4.59	0.24	2.21	0.2	0.18	1.97	0.17	0.16	1.81	0.17
LASSO	3.7	0.47	3.88	0.43	0.31	3.5	0.36	0.29	3.23	0.36
RIDGE	3.46	−0.01	−0.04	0.24	−0.01	−0.05	0.23	0	0.04	0.23
Enet	3.7	0.48	3.9	0.43	0.31	3.5	0.36	0.29	3.24	0.36
PCR	3.42	−0.22	−1.32	0.2	−0.15	−1.26	0.21	−0.12	−1.04	0.21
PLS	−0.09	0.21	1.71	0.3	0.13	1.44	0.27	0.12	1.27	0.27

Panel B: Investment Grade

Method	R^2_{OS}	Decile PS (BBW)			Quintiles (BBW)			Quintiles (BBW+FF5)		
		α_S	t_{α_S}	$S.R.$	α_S	t_{α_S}	$S.R.$	α_S	t_{α_S}	$S.R.$
OLS	−2.87	0.07	0.64	0.05	0.02	0.28	0.06	0	0.04	0.06
LASSO	3.81	0.27	2.03	0.22	0.18	1.65	0.17	0.13	1.22	0.17
RIDGE	3.83	0.07	0.53	0.15	0.05	0.54	0.14	0.05	0.49	0.14
Enet	3.81	0.27	2.02	0.22	0.18	1.66	0.17	0.13	1.22	0.17
PCR	3.48	−0.24	−1.55	0.08	−0.16	−1.49	0.08	−0.14	−1.36	0.08
PLS	1.36	0.1	0.76	0.13	0.08	0.8	0.11	0.04	0.42	0.11

Panel C: Speculative Grade

Method	R^2_{OS}	Decile PS (BBW)			Quintiles (BBW)			Quintiles (BBW+FF5)		
		α_S	t_{α_S}	$S.R.$	α_S	t_{α_S}	$S.R.$	α_S	t_{α_S}	$S.R.$
OLS	−9.38	0.25	1.4	0.1	0.16	1.39	0.05	0.16	1.35	0.05
LASSO	3.77	0.43	1.86	0.22	0.33	1.99	0.22	0.35	2.06	0.22
RIDGE	3.34	−0.03	−0.14	0.19	−0.02	−0.1	0.19	0.01	0.07	0.19
Enet	3.77	0.43	1.86	0.22	0.33	1.99	0.22	0.35	2.06	0.22
PCR	2.95	−0.38	−1.23	0.15	−0.24	−1.09	0.16	−0.21	−1	0.16
PLS	−1.56	0.15	0.97	0.15	0.16	1.33	0.23	0.2	1.61	0.23

Notes: Column 2 is the computed out-of-sample R^2_{OS} of the machine learning predicted corporate bond returns using ordinary least squares (OLS), penalized linear, LASSO, RIDGE regression, elastic net (Enet), principal component regression (PCR) and partial least square (PLS). Column 3 to 5 report the decile high-minus-low portfolio alpha (α_S), corresponding t-statistics (t_{α_S}), and monthly Sharpe Ratio (S.R.) when forming portfolios using the machine learning predicted corporate bond returns and adjusting the returns using the BBW four factor corporate bond return model. Columns 6 to 8 (9 to 11) report quintile high-minus-low portfolio alphas adjusted using the BBW (both the BBW and FF5) factor models. Alphas are reported in monthly percentages. t-statistics are reported using the Newey and West [1987] method adjusting for 3 lags in order to determine the statistical significance of the high minus low portfolio alpha coefficient. All portfolios are formed by one month lagged relative to the corporate bond excess returns. Details of the R^2_{OS} methodology and machine learning techniques are found in Appendix Section A.1.4 and A.1.5.

Table A.1.16: : Monthly Out-of-Sample Performance Statistics using the Diebold-Mariano Tests

	LASSO	RIDGE	Enet	PCR	PLS
OLS	5.88	5.55	5.88	5.38	5.81
LASSO		0	1.3	0.22	−3.69
RIDGE			0.01	0.3	−3.76
Enet				0.22	−3.69
PCR					−4.16

Notes: Table reports all pairwise Diebold and Mariano [1995] test statistics which compare the out-of-sample corporate bond return prediction amongst the machine learning models presented in Table A.1.15. A positive test statistic indicate that the model in the column outperforms that in the row model. A test statistic larger than 2 corresponds to statistically significant at the 5% level.

A.1.3 Overview of Multiple Hypothesis Testing Methods

In this section we briefly describe standard adjustments to p-values for multiple hypothesis testing that have been used in the finance literature: Bonferroni [1936] (Bonf/Bonferroni henceforth), Holm [1979] (Holm, henceforth), Benjamini and Hochberg [1995] (BH henceforth), and Benjamini and Yekutieli [2001] (BHY henceforth). For an excellent comprehensive overview of multiple hypothesis testing methods in finance see Harvey et al. [2016] and Harvey et al. [2020].

In the following descriptions of the different MHT methods we assume that there are a total of M tests and we choose the Family-wise Error rate (FWER), False-Discovery Rate (FDR), and False-Discovery Proportion (FDP) to be all equal $\alpha_{LOS} = 5\%$.

Bonferroni Method

The Bonferroni test is a single-step procedure that adjusts all p-values. The Bonferroni adjustment will reject any hypothesis with a $p - value \leq \frac{\alpha_{LOS}}{M}$, hence:

$$p_i^{Bonferroni} = \min\left[M \times p_i, 1\right].$$

The Bonferroni adjustment applies the same adjustment to each of the hypothesis tests.

Holm Method

The p-value adjustment of Holm [1979] is constructed as:

- Order the non-adjusted p-values such that $p_{(1)} \leq p_{(2)} \leq \cdots p_{(b)} \leq \cdots \leq p_{(M)}$ where $H_{(1)}, H_{(2)}, \cdots, H_{(b)}, \cdots, H_{(M)}$
- k is the minimum index such that $p_{(k)} > \frac{\alpha_{LOS}}{M+1-b}$
- Reject the null hypotheses: $H_{(1)}, \cdots, H_{(k-1)}$ but not reject $H_{(k)}, \cdots, H_{(M)}$

The equivalent adjusted p-values are then computed as:

$$p_i^{Holm} = \min\left[\max_{j \leq i}\{M - j + 1\}p_{(j)}, 1\right]$$

BHY Method

The p-value adjustment of Benjamini and Yekutieli [2001] is constructed as:

- Order the non-adjusted p-values such that $p_{(1)} \leq p_{(2)} \leq \cdots p_{(b)} \leq \cdots \leq p_{(M)}$ where $H_{(1)}, H_{(2)}, \cdots, H_{(b)}, \cdots, H_{(M)}$
- k is the maximum index such that $p_{(b)} > \frac{\alpha_{LOS} \times b}{M \sum_{j=1}^{M} \frac{1}{j}}$
- Reject the null hypotheses: $H_{(1)}, \cdots, H_{(k)}$ but not reject $H_{(k+1)}, \cdots, H_{(M)}$

The equivalent adjusted p-values are then computed as:

$$p_i^{BHY} = \min \left[p_{i+1}^{BHY}, \frac{M \times \sum_{j=1}^{M} \frac{1}{j}}{i} p_i \right] \text{ if } i \leq M-1 \text{ and } p_i^{BHY} = p_{(M)} \text{ if } i = M$$

A.1.4 Machine Learning Performance Evaluation

I compare and evaluate a variety of machine learning methods in predicting the cross-section of corporate bond returns, including the ordinary least squares (OLS) with all covariates; penalized linear regression methods such as LASSO of Tibshirani [1996], RIDGE regression of Hoerl and Kennard [1970], and elastic net (Enet) of Zhou and Hastie [2005]; dimension reduction techniques such as principal component analysis (PCA) and partial least square (PLS). I use the out-of-sample R^2_{OS} as the performance metric to assess the predictive power of individual bond return predictors. The R^2_{OS} statistic pools prediction errors across bonds and over time into a grand panel-level assessment of each model, and it measures the proportional reduction in mean squared forecast error (MSFE) for each model relative to a naive forecast of zero benchmark, which assumes that the one-month-ahead expected return on corporate bonds equals the time $t+1$ risk-free rate. To estimate the out-of-sample R^2_{OS} I follow:

$$R^2_{OS} = 1 - \frac{\sum_{(i,t) \in \tau_3} \left(r_{i,t} - \hat{r}_{i,t} \right)^2}{\sum_{(i,t) \in \tau_3} r_{i,t}^2} \tag{A.1}$$

and the most commonly used approach in the literature and divide our full sample (July 2002 to October 2019) into three disjoint time periods; (i) the first five years of training or estimation period, τ_1, (ii) the second two years of validation for tuning the hyperparameters, τ_2, and (iii) the rest of the sample as the test period, τ_3, to evaluate a model's predictive power, which represents the truly out-of-sample evaluation of the model's performance. I use the mean squared forecast error MSFE-adjusted statistic of Campbell and Thompson [2008] to test the statistical significance of R^2_{OS}. Given that there will be strong cross-sectional dependence among individual excess bond returns, as is standard in the literature I employ the modified MSFE-adjusted statistic of Clark and West [2007] based on the cross-sectional average of prediction errors from each model instead of prediction errors among individual returns. The p-value from the MSFE-adjusted statistic tests the null hypothesis that the MSFE of a naive forecast of zero is less than or equal to the MSFE of a machine learning model against the one-sided (upper-tail) alternative hypothesis that the MSFE of a naive forecast of zero is greater than the MSFE of a machine learning model (H_0: $R^2_{OS} \leq 0$ against H_A: $R^2_{OS} > 0$).

The predictive ability of two machine learning methods is compared to one another using the modified Diebold and Mariano [1995] test statistic $DM_{12} = \bar{d}_{12}/\hat{\sigma}_{\bar{d}}$ where $\bar{d}_{12}$ and $\hat{\sigma}_{\bar{d}}$ are the time-series average and Newey and West [1987] standard error (respectively) of the forecast error differentials (denoted $d_{12,t+1}$) from each model over the period $t+1$. The forecast error differential is computed as $d_{12,t+1} = \sum_{i=1}^{n_3} \left((\hat{e}^{(1)}_{i,t+1})^2 - (\hat{e}^{(2)}_{i,t+1})^2 \right) / n_{3,t+1}$ where $\hat{e}^{(1)}_{i,t+1}$ and $\hat{e}^{(1)}_{i,t+1}$ are the corporate bond return forecasts from each of model under comparison and $n_{3,t+1}$ is the number of observations.

A.1.5 Machine Learning Methodologies

The corporate bond excess return ($r_{i,t+1}$) at month $t+1$ is defined as $r_{i,t+1} = \mathbb{E}_{t+1}\left[r_{i,t+1}\right] + \epsilon_{i,t+1}$ where $\mathbb{E}_{t+1}\left[r_{i,t+1}\right] = g^*\left(z_{i,t}\right)$ is the time t expected return with the function $g\left(\cdot\right)$ is a function of the matrix set of characteristics/factors $z_{i,t} = \left(z_{1,t}, \ldots, z_{K,t}\right)^{'}$ with $i = 1, \ldots, N$ indexing the number of characteristics/factors and $t = 1, \ldots, T$ indexes the number of months.

Linear Regression

The linear regression framework assumes that $g^*\left(z_{i,t}; \theta\right) = z^{'}_{i,t}\theta$ where $\theta = \left(\theta_1, \cdots, \theta_K\right)^{'}$ can be estimated by the ordinary least squares objective function optimization problem:

$$\min_{\theta} \mathcal{L}\left(\theta\right) = \frac{1}{NT} \sum_{i=1}^{N} \sum_{t=1}^{T} \left(r_{i,t+1} - g\left(z_{i,t}; \theta\right)\right)^2 \tag{A.2}$$

The estimate of the sensitivities θ is unbiased as long as the number of characteristics/ factors (K) is small relative to the length of the time series of observations T. Since K is usually larger than T the linear regression estimates are typically inefficient and inconsistent and result in in-sample over-fitting of the estimates.

Penalized Linear Regression: LASSO, RIDGE, and Enet

To deal with the over-fitting in linear regression (i.e. when the number of characteristics/factors (K) is large relative to the length of the time series of observations T) it is typically common to include a penalty term to the objective function equation A.2 in order to impose a tradeoff between the in-sample performance and out-of-sample performance. When adding a penalty function $\min_\theta \mathcal{L}(\theta;\cdot) = \mathcal{L}(\theta) + \phi(\theta;\cdot)$ where $\phi(\theta;\cdot)$ is the penalty function of θ. The specification of the functions form of the penalty $\phi(\theta;\cdot)$ are quite general to allow for the θ that can be shrunk towards zero (or equal to zero). A generalized penalty function can be written as:

$$\phi(\theta;\lambda,\rho) = \lambda(1-\rho)\sum_{j=1}^{P}|\theta_j| + \frac{1}{2}\lambda\rho\sum_{j=1}^{P}\theta_j^2 \tag{A.3}$$

The penalized estimation in equation A.2 simplifies to simple linear regression when the hyperparameter (λ) controlling the amount of shrinkage, is $\lambda = 0$.

When $\rho = 0$ then equation A.2 corresponds to the LASSO models estimation which will set values of the vector θ exactly to zero. When $\rho = 1$ then equation A.2 corresponds to the RIDGE regression model estimation which will shrink values of the vector θ closer to zero but not exactly to zero. For a value of ρ between zero and one, then equation A.2 corresponds to the Enet regression model estimation which is a combination of the RIDGE and LASSO estimation techniques which allows for a combination of sparse and dense modeling.

Dimension Reduction: PCA and PLS

The excess monthly corporate bond return can be written as $r_{i,t+1} = z_{i,t}^{'}\theta + \epsilon_{i,t+1}$ which in matrix notation is $R = Z\theta + E$ where R is an $NT \times 1$ vector of $r_{i,t+1}$, Z is a $NT \times K$ matrix of stacked predictors $z_{i,t}$ and E is an $NT \times 1$ vector of residuals $\epsilon_{i,t+1}$.

Dimension reduction techniques are a useful way to handle the over-fitting problems in ordinary least squares regression when the number of characteristics/factors K is larger than the time-series of available data T. Dimension reduction projects a larger number of characteristics/factors into a small number of factors.

The main dimension reduction techniques are principal component analysis (PCA) and the partial least squares (PLS). PCA transforms the set of characteristics in independent components where the first component has the largest variance and then the second and so on. The decomposition is such that:

$$w_j = \max_w Var\,(Zw), \text{ s.t. } w^{'}w = 1,\ Cov\,(Zw, Zw_l) = 0,\ l = 1, 2, \cdots, j-1$$

PLS links the K linear combinations of Z for covariance maximization with the corporate bond returns (R). The weights of the j-th PLS component solve for:

$$\max_{\nu, w_j} Cov\,(R\nu, Zw_j), \text{ s.t. } w_j^{'}w_j = 1,\ Cov\,(Zw_j, Zw_l) = 0,\ l = 1, 2, \cdots, j-1$$

Appendix B

Appendix for Chapter 3

This appendix contains tables from Chapter 3.

B.1 Tables

Table B.1.1: : Implicit leverage in Options

Moneyness	Time-to-maturity (days)		
	d2mat ≤ 30	$30 <$ d2mat ≤ 179	d2mat ≥ 180
In-the-money	0.845	0.766	0.657
At-the-money	0.884	0.831	0.768
Out-of-the-money	0.915	0.881	0.839
Moneyness	Time-to-maturity (days)		
	d2mat ≤ 30	$30 <$ d2mat ≤ 179	d2mat ≥ 180
In-the-money	0.197	0.378	0.618
At-the-money	0.142	0.236	0.393
Out-of-the-money	0.118	0.204	0.359

Notes: Panel A is mean leverage for call options which is calculated as $Leverage = \frac{Ke^{-rT}N(d_2)}{\left(S_0 - \sum_{i=1}^{I} D_i e^{-rt_i}\right)N(d_1)}$ where $d_1 = \frac{\log\left(\frac{S_0 - \sum_{i=1}^{I} D_i e^{-rt_i}}{K}\right) + \left(r + \frac{\sigma^2}{2}\right)T}{\sigma\sqrt{T}}$, $d_2 = d_1 - \sigma\sqrt{T}$, S_0 is the stock price, $\sum_{i=1}^{I} D_i e^{-rt_i}$ is the present value of dividends paid over the life of the option, K is the strike price of the option, r is the risk-free interest rate, σ is the stock volatility, T is the time to expiration, and $N(\cdot)$ is the standard Normal cumulative distribution function. Panel B is mean put collateral requirements. The implicit collateral for puts is calculated as $Collateral = \frac{Ke^{-rT}N(-d_2)}{\left(S_0 - \sum_{i=1}^{I} D_i e^{-rt_i}\right)N(-d_1)}$

Table B.1.2: : Distribution of the Number of Option Contracts Traded per Month

	All Stocks				
	Mean	Median	Std. Dev.	Min.	Max.
$10 \geq$ d2mat ≤ 30	482	226	$1,334$	1	$147,196$
$30 <$ d2mat ≤ 60	378	194	893	0	$104,670$
$60 <$ d2mat ≤ 90	335	154	892	1	$100,634$
d2mat ≥ 90	288	142	585	0	$43,873$
	Top 500 Stocks				
	Mean	Median	Std. Dev.	Min.	Max.
$10 \geq$ d2mat ≤ 30	806	385	$2,044$	1	$127,942$
$30 <$ d2mat ≤ 60	552	293	$1,369$	0	$104,670$
$60 <$ d2mat ≤ 90	442	212	$1,263$	1	$100,634$
d2mat ≥ 90	366	197	719	3	$43,873$

Notes: Distribution of the Number of Option Contracts Traded per Month across different times to maturity for all stocks and top 500 largest firms based on firm size.

Table B.1.3: : Summary statistics for option-based measures of stock mispricing

	Mean	Median	Std. Dev.	Min.	Max.
IPD	−0.135	−0.035	1.952	−96.526	78.817
Imp. Fee.	0.008	0.005	0.037	−1.551	0.914
CW	−0.012	−0.006	0.069	−2.169	2.039
$\Delta CVOL$	−0.001	−0.002	0.164	−2.562	2.544
$\Delta PVOL$	−0.001	−0.002	0.165	−2.598	2.749
Skewness	0.064	0.048	0.076	−0.885	2.218
PP	0.567	0.542	0.265	0	1
OOI	0	0	0.001	−0.091	0.083
O/S	0.09	0.051	0.144	0	17.283

Notes: Summary statistics for option-based measures of stock mispricing.

Table B.1.4: : Average Indicative Short Selling Stock Fee

	# Stocks	1st Prcnt	Median	80th Prcnt	99th Prcnt
2004	3,295	0.375	0.482	0.625	6.941
2005	3,875	0.409	0.476	0.804	11.125
2006	4,109	0.392	0.464	1.352	12.838
2007	4,237	0.387	0.458	1.181	12.161
2008	4,086	0.391	0.469	1.839	17.161
2009	3,931	0.271	0.338	0.71	24.474
2010	3,936	0.335	0.373	0.897	32.819
2011	4,008	0.367	0.381	1.75	47.899
2012	3,924	0.371	0.397	1.885	41.131
2013	3,874	0.372	0.392	1.72	39.032

Notes: Average indicative short selling fee for stocks during 2004 to 2013

Table B.1.5: : Average Firm Size and Average Indicative Fee Portfolio Sorts

Panel A: Average Firm Size

	IPD	Implied Fees	CW	$\Delta CVOL$	$\Delta PVOL$	Skewness	PP	OOI	O/S
Low	2.75	9.3	3.08	4.38	4.49	8.16	15.18	7.45	3.08
2	7.96	12.49	7.65	8.57	8.51	13.58	17.83	11.26	7.65
3	13.82	10.16	10.31	10.43	10.24	8.36	9.33	3.53	10.31
4	12.12	4.82	9.82	8.94	8.99	4.21	4.89	9.72	9.82
High	5.32	1.7	5.66	4.61	4.82	2.65	2.65	7.17	5.66

Panel B: Average Indicative Fee

	IPD	Implied Fees	CW	$\Delta CVOL$	$\Delta PVOL$	Skewness	PP	OOI	O/S
Low	0.3388	0.0556	0.3707	0.175	0.1792	0.1596	0.1841	0.1659	0.0669
2	0.112	0.0313	0.0698	0.0897	0.0899	0.0765	0.1078	0.0741	0.0858
3	0.0526	0.0336	0.0384	0.0741	0.0736	0.0857	0.107	0.1115	0.1062
4	0.0499	0.0592	0.0364	0.0864	0.0819	0.1257	0.1087	0.0959	0.1335
High	0.1074	0.3072	0.1234	0.1705	0.1713	0.155	0.1001	0.1715	0.2113

Notes: Characteristics of portfolios formed from option-based mispricing measures. Panel A. Average firm sizes. Firm sizes, in billions of dollars are calculated for each portfolio each month and then averaged across months. Panel B. The proportion of stocks in quintile portfolios formed on option-based mispricing measures that are hard to borrow. A stock is defined as hard-to-borrow if the stocks indicative borrowing fee from Markit is among the highest 20% across all stocks.

Table B.1.6: : Value-Weighted Portfolio Sorts

Panel A: IPD

Port.	r_{t+1}	α_{t+1}	$t_{\alpha_{t+1}}$	r_{t+2}	α_{t+2}	$t_{\alpha_{t+2}}$	r_{t+3}	α_{t+3}	$t_{\alpha_{t+3}}$
Low	0.3543	−0.381	−2.33	0.3887	−0.4061	−2.34	0.4275	−0.4094	−1.92
2	0.4851	−0.1244	−1.25	0.4884	−0.1466	−1.2	0.4833	−0.203	−1.89
3	0.4759	−0.0934	−1.51	0.4651	−0.1472	−2.09	0.5602	−0.116	−1.76
4	0.7056	0.1628	2.48	0.7303	0.1618	1.82	0.7514	0.1572	1.81
High	0.8748	0.2553	2.95	1.0216	0.3875	3.75	1.0427	0.4299	3.52
H-L	0.5205	0.6364	3.31	0.6329	0.7936	3.59	0.6151	0.8392	2.89

Panel B: Implied Fees

Port.	r_{t+1}	α_{t+1}	$t_{\alpha_{t+1}}$	r_{t+2}	α_{t+2}	$t_{\alpha_{t+2}}$	r_{t+3}	α_{t+3}	$t_{\alpha_{t+3}}$
Low	0.6517	0.1298	1.59	0.5051	−0.0357	−0.32	0.6575	0.0776	0.57
2	0.6525	0.0961	1.39	0.5722	−0.0077	−0.14	0.6479	0.0582	0.69
3	0.4827	−0.1078	−1.96	0.7039	0.0658	0.97	0.5866	−0.057	−0.93
4	0.5816	−0.0656	−0.52	0.6339	−0.0443	−0.45	0.5891	−0.1253	−1.04
High	0.3621	−0.3873	−2.65	0.4931	−0.3061	−2.27	0.4698	−0.3829	−2.37
H-L	-0.2896	−0.5171	−2.87	−0.0121	−0.2704	−1.54	−0.1877	−0.4604	−2.31

Panel C: CW

Port.	r_{t+1}	α_{t+1}	$t_{\alpha_{t+1}}$	r_{t+2}	α_{t+2}	$t_{\alpha_{t+2}}$	r_{t+3}	α_{t+3}	$t_{\alpha_{t+3}}$
Low	0.5901	−0.1406	−0.96	0.7604	0.0015	0.01	0.6582	−0.1163	−0.7
2	0.5256	−0.0804	−0.59	0.5538	−0.09	−0.69	0.6023	−0.0387	−0.35
3	0.6532	0.1068	1.68	0.6629	0.0965	1.25	0.6562	0.0469	0.71
4	0.51	−0.0546	−0.66	0.5427	−0.0543	−0.71	0.5315	−0.0872	−0.94
High	0.5645	−0.082	−0.5	0.7768	0.0694	0.41	0.6924	−0.0397	−0.21
H-L	-0.0256	0.0586	0.25	0.0165	0.0679	0.27	0.0343	0.0766	0.26

Panel D: Skewness

Port.	r_{t+1}	α_{t+1}	$t_{\alpha_{t+1}}$	r_{t+2}	α_{t+2}	$t_{\alpha_{t+2}}$	r_{t+3}	α_{t+3}	$t_{\alpha_{t+3}}$
Low	0.7612	0.1844	1.45	0.7804	0.2002	1.39	0.9092	0.2952	2
2	0.6021	0.0918	1.69	0.662	0.1141	1.46	0.6393	0.078	1.23
3	0.6423	0.0468	0.49	0.579	−0.0559	−0.52	0.6261	−0.0273	−0.24
4	0.6836	0.0366	0.28	0.8197	0.1198	1.1	0.7085	0.0261	0.22
High	0.451	−0.2096	−1.28	0.3976	−0.2835	−1.41	0.4417	−0.3091	−1.67
H-L	-0.3102	−0.394	−1.77	−0.3828	−0.4837	−2.03	−0.4675	−0.6044	−2.3

Panel E: $\Delta CVOL$

Port.	r_{t+1}	α_{t+1}	$t_{\alpha_{t+1}}$	r_{t+2}	α_{t+2}	$t_{\alpha_{t+2}}$	r_{t+3}	α_{t+3}	$t_{\alpha_{t+3}}$
Low	0.5607	−0.0641	−0.35	0.5975	−0.1296	−0.72	1.012	0.2698	1.34
2	0.619	0.1046	1.14	0.7728	0.1269	1.34	0.5353	−0.1086	−0.99
3	0.6078	0.0655	0.71	0.6925	0.0915	1.11	0.705	0.1068	1.13

4	0.6086	0.0222	0.18	0.5605	−0.0513	−0.48	0.5791	−0.0856	−0.73
High	0.3618	−0.4087	−2.24	0.6098	−0.1354	−0.88	0.5897	−0.1975	−0.96
H-L	-0.1988	−0.3447	−1.21	0.0123	−0.0058	−0.03	−0.4224	−0.4674	−1.39

Panel F: $\Delta PVOL$

Port.	r_{t+1}	α_{t+1}	$t_{\alpha_{t+1}}$	r_{t+2}	α_{t+2}	$t_{\alpha_{t+2}}$	r_{t+3}	α_{t+3}	$t_{\alpha_{t+3}}$
Low	0.8217	0.2038	1.29	0.6306	−0.1042	−0.61	1.0225	0.2867	1.89
2	0.6273	0.1231	1.12	0.7736	0.1549	1.51	0.6471	0.0122	0.11
3	0.6012	0.0764	1.01	0.6994	0.109	1.23	0.7262	0.1243	1.54
4	0.6455	0.0483	0.52	0.5376	−0.0855	−0.72	0.5172	−0.142	−1.78
High	0.2555	−0.5231	−3.21	0.6599	−0.081	−0.6	0.4974	−0.3325	−2
H-L	-0.5662	−0.7269	−2.73	0.0293	0.0232	0.1	−0.5252	−0.6192	−2.36

Panel G: PP

Port.	r_{t+1}	α_{t+1}	$t_{\alpha_{t+1}}$	r_{t+2}	α_{t+2}	$t_{\alpha_{t+2}}$	r_{t+3}	α_{t+3}	$t_{\alpha_{t+3}}$
Low	0.5498	−0.0758	−0.63	0.6181	−0.0524	−0.49	0.6122	−0.0759	−0.7
2	0.6045	0.0524	0.78	0.545	−0.0385	−0.64	0.6292	0.0227	0.44
3	0.5775	0.0061	0.07	0.5735	−0.0418	−0.43	0.5922	−0.0488	−0.42
4	0.6404	0.0272	0.27	0.7123	0.0999	1.03	0.5782	−0.1312	−1.16
High	0.6073	−0.0379	−0.33	0.7283	0.0421	0.37	0.7843	0.0878	0.81
H-L	0.0575	0.0378	0.18	0.1102	0.0944	0.51	0.1721	0.1637	0.9

Panel H: OOI

Port.	r_{t+1}	α_{t+1}	$t_{\alpha_{t+1}}$	r_{t+2}	α_{t+2}	$t_{\alpha_{t+2}}$	r_{t+3}	α_{t+3}	$t_{\alpha_{t+3}}$
Low	0.5066	−0.147	−1.42	0.6312	−0.0309	−0.26	0.5803	−0.1599	−1.24
2	0.6168	0.1003	1.54	0.6018	0.0619	0.73	0.6056	0.0094	0.14
3	0.5303	−0.0551	−0.48	0.7245	0.1179	1.15	0.7687	0.1664	1.77
4	0.5565	0.0148	0.13	0.7313	0.1268	1.31	0.7781	0.1686	2.14
High	0.6658	−0.0437	−0.34	0.4304	−0.3469	−2.7	0.6155	−0.1287	−1.06
H-L	0.1593	0.1032	0.57	−0.2008	−0.316	−1.65	0.0352	0.0313	0.15

Panel I: O/S

Port.	r_{t+1}	α_{t+1}	$t_{\alpha_{t+1}}$	r_{t+2}	α_{t+2}	$t_{\alpha_{t+2}}$	r_{t+3}	α_{t+3}	$t_{\alpha_{t+3}}$
Low	0.8463	0.256	3.81	0.8914	0.2675	3.21	0.8955	0.2688	3.74
2	0.8082	0.1873	2.54	0.884	0.2388	2.69	0.9105	0.2383	2.68
3	0.7696	0.1531	2.17	0.8006	0.1623	1.95	0.7604	0.1226	1.82
4	0.6993	0.1066	1.33	0.6941	0.0676	0.81	0.7751	0.1553	2.02
High	0.3645	−0.1854	−4.37	0.3877	−0.2005	−2.41	0.3971	−0.2311	−3.94
H-L	-0.4818	−0.4414	−4.73	−0.5037	−0.468	−5.62	−0.4984	−0.5019	−4.73

Notes: Each month over 2004-2013, stocks are sorted into quintiles based on option based measures of stock mispricing. IPD is the average percentage difference between implied and actual stock prices. Implied borrowing fees is the stock borrowing fee implied by violations of put-call parity. CW is the difference between implied volatilities of calls and puts. Skewness is the difference in implied volatilities of a put with a delta of 0.2 and a call with a delta of 0.5. $\Delta PVOL$ is the monthly change in implied volatilities for 30 day puts with deltas of -0.5. $\Delta CVOL$ is the change analogous change in call implied volatilities. PP is the ratio of put buy volume that opens positions to the sum of put and call buy volume. OOI is the difference between synthetic positive and negative options volume. O/S is the natural logarithm of the ratio of stock and options volume. The FFC four factor model is used to calculate portfolio α. Newey and West [1987] adjusted standard errors with three lags are used. Returns and portfolio α are in percentages.

Table B.1.7: : Value Weighted Portfolio Sorts Including and Excluding Hard to Short Stocks

	α_{t+1}		α_{t+2}		α_{t+3}	
	Include	Exclude	Include	Exclude	Include	Exclude
Panel A: IPD						
Low	−0.381	−0.0597	−0.4061	0.0635	−0.4094	−0.1291
	(−2.33)	(−0.33)	(−2.34)	(0.26)	(−1.92)	(−0.78)
High	0.2553	0.3412	0.3875	0.4562	0.4299	0.5003
	(2.95)	(3.78)	(3.75)	(3.96)	(3.52)	(3.77)
High-Low	0.6364	0.4009	0.7936	0.3926	0.8392	0.6294
	(3.31)	(1.93)	(3.59)	(1.88)	(2.89)	(2.72)

	α_{t+1}		α_{t+2}		α_{t+3}	
	Include	Exclude	Include	Exclude	Include	Exclude
Panel B: Implied Fees						
Low	0.1298	0.2615	−0.0357	0.0314	0.0776	0.1758
	(1.59)	(3.09)	(−0.32)	(0.29)	(0.57)	(1.14)
High	−0.3873	−0.0499	−0.3061	0.0173	−0.3829	0.0088
	(−2.65)	(−0.33)	(−2.27)	(0.13)	(−2.37)	(0.05)
High-Low	−0.5171	−0.3114	−0.2704	−0.0141	−0.4604	−0.167
	(−2.87)	(−1.61)	(−1.54)	(−0.08)	(−2.31)	(−0.68)

	α_{t+1}		α_{t+2}		α_{t+3}	
	Include	Exclude	Include	Exclude	Include	Exclude
Panel C: CW						
Low	−0.1406	−0.0336	0.0015	0.2507	−0.1163	−0.0173
	(−0.96)	(−0.25)	(0.01)	(1.98)	(−0.7)	(−0.14)
High	−0.082	0.0446	0.0694	0.1984	−0.0397	0.1999
	(−0.5)	(0.23)	(0.41)	(1.17)	(−0.21)	(1.26)

High-Low	0.0586	0.0783	0.0679	−0.0523	0.0766	0.2172
	(0.25)	(0.3)	(0.27)	(−0.22)	(0.26)	(0.96)

	α_{t+1}		α_{t+2}		α_{t+3}	
	Include	Exclude	Include	Exclude	Include	Exclude
Panel D: $\Delta CVOL$						
Low	−0.0641	0.0325	−0.1296	0.0763	0.2698	0.2432
	(−0.35)	(0.2)	(−0.72)	(0.45)	(1.34)	(1.3)
High	−0.4087	−0.208	−0.1354	−0.0084	−0.1975	−0.1832
	(−2.24)	(−1.16)	(−0.88)	(−0.06)	(−0.96)	(−0.88)
High-Low	−0.3447	−0.2405	−0.0058	−0.0846	−0.4675	−0.4264
	(−1.21)	(−0.89)	(−0.03)	(−0.39)	(−1.39)	(−1.27)

	α_{t+1}		α_{t+2}		α_{t+3}	
	Include	Exclude	Include	Exclude	Include	Exclude
Panel E: $\Delta PVOL$						
Low	0.2038	0.23	−0.1042	0.0072	0.2867	0.3099
	(1.29)	(1.54)	(−0.61)	(0.05)	(1.89)	(2.03)
High	−0.5231	−0.3885	−0.081	0.0637	−0.3325	−0.2513
	(−3.21)	(−2.47)	(−0.6)	(0.5)	(−2)	(−1.58)
High-Low	−0.7269	−0.6185	0.0232	0.0565	−0.6192	−0.5611
	(−2.73)	(−2.35)	(0.1)	(0.26)	(−2.36)	(−2.17)

	α_{t+1}		α_{t+2}		α_{t+3}	
	Include	Exclude	Include	Exclude	Include	Exclude
Panel F: Skewness						
Low	0.1844	0.2682	0.2002	0.2171	0.2952	0.2882
	(1.45)	(2.18)	(1.39)	(1.62)	(2)	(1.94)
High	−0.2096	−0.0178	−0.2835	0.0597	−0.3091	−0.1068
	(−1.28)	(−0.11)	(−1.41)	(0.38)	(−1.67)	(−0.7)
High-Low	−0.394	−0.2861	−0.4837	−0.1574	−0.6044	−0.395
	(−1.77)	(−1.29)	(−2.03)	(−0.76)	(−2.3)	(−1.64)

	α_{t+1}		α_{t+2}		α_{t+3}	
	Include	Exclude	Include	Exclude	Include	Exclude
Panel G: PP						
Low	−0.0758	0.0597	−0.0524	0.0182	−0.0759	−0.0827
	(−0.63)	(0.47)	(−0.49)	(0.19)	(0.7)	(−0.81)
High	−0.0379	−0.0002	0.0421	0.0111	0.0878	0.1134

	(−0.33)	(0)	(0.37)	(0.93)	(0.81)	(0.99)
High-Low	0.0378	−0.0599	0.0944	0.0929	0.1637	0.1961
	(0.18)	(−0.28)	(0.51)	(0.51)	(0.9)	(1.07)

	α_{t+1}		α_{t+2}		α_{t+3}	
	Include	Exclude	Include	Exclude	Include	Exclude
Panel H: OOI						
Low	−0.147	−0.1051	−0.0309	0.0585	−0.1599	−0.0612
	(−1.42)	(−1.02)	(−0.26)	(0.53)	(−1.24)	(−0.55)
High	−0.0437	0.1434	−0.3469	−0.1844	−0.1287	−0.018
	(−0.34)	(1.09)	(−2.7)	(−1.45)	(−1.06)	(−0.13)
High-Low	0.1032	0.2486	−0.316	−0.2429	0.0313	0.0433
	(0.57)	(1.32)	(−1.65)	(−1.32)	(0.15)	(0.21)

	α_{t+1}		α_{t+2}		α_{t+3}	
	Include	Exclude	Include	Exclude	Include	Exclude
Panel I: O/S						
Low	0.256	0.2741	0.2675	0.2993	0.2688	0.2905
	(3.81)	(3.9)	(3.21)	(3.75)	(3.74)	(3.87)
High	−0.1854	−0.0662	−0.2005	−0.0714	−0.2311	−0.1128
	(−4.37)	(−1.55)	(−2.41)	(−1.81)	(−3.94)	(−2.37)
High-Low	−0.4414	−0.3402	−0.468	−0.3706	−0.5019	−0.4033
	(−4.37)	(−3.57)	(−5.62)	(−3.57)	(−4.73)	(−4.03)

Notes: Four-factor average alphas for quintile portfolios formed on the basis of option-based measures of stock mispricing when the portfolio include and exclude hard-to-borrow stocks. A stock is defined as hard-to-borrow if the indicative fee to borrow the stock is among the highest 20% across all stocks. Alphas are in percent (i.e. 0.61 is 61 basis points) and are for the first three months following portfolio formation. T-statistics are in parentheses under the alphas. They are calculated using Newey-West adjusted standard errors with three lags.

Table B.1.8: : High, Low, and H-L Returns and Alphas over Time

	IPD			Skewness			O/S		
Month	High	Low	H-L	High	Low	H-L	High	Low	H-L
t+1	0.2553	−0.381	0.6364	−0.2096	0.1844	−0.394	−0.1854	0.256	−0.4414
	(2.95)	(−2.33)	(3.31)	(−1.28)	(1.45)	(−1.77)	(−4.37)	(3.81)	(−4.73)
t+2	0.3875	−0.4061	0.7936	−0.2835	0.2002	−0.4837	−0.2005	0.2675	−0.468
	(3.75)	(−2.34)	(3.59)	(−1.41)	(1.39)	(−2.03)	(−4.3)	(3.53)	(−4.76)
t+3	0.4299	−0.4094	0.8392	−0.3091	0.2952	−0.6044	−0.2331	0.2688	−0.5019

	(3.52)	(−1.92)	(2.89)	(−1.67)	(2)	(−2.3)	(−3.94)	(3.74)	(−4.73)
t+4	0.3379	−0.4127	0.7506	−0.2865	0.4246	−0.711	−0.1953	0.2728	−0.4681
	(3.18)	(−1.53)	(2.58)	(−1.24)	(2.62)	(−2.2)	(−3.28)	(3.56)	(−4.24)
t+5	0.3059	−0.6516	0.9575	−0.4286	0.4739	−0.9025	−0.1934	0.2135	−0.4068
	(3.13)	(−2.11)	(2.8)	(−1.61)	(3.04)	(−2.66)	(−3.03)	(2.49)	(−3.57)
t+6	0.5547	−0.3136	0.8683	−0.3205	0.3658	−0.6864	−0.1873	0.2021	−0.3894
	(7.16)	(−1.23)	(3.07)	(−1.4)	(2.13)	(−2.07)	(−3.04)	(2.72)	(−3.93)
t+7	0.1334	−0.1236	0.257	−0.2575	0.1835	−0.441	−0.1252	0.2436	−0.3688
	(1.33)	(−0.54)	(0.94)	(−1.25)	(1.33)	(−1.52)	(−2.07)	(3.4)	(−3.6)
t+8	-0.0687	0.0855	−0.1542	−0.0488	0.1066	−0.1554	−0.1179	0.2319	−0.3498
	(−0.6)	(0.36)	(−0.57)	(−0.26)	(0.81)	(−0.59)	(−1.74)	(2.83)	(−3.32)
t+9	0.0527	0.0499	0.0028	−0.1092	0.2991	−0.4082	−0.1378	0.1845	−0.3223
	(0.42)	(0.21)	(0.01)	(−0.48)	(1.87)	(−1.26)	(−2.31)	(2.19)	(−2.94)
t+10	0.1536	−0.3196	0.4732	−0.1881	0.1998	−0.3879	−0.0935	0.1117	−0.2051
	(1.13)	(−2.08)	(2.25)	(−0.81)	(1.51)	(−1.46)	(−1.45)	(1.52)	(−1.87)
t+11	0.1692	−0.0545	0.2237	−0.1086	0.0118	−0.1204	−0.101	0.2051	−0.3061
	(1.31)	(−0.3)	(0.91)	(−0.51)	(0.08)	(−0.46)	(−1.46)	(2.33)	(−2.8)
t+12	0.3895	0.3044	0.0852	−0.0381	−0.0782	0.0401	−0.1497	0.186	−0.3357
	(2.79)	(1.07)	(0.27)	(−0.16)	(−0.6)	(0.14)	(−2.08)	(2.44)	(−3.33)
Sum	3.1009	−2.6323	5.7333	−2.5881	2.6667	−5.2548	−1.9201	2.6435	−4.5634

Notes: Fama-French four factor alphas for the 12 months following portfolio formation for quintile portfolios formed on option-based measures of stock mispricing. Portfolios are value-weighted. Average alphas for the high and low quintile portfolios, along with the difference in alphas are reported. Each month over 2004-2013, stocks are sorted into quintiles based on option based measures of stock mispricing. IPD is the average percentage difference between implied and actual stock prices. Implied borrowing fees is the stock borrowing fee implied by violations of put-call parity. Skewness is the difference in implied volatilities of a put with a delta of 0.2 and a call with a delta of 0.5. O/S is the natural logarithm of the ratio of stock and options volume. The FFC four factor model is used to calculate portfolio α. Newey and West [1987] adjusted standard errors with three lags are used. Returns and portfolio α are in percentages.

Table B.1.9: : Correlations

	IPD	Imp. Fee.	CW	$\Delta CVOL$	$\Delta PVOL$	Skewness	PP	OOI	O/S
IPD	1								
Imp. Fee.	-0.232	1							
CW	0.216	−0.5	1						
$\Delta CVOL$	0.031	−0.046	0.052	1					
$\Delta PVOL$	0.031	−0.005	0.019	0.57	1				
Skew	-0.097	0.175	−0.205	−0.037	−0.034	1			
PP	-0.063	0.067	−0.042	0.013	0.015	0.172	1		
OOI	0.054	−0.081	0.11	0.061	0.047	−0.084	−0.155	1	
O/S	-0.057	0.02	−0.091	0.012	0.015	−0.067	−0.317	−0.04	1

Notes: Correlations of monthly quintile portfolio assignments for various measures of option-based stock mispricing. IPD is the average percentage difference between implied and actual stock prices. Implied borrowing fees is the stock borrowing fee implied by violations of put-call parity. CW is the difference between implied volatilities of calls and puts. Skewness is the difference in implied volatilities of a put with a delta of 0.2 and a call with a delta of 0.5. $\Delta PVOL$ is the monthly change in implied volatilities for 30 day puts with deltas of -0.5. $\Delta CVOL$ is the change analogous change in call implied volatilities. PP is the ratio of put buy volume that opens positions to the sum of put and call buy volume. OOI is the difference between synthetic positive and negative options volume. O/S is the natural logarithm of the ratio of stock and options volume.

Table B.1.10: : Sharpe Ratios over Time

	IPD	Skewness	O/S
T+1	0.1804	0.1858	0.1591
T+2	0.2114	0.1891	0.1625
T+3	0.2245	0.1943	0.1875
T+4	0.2096	0.1948	0.2184
T+5	0.2022	0.1843	0.224
T+6	0.2575	0.1805	0.1981
T+7	0.1708	0.1896	0.1838
T+8	0.1378	0.1932	0.1742
T+9	0.1547	0.1742	0.1952
T+10	0.1721	0.1632	0.1876
T+11	0.1718	0.1711	0.1433
T+12	0.2132	0.167	0.1196

Notes: Sharpe ratios of value-weighted quintile portfolios of underpriced stocks formed using IPD, O/S, and skewness. Each month over 2004-2013, stocks are sorted into five quintiles for each of the mis-pricing measures: IPD, O/S, and skewness. For the high IPD, low O/S, and low skewness quintile portfolios, Sharpe ratios are calculated for each of the following 12 months. The excess return for the 1st month after portfolio formation is computed by subtracting the risk-free rate from the portfolio return. The time-series average monthly excess return for the first month after portfolio formation is divided by the time series standard deviation of first month excess returns to calculate the Sharpe ratio for the first month. Sharpe ratios for months 2-12 are calculated analogously. For comparison, the mean monthly Sharpe ratio is 0.1219 for the S&P 500 and 0.1236 for the CRSP value-weighted index over the same period.

Table B.1.11: : Value-Weighted Double Sorted Portfolio Alphas

	O/S					
	Low	2	3	4	High	H-L
Low IPD	0.1527	−0.1766	−0.0408	−0.7992	−0.9631	−1.1158
	(1.07)	(−0.8)	(−0.19)	(−2.16)	(−2.43)	(−2.69)
2	0.3163	0.0535	0.2865	−0.2075	−0.5291	−0.8453
	(3.04)	(0.31)	(1.68)	(−1.14)	(−1.97)	(−2.6)
3	0.3095	0.0712	−0.125	0.1403	−0.1861	−0.4957
	(2.87)	(0.57)	(−0.82)	(0.98)	(−1.42)	(−2.62)
4	0.0554	0.133	0.2802	0.2251	−0.0715	−0.127
	(0.43)	(0.82)	(2.13)	(1.87)	(−0.53)	(−0.71)
High IPD	0.2207	0.4794	0.4487	0.2202	0.3261	0.1054
	(1.56)	(3.41)	(2.75)	(1.09)	(1.49)	(0.38)
High - Low	0.068	0.656	0.4895	1.0194	1.2892	
	(0.36)	(3.7)	(1.77)	(2.35)	(2.93)	
	Skewness					
	Low	2	3	4	High	H-L
Low IPD	0.2179	0.1765	−0.4943	−1.1891	−0.3788	−0.5967
	(0.78)	(0.76)	(−2.23)	(−3.99)	(−1.34)	(−1.49)
2	−0.0288	0.0159	−0.0762	0.0229	−0.348	−0.3192
	(−0.18)	(0.09)	(−0.36)	(0.11)	(−1.29)	(−1.08)
3	0.1307	0.2233	−0.0739	0.1176	−0.2561	−0.3868
	(0.86)	(1.69)	(0.54)	(0.67)	(−1.47)	(−1.5)
4	0.2534	0.0329	−0.031	0.1503	0.2119	−0.0415
	(1.42)	(0.25)	(−0.21)	(0.82)	(1)	(−0.14)
High IPD	0.149	0.1495	0.5096	0.268	0.382	0.233
	(0.57)	(0.92)	(3.09)	(1.62)	(1.9)	(0.74)
High - Low	−0.0689	−0.027	1.0039	1.457	0.7608	
	(−0.18)	(−0.09)	(3.68)	(4.07)	(2.17)	
	Skewness					
	Low	2	3	4	High	H-L
Low Imp. Fee	0.0039	0.293	0.3379	0.3641	−0.1953	−0.1991
	(0.02)	(1.96)	(2.05)	(1.62)	(−0.75)	(−0.61)
2	0.2416	0.0946	0.3048	0.0671	0.3675	0.1259
	(1.51)	(0.69)	(2.54)	(0.33)	(1.81)	(0.42)
3	0.2489	0.0224	−0.3119	−0.0617	−0.2961	−0.5451
	(1.71)	(0.16)	(−2.06)	(−0.36)	(−1.24)	(−1.76)
4	−0.0931	−0.0134	0.128	−0.2074	−0.1196	−0.0265
	(−0.48)	(−0.08)	(0.69)	(−1.12)	(−0.37)	(−0.09)
High Imp. Fee	0.1257	−0.3088	−0.6165	−0.5463	−0.658	−0.7837
	(0.52)	(−1.47)	(−3.12)	(−1.93)	(−2.01)	(−2.22)
High - Low	0.1218	−0.6018	−0.954	−0.9104	−0.4628	

(0.37)	(−2.14)	(−3.53)	(−2.66)	(−1.43)

Notes: In Panel A, portfolios formed by double sorts on the log of the ratio of option to stock volume (O/S), and on the percentage difference between the actual stock price and the price implied by options (IPD). In Panel B, portfolios formed by double sorts on skewness, measured as the difference between the implied volatility of a short-term out-of-the-money put and the implied volatility of an at-the-money call, and the percentage difference between the actual stock price and the price implied by options (IPD). In Panel C, portfolios formed by double sorts on skewness, measured as the difference between the implied volatility of a short-term out-of-the-money put and the implied volatility of an at-the-money call, and implied stock borrowing fees estimated from put-call parity violations.

Table B.1.12: : Percentage of stocks in portfolio that are hard-to-borrow

	O/S				
	Low	2	3	4	High
$LowIPD$	0.123	0.195	0.28	0.415	0.653
2	0.045	0.065	0.089	0.134	0.204
3	0.035	0.039	0.052	0.054	0.075
4	0.036	0.039	0.045	0.055	0.066
$HighIPD$	0.064	0.08	0.09	0.106	0.149
	Skewness				
	Low	2	3	4	High
$LowIPD$	0.309	0.258	0.297	0.38	0.417
2	0.121	0.078	0.095	0.117	0.138
3	0.065	0.036	0.036	0.049	0.073
4	0.062	0.036	0.035	0.042	0.07
$HighIPD$	0.193	0.077	0.066	0.082	0.1
	Skewness				
	Low	2	3	4	High
$LowFees$	0.088	0.046	0.04	0.05	0.054
2	0.039	0.025	0.027	0.031	0.035
3	0.039	0.028	0.028	0.034	0.038
4	0.068	0.055	0.054	0.056	0.06
$HighFees$	0.265	0.284	0.329	0.343	0.31

Notes: Percentage of stocks in portfolio that are hard-to-borrow. Hard-to-borrow is defined as having borrowing fees that are among the highest 20% across all stocks. Panel A. Portfolios formed by double sorts on the log of the ratio of option to stock volume (O/S), and on the percentage difference between the actual stock price and the price implied by options (IPD). Panel B. Portfolios formed by double sorts on skewness, measured as the difference between the implied volatility of a short-term out-of-the-money put and the implied volatility of an at-the-money call, and the percentage difference between the actual stock price and the price implied by options (IPD). Panel C. Portfolios formed by double sorts on skewness, measured as the difference between the implied volatility of a short-term out-of-the-money put and the implied volatility of an at-the-money call, and implied stock borrowing fees estimated from put-call parity violations.

Table B.1.13 reports returns and four-factor Fama-French-Carhart alphas for equal-weighted quintile portfolios formed using nine option-based measures of stock mispricing. Panel A reports returns and alphas of portfolios formed on IPD and implied lending fees, the two measures that are based on differences between actual and implied stock prices. Both of these measures successfully predict stock returns in each of the three months following portfolio formation. A long-short strategy of buying the quintile of stocks with the highest IPD, that is the greatest differences between implied and actual stock prices, and selling the quintile of stocks with the lowest IPD, produces four-factor alphas of 77 basis points, 73 basis points, and 43 basis points in the three months after portfolio formation. Similarly, the long-short strategy of buying the quintile of stocks with low implied lending fees and selling the portfolio with high implied lending fees produces four-factor alphas of 77 basis points, 58 basis points and 53 basis points for the three months following portfolio formation. These alphas, like the alphas from the IPD long-short strategy are all statistically significant at any conventional level. It is significant that investors could wait one or two months after the calculation of IPD or implied fees and still earn abnormal returns. This suggests that the returns are not an artifact of microstructure noise. They could not, for example, be a result of stale trade prices or bid-ask bounce.

The ability of these measures to predict stock returns appears to come almost entirely from the short-side. In Panel A, the portfolio with the lowest IPD, that is the lowest implied prices relative to actual prices, has an alpha of -62 basis points with a t-statistic of -4.86 for month t+1. In contrast, the portfolio with the highest IPD, or greatest implied price relative to actual prices, has an alpha of 16 basis points with a t-statistic of 1.53. Similarly, the portfolio of stocks with high implied lending fees earns abnormal returns of -65 basis points with a t-statistic of -6.65 for the month after portfolio formation. The quintile of stocks with low implied fees earns abnormal returns of 12 basis points with a t-statistic of 2.23. Results for months t+2 and t+3 are similar for portfolios based on both IPD and implied lending fees. In both months t+2 and t+3, significant negative abnormal returns are earned in the bearish portfolio, but none of the other portfolios have significantly positive alphas.

These results suggest that for equal-weighted portfolios, which consist mainly of small stocks, the costs of short-selling and short-sale restrictions are behind the ability of option-based measures to predict stock returns. In each case, negative alphas, which provide profit opportunities for short-sellers, are larger and more significant than positive alphas. These portfolios are equal-weighted, so small stocks, which may be difficult to short, make up a significant part of the portfolios. It seems likely that the high and low portfolios in particular may be heavily weighted with small stocks.

Panels C to F reports results for portfolios formed using option based measures of mispricing derived from implied volatilities. Sorts based on Skewness, $\Delta PVOL$, and CW all produce portfolios that provide significant abnormal returns from long-short strategies in one or more months following portfolio formation. The abnormal returns produced by skewness, $\Delta CVOL$, and $\Delta PVOL$ are small relative to those produced by measures based on differences in implied and actual stock prices. Long-short alphas from portfolios based on CW are of similar magnitude to the alphas portfolios produced by IPD and implied lending fees. The other implied volatility measures generate smaller long-short portfolio alphas.

The abnormal returns earned by portfolios sorted on these measures also come from the short side. For example, the low CW portfolio, that is the portfolio where implied volatilities of calls are low relative to the implied volatilities of puts, earns abnormal returns of -61 basis points, -53 basis points and -42 basis points in the three months following portfolio formation. Each of these is significantly less than zero at the 1% level. In contrast, the abnormal returns of the high CW portfolio are less than 14 basis points in each of the three succeeding months and never significant. Similar results are obtained from portfolios based on skewness and $\Delta PVOL$.

Panels G to I provides results for portfolios formed using options order imbalance (OOI), Pan Poteshman (PP) and the log ratio of option to stock volume (O/S). Each of these measures is based on trading volume. The results for OOI and PP are weak. A long-short strategy produces abnormal returns for OOI but only for the first month following portfolio formation. The long-short strategy fails to produce significant abnormal returns at all for PP. This is not entirely surprising though as the PP measure was originally used to predict returns over shorter intervals. For O/S, a long-short strategy produces significant abnormal returns in each month. The returns again come primarily from the short-side. Again, for all these measures, short-sale restrictions appear to be the source of returns to option based strategies.

We also calculate average returns and alphas for the equal-weighted quintile portfolios for the two five year periods 2003-2008 and 2009-2013. Results (not shown) are very similar for the two subperiods. In both periods, most of these option-based measures are able to find overpriced stocks that underperform by statistically significant amounts. Quintile portfolios of underpriced stocks may have positive alphas, but they are generally insignificant in both periods. SEC Rule 10b-21, which cracked down on naked shorting and failures to deliver, became effective in mid-October 2008, very close to the end of the first subperiod. Our finding that option-based measures of mispricing had similar predictive power in both subperiods suggests that Rule 10b-21 had little impact on bearish investors decisions whether to trade stock or options.

All of these measures of stock mispricing are taken from options, so it might seem that they contain the same information. We find, though, that portfolio sortings across measures have surprisingly low correlations (see Table C.1.2 in the text). Hence it seems possible that double-sorts on these measures may produce larger returns.

In Panel A of Table B.1.14 we first sort stocks into quintiles by IPD, and then sort each IPD quintile into quintiles by O/S. In Panel B, we sort first on IPD, and then on skewness. In Panel C, we first sort stocks by implied borrowing fees and then by skewness. These double sorts do produce portfolios of stocks with very large negative abnormal returns. In Panel A, the low IPD/High O/S portfolio earn an average alpha of -1.0192 in the first month after portfolio formation. The low IPD/high skewness portfolio in Panel B earns an abnormal return of -1.3620% for month t+1. Finally, the high implied fee/high skewness portfolio in Panel C earns an alpha of -0.9038 in the month following portfolio formation. As Table B.1.12 in the text shows, however, the portfolios that appear to generate large returns from shortselling are very heavily weighted with hard-to-borrow stocks. In contrast, none of the double sorts yields a portfolio with especially large positive alphas.

Table B.1.13: : Equally-Weighted Portfolio Sorts

Panel A: IPD									
Port.	r_{t+1}	α_{t+1}	$t_{\alpha_{t+1}}$	r_{t+2}	α_{t+2}	$t_{\alpha_{t+2}}$	r_{t+3}	α_{t+3}	$t_{\alpha_{t+3}}$
Low	0.2639	−0.6168	−4.86	0.3078	−0.6199	−4.74	0.532	−0.3921	−2.75
2	0.7817	0.0174	0.24	0.7619	−0.0459	−0.69	0.7586	−0.0637	−0.92
3	0.7289	0.0032	0.04	0.7647	−0.0049	−0.06	0.722	−0.0391	−0.44
4	0.7757	0.0374	0.4	0.87	0.0897	0.89	0.8493	0.074	0.78
High	1.0051	0.1566	1.53	0.9691	0.1068	1.02	0.8967	0.0356	0.39
H-L	0.7412	0.7734	5.91	0.6613	0.7268	5.16	0.3647	0.4277	3.05

Panel B: Implied Fees									
Port.	r_{t+1}	α_{t+1}	$t_{\alpha_{t+1}}$	r_{t+2}	α_{t+2}	$t_{\alpha_{t+2}}$	r_{t+3}	α_{t+3}	$t_{\alpha_{t+3}}$
Low	0.8509	0.1223	2.23	0.8169	0.0441	0.72	0.8381	0.0702	0.94
2	0.9148	0.2006	2.7	0.8479	0.0962	1.15	0.8708	0.126	1.63
3	0.8063	0.0773	1.07	0.8597	0.0912	1.07	0.8212	0.0642	0.79
4	0.7201	−0.0453	−0.55	0.8442	0.0283	0.37	0.8789	0.0638	0.67
High	0.1994	−0.6524	−6.65	0.354	−0.5336	−5.58	0.4271	−0.4609	−3.95
H-L	-0.6515	−0.7747	−7.68	−0.4629	−0.5777	−5.47	−0.4109	−0.531	−4.11

Panel C: CW									
Port.	r_{t+1}	α_{t+1}	$t_{\alpha_{t+1}}$	r_{t+2}	α_{t+2}	$t_{\alpha_{t+2}}$	r_{t+3}	α_{t+3}	$t_{\alpha_{t+3}}$
Low	0.2523	−0.6053	−4.42	0.3814	−0.5282	−4.42	0.4894	−0.4203	−3.28
2	0.7937	0.0274	0.39	0.7317	−0.0683	−0.89	0.7755	−0.0203	−0.26
3	0.815	0.0898	1.3	0.9612	0.2084	2.95	0.8913	0.129	1.84
4	0.9219	0.1815	3.06	0.8957	0.1158	1.67	0.8884	0.111	1.49

High	0.9992	0.112	0.9	1.0526	0.1359	0.89	1.0234	0.1322	0.95
H-L	0.7469	0.7173	5.48	0.6712	0.6641	4.24	0.534	0.5526	3.78

Panel D: Skewness									
Port.	r_{t+1}	α_{t+1}	$t_{\alpha_{t+1}}$	r_{t+2}	α_{t+2}	$t_{\alpha_{t+2}}$	r_{t+3}	α_{t+3}	$t_{\alpha_{t+3}}$
Low	0.9702	0.0926	0.64	0.9637	0.0566	0.39	0.9982	0.1204	0.88
2	0.879	0.1596	1.75	0.8551	0.0845	1.03	0.8678	0.0893	1.05
3	0.7277	−0.0388	−0.48	0.822	0.0187	0.2	0.756	−0.0404	−0.49
4	0.6271	−0.1814	−2.53	0.757	−0.0988	−1.57	0.7907	−0.0394	−0.54
High	0.5275	−0.2549	−2.69	0.5454	−0.266	−2.72	0.5579	−0.2819	−2.65
H-L	-0.4427	−0.3475	−1.89	−0.4184	−0.3226	−1.85	−0.4403	−0.4026	−2.31

Panel E: $\Delta CVOL$									
Port.	r_{t+1}	α_{t+1}	$t_{\alpha_{t+1}}$	r_{t+2}	α_{t+2}	$t_{\alpha_{t+2}}$	r_{t+3}	α_{t+3}	$t_{\alpha_{t+3}}$
Low	0.5308	−0.2933	−2.39	0.7063	−0.1989	−1.62	0.848	−0.0466	−0.46
2	0.8168	0.0847	1.25	0.8551	0.0573	0.75	0.8266	0.0129	0.14
3	0.8072	0.0758	0.92	0.8975	0.1177	1.44	0.9306	0.1266	1.47
4	0.8163	0.0475	0.68	0.8421	0.0525	0.63	0.845	0.0119	0.14
High	0.7767	−0.1264	−1.18	0.731	−0.1736	−1.81	0.7502	−0.1617	−1.33
H-L	0.2459	0.1669	1.22	0.0246	0.0253	0.18	−0.0978	−0.1151	−0.78

Panel F: $\Delta PVOL$									
Port.	r_{t+1}	α_{t+1}	$t_{\alpha_{t+1}}$	r_{t+2}	α_{t+2}	$t_{\alpha_{t+2}}$	r_{t+3}	α_{t+3}	$t_{\alpha_{t+3}}$
Low	0.8012	−0.0163	−0.13	0.6605	−0.2404	−2.07	0.8703	−0.0372	−0.41
2	0.7859	0.0481	0.66	0.9083	0.1009	1.29	0.9015	0.0942	1.1
3	0.8253	0.1066	1.24	0.935	0.1603	1.87	0.9102	0.1073	1.33
4	0.7812	0.0032	0.05	0.8459	0.0418	0.5	0.7959	−0.033	−0.41
High	0.5544	−0.353	−3.43	0.6823	−0.2077	−2.23	0.7224	−0.1886	−1.56
H-L	-0.2469	−0.3366	−2.55	0.0218	0.0327	0.25	−0.1479	−0.1514	−1.19

Panel G: PP									
Port.	r_{t+1}	α_{t+1}	$t_{\alpha_{t+1}}$	r_{t+2}	α_{t+2}	$t_{\alpha_{t+2}}$	r_{t+3}	α_{t+3}	$t_{\alpha_{t+3}}$
Low	0.3806	−0.4407	−4.88	0.5994	−0.2514	−2.67	0.7145	−0.1728	−2.21
2	0.8083	0.0661	0.93	0.8439	0.0644	1.01	0.8324	0.0205	0.2
3	1.0514	0.2439	2.12	1.0287	0.1976	1.92	0.918	0.0542	0.69
4	0.9208	0.1215	1.24	1.0303	0.1742	1.91	1.0336	0.1419	1.28
High	0.9194	0.0189	0.17	0.7201	−0.2181	−1.81	0.8298	−0.1157	−1.01
H-L	0.5387	0.4596	3.83	0.1206	0.0333	0.3	0.1153	0.0571	0.48

Panel H: OOI									
Port.	r_{t+1}	α_{t+1}	$t_{\alpha_{t+1}}$	r_{t+2}	α_{t+2}	$t_{\alpha_{t+2}}$	r_{t+3}	α_{t+3}	$t_{\alpha_{t+3}}$
Low	0.6282	−0.2846	−1.69	0.5976	−0.3576	−2.44	0.6183	−0.333	−2.33
2	0.7396	−0.045	−0.038	0.7539	−0.0819	−0.67	0.8445	−0.0253	−0.25
3	0.767	0.0021	0.03	0.803	−0.0274	−0.27	0.8138	−0.0296	−0.34
4	0.6597	−0.1063	−1.06	0.7392	−0.0528	−0.53	0.8005	−0.0499	−0.56
High	0.6852	−0.0936	−1	0.71	−0.0921	−0.89	0.7971	−0.0346	−0.41
H-L	0.0569	0.1911	0.89	0.1124	0.2655	1.26	0.1788	0.2985	1.85

Panel I: O/S									
Port.	r_{t+1}	α_{t+1}	$t_{\alpha_{t+1}}$	r_{t+2}	α_{t+2}	$t_{\alpha_{t+2}}$	r_{t+3}	α_{t+3}	$t_{\alpha_{t+3}}$
Low	0.8785	0.1142	1.55	0.9473	0.1476	2.02	0.9143	0.1145	1.86
2	0.8716	0.0706	0.92	0.8259	−0.0206	−0.31	0.8316	−0.011	−0.14
3	0.7821	−0.0026	−0.03	0.8961	0.0688	0.85	0.834	0.0191	0.31
4	0.6911	−0.1187	−1.31	0.6799	−0.1465	−1.63	0.7673	−0.0485	−0.52
High	0.4044	−0.3728	−4.25	0.4706	−0.3627	−3.95	0.4797	−0.3615	−3.64
H-L	-0.4741	−0.487	−4.18	−0.4768	−0.5103	−4.15	−0.4346	−0.476	−4.53

Notes: Each month over 2004-2013, stocks are sorted into quintiles based on option based measures of stock mispricing. IPD is the average percentage difference between implied and actual stock prices. Implied borrowing fees is the stock borrowing fee implied by violations of put-call parity. CW is the difference between implied volatilities of calls and puts. Skewness is the difference in implied volatilities of a put with a delta of 0.2 and a call with a delta of 0.5. $\Delta PVOL$ is the monthly change in implied volatilities for 30 day puts with deltas of -0.5. $\Delta CVOL$ is the change analogous change in call implied volatilities. PP is the ratio of put buy volume that opens positions to the sum of put and call buy volume. OOI is the difference between synthetic positive and negative options volume. O/S is the natural logarithm of the ratio of stock and options volume. The FFC four factor model is used to calculate portfolio α. Newey and West [1987] adjusted standard errors with three lags are used. Returns and portfolio α are in percentages.

Table B.1.14: : Equal-Weighted Double Sorted Portfolio Alphas

	O/S					
	Low	2	3	4	High	H-L
Low IPD	−0.0032	−0.3225	−0.6042	−0.8547	−1.0192	−1.0159
	(−0.02)	(−1.66)	(−3.45)	(−4.29)	(−4.31)	(−3.56)
2	0.3213	0.0379	0.1625	−0.06	−0.3714	−0.6928
	(2.06)	(0.28)	(1.14)	(−0.45)	(−2.18)	(−2.81)
3	0.2734	0.0392	−0.0615	−0.0246	−0.2433	−0.5166
	(2.41)	(0.32)	(−0.45)	(−0.17)	(−1.72)	(−2.77)
4	−0.0235	−0.0183	0.1824	0.0744	−0.2594	−0.2359
	(−0.21)	(−0.13)	(1.22)	(0.65)	(−1.59)	(−1.47)
High IPD	−0.1633	0.0669	0.2611	0.157	0.0865	0.2498
	(−0.99)	(0.49)	(1.57)	(0.9)	(0.59)	(1.35)
High - Low	−0.16	0.3894	0.8653	1.0117	1.1057	
	(−0.67)	(1.83)	(3.53)	(3.75)	(3.89)	

	Skewness					
	Low	2	3	4	High	H-L
Low IPD	0.045	−0.0109	−0.4105	−1.0919	−1.3169	−1.362
	(0.14)	(−0.06)	(−2.2)	(−5.57)	(−5.08)	(−3.48)
2	0.1229	0.1622	0.0464	−0.0995	−0.0311	−0.154
	(0.67)	(1.33)	(−0.31)	(−0.72)	(−0.18)	(−0.59)
3	−0.0077	0.0483	−0.018	0.1518	−0.2061	−0.1984
	(−0.05)	(0.38)	(−0.14)	(0.96)	(−1.65)	(−0.98)
4	0.1488	−0.0195	−0.0537	−0.026	−0.0573	−0.2061
	(0.95)	(−0.12)	(−0.38)	(−0.18)	(−0.4)	(−0.98)
High IPD	−0.1243	0.3437	0.108	−0.1192	0.2228	0.347
	(−0.53)	(2.03)	(0.84)	(−0.82)	(1.32)	(1.48)
High - Low	−0.1693	0.3546	0.5185	0.9727	1.5397	
	(−0.61)	(1.66)	(2.24)	(4.15)	(4.52)	

	Skewness					
	Low	2	3	4	High	H-L
Low Imp. Fee	−0.1743	0.163	0.1755	0.2693	0.1886	0.3629
	(−1.17)	(1.52)	(1.59)	(2.42)	(1.44)	(1.69)
2	0.1946	0.1531	0.2036	0.2062	0.2276	0.0329
	(1.37)	(1.25)	(2)	(1.49)	(1.59)	(0.16)
3	0.2358	0.0892	−0.074	0.2263	−0.112	−0.3478
	(1.65)	(0.69)	(−0.6)	(1.59)	(−0.85)	(−1.72)
4	0.0007	−0.0493	−0.0765	−0.1348	0.0072	0.0065
	(0)	(−0.34)	(−0.55)	(−1.29)	(0.06)	(0.03)
High Imp. Fee	−0.2159	−0.4013	−0.9707	−0.7615	−0.9038	−0.6879
	(−0.99)	(−2.52)	(−6.73)	(−4.4)	(−4.82)	(−2.45)
High - Low	−0.0416	−0.5643	−1.1462	−1.0308	−1.0924	
	(−0.16)	(−3.23)	(−6.99)	(−5.13)	(5.82)	

Notes: Panel A. Portfolios formed from double sorts of stocks on IPD and O/S. All stocks, equal-weighted portfolios. Panel B. Portfolios formed from double sorts of stocks on IPD and Skewness. All stocks, equal-weighted portfolios. Panel C. Portfolios formed from double sorts of stocks on implied borrowing fees and skewness. All stocks, equal-weighted portfolios.

Appendix C

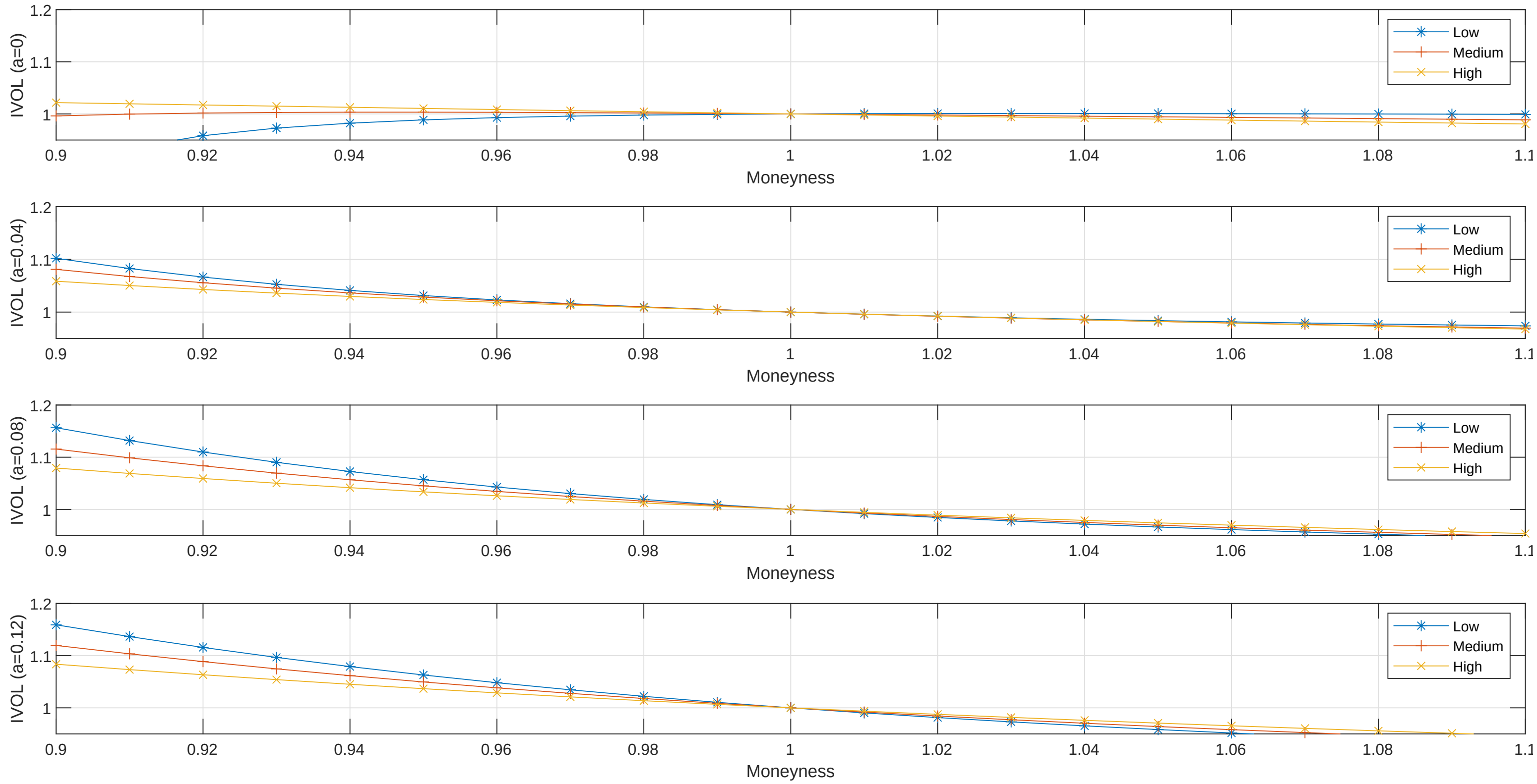
Appendix for Chapter 4

This appendix contains additional proofs from Chapter 4.

C.1 Mathematical Proofs, Figures, and Tables

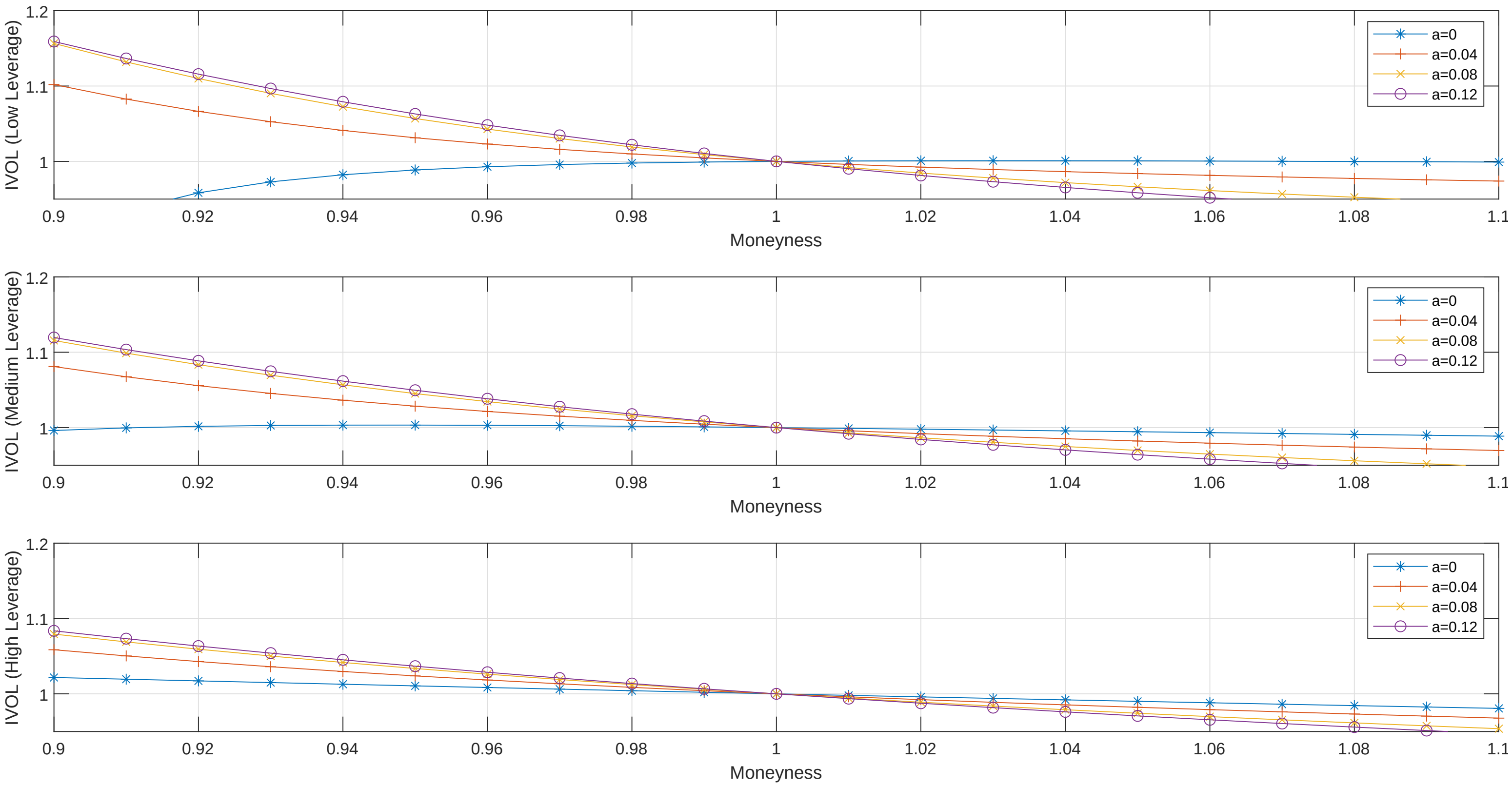
C.1.1 Figures

Figure C.1.1: : Impact of Leverage Ratio on Implied Volatility Smile For Different Accounting Transparency



Volatility smile for low (blue), medium (red), or high (yellow) leverage ratios. The panel closest to the top of the page is where the accounting precision is set to zero, (observed asset values). Normalized volatility smile as a function of moneyness.

Figure C.1.2: : Impact of Accounting Transparency on Implied Volatility Smile For Different Leverage Ratios



Volatility smile for $a = 0.00$ (blue), $a = 0.04$ (red), $a = 0.08$ (yellow), and $a = 0.12$ (purple) accounting noise. Each of Panels represent low, medium, and high leverage ratio.

Figure C.1.3: : Change in Implied Volatility Skewness with respect to Leverage for different Accounting Transparency

Change in Implied Volatility Skewness with respect to Leverage for three different levels of Accounting Transparency of $a = 0.04$, $a = 0.08$, and $a = 0.12$ for three different levels of moneyness

C.1.2 Tables

Table C.1.1: : Summary Statistics

			Percentiles			
Variable Name	N	Mean	Std. Dev.	25th	Median	75th
Slope Skew	140611	−0.58	0.49	−0.72	−0.46	−0.28
Slope Skew v2	141201	−1.01	0.83	−1.41	−0.89	−0.46
Disp	181540	0.16	0.41	0.02	0.04	0.11
Nanalyst	290063	5.54	6.79	0.00	3.00	8.00
Leverage	319606	0.20	0.22	0.01	0.12	0.32
Momentum	307111	0.06	0.43	−0.18	0.02	0.22
M/B	319606	3.69	5.68	1.22	2.06	3.77
Beta	272691	1.52	1.69	0.49	1.26	2.28
Size	311655	19.48	2.17	17.9	19.42	20.94
Idio Skew	312269	0.4	1.32	−0.17	0.33	0.92
Atm Ivol	133414	0.49	0.24	0.32	0.43	0.61

Notes: The table presents quarterly summary statistics for main variables. We compute the dispersion in analyst forecast (Disp) and number of analyst covering stock (Nanalyst) from I/B/E/S quarterly earnings data. M/B is the ratio of market and book equity computed using quarterly data from COMPUSTAT. Leverage is the book value of debt divided by sum of debt and market value of equity using quarterly data from COMPUSTAT. M/B and Leverage are lagged by one quarter in order to account for the timing of the release of accounting statements. Size is the log of the product of the stock price and shares outstanding (times 1000) from CRSP monthly stock files (of firms with share codes 10 and 11 common shares). Momentum is the past 6 month cumulative monthly stock returns from CRSP. Beta is the stock beta with the market estimated from 36 months rolling regressions adjusted by 3 months of lags for asynchronous trading as per the Dimson [1979] adjustment. Idio Skew is the idiosyncratic skewness of daily returns estimated quarterly using daily CRSP stock returns. Atm Ivol is average of call and put contract implied volatility with $|\Delta| = 0.5$ and 30 days to maturity, and using OptionMetrics Volatility surface. Slope Skew (and v2) and Atm Ivol are computed daily on a firm level and then averaged over the quarter. The sample period is quarterly observations from January 1997 to December 2017.

Table C.1.2: : Correlations

Variable Names	Correlations										
	Slope Skew	Slope Skew v2	Disp	Nanalyst	Leverage	Momentum	M/B	Beta	Size	Idio Skew	Atm Ivol
Slope Skew	1.00	0.63	-0.08	0.19	-0.13	0.11	0.10	0.01	0.21	$-3.7e-3$	0.03
Slope Skew v2	0.63	1.00	0.08	-0.10	-0.16	0.05	0.10	0.19	-0.29	0.02	0.51
Disp	-0.08	0.08	1.00	-0.09	0.07	-0.05	-0.02	0.09	-0.18	0.02	0.20
Nanalyst	0.19	-0.10	-0.09	1.00	-0.05	0.01	0.09	-0.10	0.76	-0.11	-0.30
Leverage	-0.13	-0.16	0.07	-0.05	1.00	-0.05	-0.18	-0.06	-0.08	0.01	-0.07
Momentum	0.11	0.05	-0.05	0.01	-0.05	1.00	0.08	0.01	0.16	0.05	-0.13
M/B	0.1	0.10	-0.02	0.09	-0.18	0.08	1.00	0.08	0.12	$2e-3$	0.09
Beta	0.01	0.19	0.09	-0.1	-0.06	0.01	0.08	1.00	-0.16	0.05	0.30
Size	0.21	-0.29	-0.18	0.76	-0.08	0.16	0.12	-0.16	1.00	-0.12	-0.59
Idio Skew	$-3.7e-3$	0.02	0.02	-0.11	0.01	0.05	$2e-3$	0.05	-0.12	1.00	0.03
Atm Ivol	0.03	0.51	0.20	-0.30	-0.07	-0.13	0.09	0.30	-0.59	0.03	1.00

Notes: Table contains pooled correlations between all control and accounting quality measures from Table C.1.1. The sample period is quarterly observations from January 1997 to December 2017.

Table C.1.3: : FM Regression with Nanalyst as Accounting Transparency (Options with $15 \leq$ d2mat ≤ 45)

	Less Accounting Transparency						
Variable	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Leverage	-0.22	-0.21	-0.19	-0.28	-0.27	-0.27	-0.21
	(-6.09)	(-5.98)	(-6.13)	(-9.79)	(-9.58)	(-9.74)	(-8.16)
Momentum		0.20	0.21	0.13	0.13	0.13	0.13
		(8.83)	(8.75)	(5.78)	(5.55)	(5.89)	(4.63)
Beta			$-2.2e-3$	0.02	0.02	0.02	0.01
			(-0.44)	(3.56)	(3.6)	(3.86)	(2.07)
Size				0.14	0.14	0.14	0.15
				(11.54)	(11.47)	(11.41)	(13.88)
M/B					$3.1e-3$	$3.1e-3$	$1.8e-3$
					(3.72)	(3.74)	(1.69)
Idio Skew						-0.01	-0.01
						(-3.96)	(-3.42)
Atm Ivol							0.17
							(2.92)
R^2	1.03	2.87	3.20	8.88	8.92	9.11	9.59
N obs	362.00	362.00	341.00	341.00	341.00	341.00	294.00

	More Accounting Transparency						
Variable	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Leverage	-0.34	-0.33	-0.32	-0.33	-0.32	-0.32	-0.29
	(-17.36)	(-17.51)	(-16.77)	(-17.5)	(-18.21)	(-18.28)	(-17.37)
Momentum		0.18	0.20	0.15	0.14	0.14	0.14
		(10.36)	(10.95)	(8.72)	(8.43)	(8.39)	(9.09)
Beta			$3.1e-3$	0.02	0.02	0.02	0.01
			(0.75)	(3.4)	(3.37)	(3.41)	(3.48)
Size				0.05	0.05	0.05	0.06
				(8.75)	(8.8)	(8.79)	(7.27)
M/B					$1.9e-3$	$1.9e-3$	$1e-3$
					(2.89)	(2.88)	(2.03)
Idio Skew						$-1.8e-3$	$-1.9e-3$
						(-1.37)	(-1.25)
Atm Ivol							0.25
							(3.73)
R^2	3.00	4.75	5.04	9.39	9.51	9.55	10.32
N obs	1236.00	1236.00	1175.00	1175.00	1175.00	1175.00	991.00

Notes: Quarterly regressions are estimated using Fama and Macbeth [1973] from January 1997 to December 2017 on all firms with listed equity options. Panel A (B), each of the 7 FM regression models is estimated using the subset of data each quarter that is below (above) 50th percentile level of Accounting Transparency. The variable (and sign) of interest is Leverage in regression (1). Control variables are added individually from regressions (2) to (7): Momentum, Beta, Size, M/B, Idio Skew, and Atm Ivol. R-squared (R^2) and quarterly number of observations (N obs) are computed in each regression. t-statistics are computed using Newey and West [1987] standard errors with 3 lags and reported in parentheses below the coefficient. The dependent variable is the Slope Skew.

Table C.1.4: : FM Regression with Disp as Accounting Transparency (Options with $15 \leq$ d2mat ≤ 45)

	Less Accounting Transparency						
Variable	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Leverage	-0.23	-0.21	-0.19	-0.28	-0.27	-0.27	-0.24
	(-9.32)	(-8.83)	(-7.75)	(-14.18)	(-13.69)	(-13.69)	(-11.4)
Momentum		0.20	0.22	0.14	0.14	0.14	0.14
		(9.39)	(9.95)	(7.98)	(7.67)	(7.72)	(7.16)
Beta			-0.01	0.01	0.01	0.01	0.01
			(-1.75)	(2.16)	(2.04)	(2.09)	(2.98)
Size				0.11	0.11	0.11	0.11
				(10.76)	(10.79)	(10.76)	(9.52)
M/B					$2.5e-3$	$2.5e-3$	$2.9e-3$
					(4.62)	(4.56)	(4.47)
Idio Skew						$-3e-3$	$-3.7e-3$
						(-1.71)	(-1.8)
Atm Ivol							0.01
							(0.26)
R^2	1.66	3.94	4.41	14.14	14.18	14.21	14.62
N obs	668.00	668.00	628.00	628.00	628.00	628.00	520.00

	More Accounting Transparency						
Variable	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Leverage	-0.40	-0.39	-0.38	-0.40	-0.38	-0.38	-0.33
	(-14.27)	(-14.3)	(-13.92)	(-14.67)	(-16.41)	(-16.41)	(-16.03)
Momentum		0.15	0.16	0.14	0.13	0.14	0.12
		(9.95)	(9.74)	(9.5)	(9.52)	(9.46)	(8.32)
Beta			0.01	0.02	0.02	0.02	0.02
			(4.03)	(4.77)	(4.76)	(4.8)	(4.29)
Size				0.04	0.04	0.04	0.05
				(5.43)	(5.4)	(5.4)	(5.98)
M/B					$2.6e-3$	$2.6e-3$	$1.4e-3$
					(3.04)	(3.04)	(2.48)
Idio Skew						$-1.6e-3$	$-0.1e-3$
						(-0.99)	(-0.08)
Atm Ivol							0.33
							(5.43)
R^2	3.71	4.83	4.95	8.27	8.42	8.46	9.35
N obs	814.00	814.00	780.00	780.00	780.00	780.00	675.00

Notes: Quarterly regressions are estimated using Fama and Macbeth [1973] from January 1997 to December 2017 on all firms with listed equity options. Panel A (B), each of the 7 FM regression models is estimated using the subset of data each quarter that is below (above) 50th percentile level of Accounting Transparency. The variable (and sign) of interest is Leverage in regression (1). Control variables are added individually from regressions (2) to (7): Momentum, Beta, Size, M/B, Idio Skew, and Atm Ivol R-squared (R^2) and quarterly number of observations (N obs) are computed in each regression. t-statistics are computed using Newey and West [1987] standard errors with 3 lags and reported in parentheses below the coefficient. The dependent variable is the Slope Skew.

Table C.1.5: : FM Regression with Nanalyst as Accounting Transparency (Avg d2mat)

Less Accounting Transparency							
Variable	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Leverage	−0.10	−0.10	−0.09	−0.11	−0.10	−0.10	−0.08
	(−6.35)	(−6.38)	(−5.96)	(−7.16)	(−7.06)	(−7.15)	(−5.9)
Momentum		0.05	0.06	0.04	0.04	0.04	0.04
		(5.49)	(6.05)	(4.46)	(4.15)	(4.6)	(3.38)
Beta			$3.2e-3$	0.01	0.01	0.01	$3.1e-3$
			(1.81)	(4.09)	(4.05)	(4.31)	(1.44)
Size				0.03	0.03	0.03	0.04
				(7.91)	(7.9)	(7.95)	(7.29)
M/B					$1.7e-3$	$1.7e-3$	$1.2e-3$
					(3.82)	(3.79)	(2.61)
Idio Skew						−0.01	−0.01
						(−4.18)	(−4.26)
Atm Ivol							0.11
							(4.74)
R^2	0.77	1.77	1.94	3.51	3.68	4.07	5.31
N obs	380.00	380.00	358.00	358.00	358.00	358.00	308.00

More Accounting Transparency							
Variable	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Leverage	−0.13	−0.12	−0.11	−0.12	−0.11	−0.11	−0.1
	(−15.95)	(−17.46)	(−17.05)	(−17.8)	(−19.03)	(−19.14)	(−19.84)
Momentum		0.06	0.06	0.05	0.05	0.05	0.05
		(7.79)	(8.09)	(7.67)	(7.54)	(7.54)	(7.17)
Beta			$3.6e-3$	0.01	0.01	0.01	$3.3e-3$
			(2.41)	(4.03)	(4)	(4.04)	(2.53)
Size				0.01	0.01	0.01	0.02
				(6.57)	(6.69)	(6.7)	(5.64)
M/B					$0.8e-3$	$0.8e-3$	$0.3e-3$
					(2.89)	(2.89)	(1.58)
Idio Skew						$-1.2e-3$	$-1e-3$
						(−1.78)	(−1.43)
Atm Ivol							0.13
							(4.08)
R^2	2.1	3.21	3.47	5.24	5.37	5.42	6.74
N obs	1255.00	1255.00	1194.00	1194.00	1194.00	1193.00	1005.00

Notes: Quarterly regressions are estimated using Fama and Macbeth [1973] from January 1997 to December 2017 on all firms with listed equity options. Panel A (B), each of the 7 FM regression models is estimated using the subset of data each quarter that is below (above) 50th percentile level of Accounting Transparency. The variable (and sign) of interest is Leverage in regression (1). Control variables are added individually from regressions (2) to (7): Momentum, Beta, Size, M/B, Idio Skew, and Atm Ivol R-squared (R^2) and quarterly number of observations (N obs) are computed in each regression. t-statistics are computed using Newey and West [1987] standard errors with 3 lags and reported in parentheses below the coefficient. The dependent variable is the Slope Skew v3.

Table C.1.6: : FM Regression with Disp as Accounting Transparency (Avg d2mat)

	Less Accounting Transparency						
Variable	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Leverage	−0.10	−0.10	−0.09	−0.11	−0.11	−0.11	−0.10
	(−12.2)	(−13.48)	(−13.41)	(−19.46)	(−18.87)	(−18.91)	(−15.33)
Momentum		0.07	0.08	0.05	0.05	0.05	0.05
		(8.01)	(8.05)	(7.07)	(6.82)	(6.86)	(5.87)
Beta			$-0.3e-3$	$4.5e-3$	$4.3e-3$	$4.4e-3$	$3.7e-3$
			(−0.17)	(3.05)	(2.98)	(3)	(2.71)
Size				0.03	0.03	0.03	0.03
				(10)	(10.09)	(10.05)	(7.82)
M/B					$1.1e-3$	$1.1e-3$	$0.9e-3$
					(3.61)	(3.65)	(3.09)
Idio Skew						$-0.9e-3$	$-0.6e-3$
						(−1.06)	(−0.64)
Atm Ivol							0.05
							(2)
R^2	1.37	2.94	3.32	8.20	8.28	8.33	9.00
N obs	689.00	689.00	648.00	648.00	648.00	648.00	535.00

	More Accounting Transparency						
Variable	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Leverage	−0.13	−0.13	−0.12	−0.13	−0.12	−0.12	−0.10
	(−11.24)	(−11.44)	(−10.74)	(−10.97)	(−11.86)	(−11.88)	(−11.72)
Momentum		0.05	0.05	0.05	0.05	0.05	0.04
		(6.91)	(7.67)	(7.84)	(7.84)	(8.04)	(6.66)
Beta			0.01	0.01	0.01	0.01	$3.8e-3$
			(5.23)	(5.03)	(5.01)	(5.06)	(2.85)
Size				0.01	0.01	0.01	0.01
				(3.1)	(2.98)	(3)	(4.12)
M/B					$1.2e-3$	$1.1e-3$	$0.7e-3$
					(3.44)	(3.42)	(2.79)
Idio Skew						$-1.5e-3$	$-1.2e-3$
						(−2.23)	(−1.86)
Atm Ivol							0.15
							(5.2)
R^2	2.08	2.83	2.92	4.25	4.38	4.41	5.38
N obs	825.00	825.00	791.00	791.00	791.00	791.00	683.00

Notes: Quarterly regressions are estimated using Fama and Macbeth [1973] from January 1997 to December 2017 on all firms with listed equity options. Panel A (B), each of the 7 FM regression models is estimated using the subset of data each quarter that is below (above) 50th percentile level of Accounting Transparency. The variable (and sign) of interest is Leverage in regression (1). Control variables are added individually from regressions (2) to (7): Momentum, Beta, Size, M/B, Idio Skew, and Atm Ivol. R-squared (R^2) and quarterly number of observations (N obs) are computed in each regression. t-statistics are computed using Newey and West [1987] standard errors with 3 lags and reported in parentheses below the coefficient. The dependent variable is the Slope Skew.

Table C.1.7: : FM Regression with Nanalyst as Accounting Transparency (Options with $15 \leq$ d2mat ≤ 45)

	Less Accounting Transparency						
Variable	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Leverage	-0.53	-0.53	-0.5	-0.44	-0.41	-0.41	-0.28
	(-8.95)	(-9.03)	(-9)	(-7.8)	(-7.34)	(-7.39)	(-7.16)
Momentum		0.09	0.11	0.16	0.15	0.15	0.12
		(2.91)	(4.03)	(5.96)	(5.58)	(5.8)	(4.17)
Beta			0.08	0.07	0.06	0.06	0.02
			(11.73)	(11.24)	(10.95)	(10.85)	(3.14)
Size				-0.09	-0.09	-0.09	0.07
				(-6.82)	(-6.96)	(-7.02)	(4.71)
M/B					0.01	0.01	$-0.9e-3$
					(9.78)	(9.56)	(-0.63)
Idio Skew						-0.01	-0.01
						(-2.04)	(-2.95)
Atm Ivol							1.61
							(18.31)
R^2	2.53	3.98	6.11	8.23	8.67	8.96	17.74
N obs	366.00	366.00	345.00	345.00	345.00	345.00	296.00

	More Accounting Transparency						
Variable	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Leverage	-0.73	-0.73	-0.71	-0.66	-0.61	-0.62	-0.5
	(-11.67)	(-12.65)	(-11.18)	(-11.39)	(-11.65)	(-11.7)	(-13.44)
Momentum		0.01	0.02	0.13	0.12	0.12	0.19
		(0.18)	(0.31)	(3.25)	(2.95)	(2.85)	(6.56)
Beta			0.14	0.11	0.11	0.11	0.03
			(8.16)	(8.52)	(8.41)	(8.46)	(4.39)
Size				-0.13	-0.13	-0.13	-0.01
				(-10.35)	(-10.65)	(-10.64)	(-0.84)
M/B					0.01	0.01	$-1.4e-3$
					(4.92)	(4.91)	(-1.91)
Idio Skew						$2.8e-3$	$2.4e-3$
						(1.34)	(1.21)
Atm Ivol							2.33
							(11.61)
R^2	4.29	6.56	11.48	19.5	19.74	19.78	31.40
N obs	1238.00	1238.00	1177.00	1177.00	1177.00	1177.00	992.00

Notes: Quarterly regressions are estimated using Fama and Macbeth [1973] from January 1997 to December 2017 on all firms with listed equity options. Panel A (B), each of the 7 FM regression models is estimated using the subset of data each quarter that is below (above) 50th percentile level of Accounting Transparency. The variable (and sign) of interest is Leverage in regression (1). Control variables are added individually from regressions (2) to (7): Momentum, Beta, Size, M/B, Idio Skew, and Atm Ivol R-squared (R^2) and quarterly number of observations (N obs) are computed in each regression. t-statistics are computed using Newey and West [1987] standard errors with 3 lags and reported in parentheses below the coefficient. The dependent variable is the Slope Skew v2.

Table C.1.8: : FM Regression with Disp as Accounting Transparency (Options with $15 \leq$ d2mat ≤ 45)

	Less Accounting Transparency						
Variable	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Leverage	-0.54	-0.53	-0.49	-0.46	-0.42	-0.42	-0.33
	(-9.83)	(-10.04)	(-9.6)	(-9.44)	(-8.82)	(-8.85)	(-11.27)
Momentum		0.08	0.09	0.11	0.1	0.09	0.14
		(2.31)	(3)	(3.73)	(3.16)	(3.08)	(4.98)
Beta			0.06	0.06	0.06	0.06	0.02
			(12.36)	(12.34)	(11.92)	(11.79)	(4.67)
Size				-0.03	-0.04	-0.04	0.05
				(-4.94)	(-5.22)	(-5.19)	(6.2)
M/B					0.01	0.01	$1.6e-3$
					(8.85)	(8.77)	(2.58)
Idio Skew						$3.5e-3$	$0.3e-3$
						(1.61)	(0.15)
Atm Ivol							1.53
							(14.62)
R^2	4.61	6.20	8.19	9.67	10.25	10.28	20.42
N obs	676.00	676.00	636.00	636.00	636.00	636.00	526.00

	More Accounting Transparency						
Variable	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Leverage	-1.11	-1.10	-1.05	-0.97	-0.93	-0.93	-0.63
	(-13.36)	(-13.65)	(-12.28)	(-12.45)	(-13.03)	(-13.01)	(-11.74)
Momentum		0.14	0.15	0.2	0.19	0.19	0.18
		(2.15)	(3.23)	(5.25)	(5.13)	(5.09)	(6.07)
Beta			0.16	0.12	0.12	0.12	0.04
			(7.98)	(8.58)	(8.6)	(8.66)	(4.55)
Size				-0.15	-0.15	-0.15	-0.03
				(-11.15)	(-11.86)	(-11.85)	(-2.56)
M/B					0.01	0.01	$0.2e-3$
					(3.71)	(3.71)	(0.23)
Idio Skew						$-0.3e-3$	$3.4e-3$
						(-0.11)	(1.33)
Atm Ivol							2.66
							(11.97)
R^2	6.04	8.09	12.41	22.04	22.22	22.26	31.63
N obs	813.00	813.00	779.00	779.00	779.00	779.00	673.00

Notes: Quarterly regressions are estimated using Fama and Macbeth [1973] from January 1997 to December 2017 on all firms with listed equity options. Panel A (B), each of the 7 FM regression models is estimated using the subset of data each quarter that is below (above) 50th percentile level of Accounting Transparency. The variable (and sign) of interest is Leverage in regression (1). Control variables are added individually from regressions (2) to (7): Momentum, Beta, Size, M/B, Idio Skew, and Atm Ivol R-squared (R^2) and quarterly number of observations (N obs) are computed in each regression. t-statistics are computed using Newey and West [1987] standard errors with 3 lags and reported in parentheses below the coefficient. The dependent variable is the Slope Skew v2.

Table C.1.9: : FM Regression with Nanalyst as Accounting Transparency (Avg d2mat)

	Less Accounting Transparency						
Variable	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Leverage	-0.27	-0.27	-0.25	-0.20	-0.18	-0.18	-0.13
	(-10.82)	(-11.22)	(-10.36)	(-7.49)	(-7.06)	(-7.07)	(-5)
Momentum		$4.2e-3$	0.01	0.06	0.05	0.06	0.04
		(0.22)	(0.86)	(4.19)	(3.74)	(4.15)	(2.44)
Beta			0.04	0.03	0.03	0.03	0.01
			(12.14)	(10.48)	(10.06)	(10.15)	(2.49)
Size				-0.09	-0.09	-0.09	$2.3e-3$
				(-9.15)	(-9.22)	(-9.31)	(0.27)
M/B					0.01	0.01	$-0.4e-3$
					(8.71)	(8.82)	(-0.54)
Idio Skew						-0.01	-0.01
						(-2.23)	(-2.72)
Atm Ivol							0.90
							(13.88)
R^2	1.67	2.73	4.74	9.49	9.83	10.14	17.5
N obs	383.00	383.00	361.00	361.00	361.00	361.00	309.00

	More Accounting Transparency						
Variable	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Leverage	-0.34	-0.34	-0.33	-0.29	-0.27	-0.27	-0.21
	(-13.83)	(-15.01)	(-11.22)	(-11.34)	(-11.89)	(-11.95)	(-10.5)
Momentum		-0.04	-0.04	0.06	0.05	0.05	0.1
		(-1.07)	(-1.41)	(2.84)	(2.51)	(2.42)	(5.66)
Beta			0.09	0.06	0.06	0.06	0.01
			(7.96)	(8.22)	(8.09)	(8.13)	(4.22)
Size				-0.1	-0.1	-0.1	-0.03
				(-9.84)	(-9.96)	(-9.96)	(-5.15)
M/B					$4e-3$	$4e-3$	$-0.8e-3$
					(5.47)	(5.49)	(-2.01)
Idio Skew						$0.5e-3$	$0e-3$
						(0.42)	$(0e-3)$
Atm Ivol							1.31
							(10.79)
R^2	2.96	5.28	10.74	23.85	24.07	24.11	35.52
N obs	1257.00	1257.00	1194.00	1194.00	1194.00	1194.00	1005.00

Notes: Quarterly regressions are estimated using Fama and Macbeth [1973] from January 1997 to December 2017 on all firms with listed equity options. Panel A (B), each of the 7 FM regression models is estimated using the subset of data each quarter that is below (above) 50th percentile level of Accounting Transparency. The variable (and sign) of interest is Leverage in regression (1). Control variables are added individually from regressions (2) to (7): Momentum, Beta, Size, M/B, Idio Skew, and Atm Ivol. R-squared (R^2) and quarterly number of observations (N obs) are computed in each regression. t-statistics are computed using Newey and West [1987] standard errors with 3 lags and reported in parentheses below the coefficient. The dependent variable is the Slope Skew v2.

Table C.1.10: : FM Regression with Disp as Accounting Transparency (Avg d2mat)

	Less Accounting Transparency						
Variable	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Leverage	-0.30	-0.29	-0.27	-0.23	-0.21	-0.21	-0.17
	(-13.2)	(-13.68)	(-13.55)	(-12.45)	(-11.67)	(-11.73)	(-15.2)
Momentum		0.01	0.01	0.05	0.04	0.04	0.06
		(0.27)	(0.79)	(3.25)	(2.64)	(2.62)	(4.42)
Beta			0.04	0.03	0.03	0.03	0.01
			(11.07)	(11.41)	(10.77)	(10.71)	(4.57)
Size				-0.05	-0.05	-0.05	$2.5e-3$
				(-8.9)	(-8.95)	(-8.95)	(0.58)
M/B					0.01	0.01	$0.7e-3$
					(8.72)	(8.78)	(1.76)
Idio Skew						$1.4e-3$	$-0.1e-3$
						(1.17)	(-0.1)
Atm Ivol							0.87
							(12.73)
R^2	3.93	5.18	7.34	11.98	12.53	12.58	22.61
N obs	695.00	695.00	654.00	654.00	654.00	653.00	539.00

	More Accounting Transparency						
Variable	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Leverage	-0.51	-0.50	-0.47	-0.41	-0.39	-0.39	-0.23
	(-13.71)	(-13.91)	(-11.19)	(-11.15)	(-11.64)	(-11.62)	(-7.87)
Momentum		0.05	0.05	0.09	0.09	0.09	0.08
		(1.17)	(1.93)	(4.72)	(4.54)	(4.61)	(4.86)
Beta			0.09	0.06	0.06	0.06	0.02
			(7.67)	(8.52)	(8.54)	(8.59)	(4.4)
Size				-0.11	-0.11	-0.11	-0.04
				(-10.43)	(-10.71)	(-10.7)	(-5.85)
M/B					$3.8e-3$	$3.8e-3$	$0.3e-3$
					(4.52)	(4.51)	(0.69)
Idio Skew						$-1.6e-3$	$-0.3e-3$
						(-1.03)	(-0.24)
Atm Ivol							1.43
							(11.09)
R^2	4.00	6.07	10.63	25.54	25.67	25.72	34.6
N obs	824.00	824.00	790.00	790.00	790.00	790.00	683.00

Notes: Quarterly regressions are estimated using Fama and Macbeth [1973] from January 1997 to December 2017 on all firms with listed equity options. Panel A (B), each of the 7 FM regression models is estimated using the subset of data each quarter that is below (above) 50th percentile level of Accounting Transparency. The variable (and sign) of interest is Leverage in regression (1). Control variables are added individually from regressions (2) to (7): Momentum, Beta, Size, M/B, Idio Skew, and Atm Ivol. R-squared (R^2) and quarterly number of observations (N obs) are computed in each regression. t-statistics are computed using Newey and West [1987] standard errors with 3 lags and reported in parentheses below the coefficient. The dependent variable is the Slope Skew v2.

C.1.3 Mathematical Proofs

We re-write equation $g\left(u|z,y\right)$ in the following form.

$$g\left(u|z,y\right)=\frac{L_1}{L_0}\times e^{\frac{-(u-M_1)^2}{2\psi}}-\frac{L_2}{L_0}\times e^{\frac{-(u-M_2)^2}{2\psi}} \tag{C.1}$$

where $\tilde{u}=u-\underline{\nu}$,

$$\begin{aligned}
M_1 &= \underline{\nu}+\frac{(\tilde{y}\sigma^2t+a^2(\tilde{z}_0+mt))}{(a^2+\sigma^2t)}\\
M_2 &= \underline{\nu}+\frac{(\tilde{y}\sigma^2t-a^2(\tilde{z}_0-mt))}{(a^2+\sigma^2t)}\\
\psi &= \frac{\sigma^2a^2t}{(a^2+\sigma^2t)}\\
L_0 &= \left(1/\sqrt{\alpha/2\pi}\right)\left[\Phi\left(\sqrt{\alpha}\beta\right)-\Phi\left(\sqrt{\alpha}\left(\beta-\frac{2\tilde{z}_0}{\sigma^2t\alpha}\right)\right)e^{\left(-\frac{\alpha\eta}{2}+\frac{2\tilde{z}_0\left(\tilde{z}_0-\alpha\beta\sigma^2t\right)}{\sigma^4t^2\alpha}\right)}\right]\\
L_1 &= e^{\frac{-\left(a^2+\sigma^2t\right)}{2\sigma^2a^2t}\left[\frac{-\left(\tilde{y}\sigma^2t+a^2(\tilde{z}_0+mt)\right)^2}{\left(a^2+\sigma^2t\right)^2}+\frac{\left(\tilde{y}^2\sigma^2t+a^2(\tilde{z}_0+mt)^2\right)}{(a^2+\sigma^2t)}\right]}\\
L_2 &= e^{\frac{-\left(a^2+\sigma^2t\right)}{2\sigma^2a^2t}\left[\frac{-\left(\tilde{y}\sigma^2t-a^2(\tilde{z}_0-mt)\right)^2}{\left(a^2+\sigma^2t\right)^2}+\frac{\left(\tilde{y}^2\sigma^2t+a^2(\tilde{z}_0+mt)^2\right)}{(a^2+\sigma^2t)}\right]}
\end{aligned} \tag{C.2}$$

We then compute a closed form expression for an option pricing model of Toft and Pryck [1997] in the spirit of Duffie and Lando [2001] where the firm value is imperfectly observed. We derive the call option pricing model as:

$$\begin{aligned}
DL_{CALL} &= \int_{\underline{\nu}}^{\infty} TP_{CALL}\left(e^u,V_B,C,r,\delta,\tau,\sigma_A,T,K,t\right)g\left(u|z,y\right)du\\
&= \int_{\underline{\nu}}^{\infty} e^{u+(-r+m+\sigma^2/2)(T-t)}\left[\Phi\left(-z_1\right)-e^{(2m/\sigma^2+2)(\underline{\nu}-u)}\Phi\left(-z_2\right)\right]g\left(u|z,y\right)du
\end{aligned}$$

$$
\begin{aligned}
&+\int_{\underline{\nu}}^{\infty} Be^{-\gamma(u-\underline{\nu})}\left[\Phi\left(-z_3\right)-e^{(2m/\sigma^2-2\gamma)(\underline{\nu}-u)}\Phi\left(-z_4\right)\right] g\left(u|z,y\right)du \\
&-\int_{\underline{\nu}}^{\infty} e^{-r(T-t)}\left(A+K\right)\left[\Phi\left(-z_5\right)-e^{(2m/\sigma^2)(\underline{\nu}-u)}\Phi\left(-z_6\right)\right] g\left(u|z,y\right)du
\end{aligned}
\tag{C.3}
$$

where

$$
\begin{aligned}
-z_1 &= \tfrac{-y^*+\left(m(T-t)+u+\sigma^2(T-t)\right)}{\sigma\sqrt{T-t}} & -z_2 &= \tfrac{-y^*+\left(m(T-t)-u+2\underline{\nu}+\sigma^2(T-t)\right)}{\sigma\sqrt{T-t}} \\
-z_3 &= \tfrac{-y^*+\left(m(T-t)+u-\gamma\sigma^2(T-t)\right)}{\sigma\sqrt{T-t}} & -z_4 &= \tfrac{-y^*+\left(m(T-t)-u+2\underline{\nu}-\gamma\sigma^2(T-t)\right)}{\sigma\sqrt{T-t}} \\
-z_5 &= \tfrac{-y^*+(m(T-t)+u)}{\sigma\sqrt{T-t}} & -z_6 &= \tfrac{-y^*+(m(T-t)-u+2\underline{\nu})}{\sigma\sqrt{T-t}}
\end{aligned}
\tag{C.4}
$$

To compute the integral in equation C.21, note that equation C.19 is essentially a difference in the Normally distributed random variable probability distribution functions (PDF) being integrated over the TP option pricing formula, which is a function of Normal distributed cumulative distribution functions (CDFS). As such the end product will be an integral, over the range of either $(-\infty,\underline{\nu})$ for a put or $(\underline{\nu},\infty)$ for a call, of the product of a Normally distributed CDF and the Normally PDF. To compute this integral we make use of the equations $10,010.1$ and $10,010.4$ in Owen [1980] (stated in Lemmas C.1.3 and C.1.4 respectively) which allows for a closed form expression of the integral. The result itself is then just a function of the Bivariate Normally distributed cumulative probability distribution function.

Lemma C.1.1. *Let A and B are real valued constants and $Z\sim\mathcal{N}(0,1)$, with probability (cumulative) distribution function denoted $\phi(z)$ $(\Phi(z))$ respectively then*

$$
\int_{-\infty}^{\underline{\nu}}\Phi(A+Bz)\phi(z)dz=\Phi_2\left(\frac{A}{\sqrt{1+B^2}},\underline{\nu};\frac{-B}{\sqrt{1+B^2}}\right)
\tag{C.5}
$$

where $\Phi_2\left(z_1,z_2;\rho\right)$ is the cumulative bivariate normal distribution of two joint random variables Z_1 and Z_2 with correlation ρ.

Lemma C.1.2. *Let A and B are real valued constants and $Z\sim\mathcal{N}(0,1)$, with probability (cumulative) distribution function denoted $\phi(z)$ $(\Phi(z))$ respectively then*

$$
\int_h^k\Phi(A+Bx)\phi(x)dx=\int_{-\infty}^{A/\sqrt{1+B^2}}\phi(x)\Phi\left(k\sqrt{B^2+1}+Bx\right)dx
$$

$$
\begin{aligned}
&- \int_{-\infty}^{A/\sqrt{1+B^2}} \phi(x)\Phi\left(h\sqrt{B^2+1}+Bx\right)dx \\
&= \Phi_2\left(\frac{A}{\sqrt{1+B^2}}, k; \frac{-B}{\sqrt{1+B^2}}\right) - \Phi_2\left(\frac{A}{\sqrt{1+B^2}}, h; \frac{-B}{\sqrt{1+B^2}}\right) \qquad \text{(C.6)}
\end{aligned}
$$

Taking $k \longrightarrow \infty$ in equation C.26 yields:

$$
\begin{aligned}
\int_h^{\infty} \Phi(A+Bx)\phi(x)dx &= \lim_{k\to\infty} \Phi_2\left(\frac{A}{\sqrt{1+B^2}}, k; \frac{-B}{\sqrt{1+B^2}}\right) - \Phi_2\left(\frac{A}{\sqrt{1+B^2}}, h; \frac{-B}{\sqrt{1+B^2}}\right) \\
&= \Phi\left(\frac{A}{\sqrt{1+B^2}}\right) - \Phi_2\left(\frac{A}{\sqrt{1+B^2}}, h; \frac{-B}{\sqrt{1+B^2}}\right) \qquad \text{(C.7)}
\end{aligned}
$$

In order to solve for our option pricing model with incomplete accounting information, given by H_t, in closed form, the function $g\left(u|y, z_0, t\right)$ can be re-written as a difference of Normal distribution probability density functions with different means and variances in the general form (see appendix for details of derivation and notation):

$$
g\left(u|z_0, y\right) = \frac{L_1}{L_0} \times e^{\frac{-(u-M_1)^2}{2\psi}} - \frac{L_2}{L_0} \times e^{\frac{-(u-M_2)^2}{2\psi}} \qquad \text{(C.8)}
$$

Our formula for H_t involves the integral of the product of the TP option pricing formula (denoted $h_t\left(u\right)$) and the density function $g\left(u|y, z_0, t\right)$. Expanding the product results in computing twelve integrals of the form in equation 4.3.17 (see Owen [1980]). The resulting sum of the twelve integrals results in a closed for solution, up to Bivariate Normal probability function, of the expression H_t.

Next we apply the framework of Duffie and Lando [2001] to the above call option option pricing formulas above:

$$
g\left(u|z, y\right) = \frac{\sqrt{\alpha/(2\pi)}\left[1-e^{\left(\frac{-2\tilde{z}_0\tilde{u}}{\sigma^2 t}\right)}\right] e^{\left(\frac{-(\tilde{y}-\tilde{u})^2}{2a^2}\right)} e^{\left(\frac{-(\tilde{u}-\tilde{z}_0-mt)^2}{2\sigma^2 t}\right)}}{\Phi\left(\sqrt{\alpha}\beta\right) - \Phi\left(\sqrt{\alpha}\left(\beta-\frac{2\tilde{z}_0}{\sigma^2 t\alpha}\right)\right) e^{\left(-\frac{\alpha\eta}{2}+\frac{2\tilde{z}_0\left(\tilde{z}_0-\alpha\beta\sigma^2 t\right)}{\sigma^4 t^2\alpha}\right)}} \qquad \text{(C.9)}
$$

where

$$
\begin{array}{ll}
\tilde{z}_0 = z_0 - \underline{\nu} & \alpha = \frac{\sigma^2 t + a^2}{a^2 \sigma^2 t} \\
\tilde{u} = u - \underline{\nu} & \beta = \frac{\sigma^2 t \tilde{y} + a^2 (\tilde{z}_0 + mt)}{\sigma^2 t + a^2} \\
\tilde{y} = y - \underline{\nu} & \eta = \frac{a^2 \sigma^2 t (\tilde{y} - (\tilde{z}_0 + mt))^2}{\left(\sigma^2 t + a^2\right)^2}
\end{array}
\tag{C.10}
$$

The denominator of equation C.9 is a constant in u since it does not depend on u whereas the numerator of equation C.9 can be broken down (working in the exponential term) as:

$$
\begin{aligned}
&\sigma^2 t \left(\tilde{y} - \tilde{u}\right)^2 + a^2 \left(\tilde{u} - \tilde{z}_0 - mt\right)^2 \\
&= \sigma^2 t \left(\tilde{y}^2 - 2\tilde{y}\tilde{u} + \tilde{u}^2\right) + a^2 \left(\tilde{u}^2 - 2\tilde{u}\left(\tilde{z}_0 + mt\right) + \left(\tilde{z}_0 + mt\right)^2\right) \\
&= \left(a^2 + \sigma^2 t\right)\tilde{u}^2 - 2\tilde{u}\left(\tilde{y}\sigma^2 t + a^2(\tilde{z}_0 + mt)\right) + \left(\tilde{y}^2\sigma^2 t + a^2\left(\tilde{z}_0 + mt\right)^2\right) \\
&= \left(a^2 + \sigma^2 t\right)\left[\tilde{u}^2 - 2\tilde{u}\frac{(\tilde{y}\sigma^2 t + a^2(\tilde{z}_0 + mt))}{(a^2 + \sigma^2 t)} + \left(\frac{(\tilde{y}\sigma^2 t + a^2(\tilde{z}_0 + mt))}{(a^2 + \sigma^2 t)}\right)^2\right] - \frac{(\tilde{y}\sigma^2 t + a^2(\tilde{z}_0 + mt))^2}{(a^2 + \sigma^2 t)} + \left(\tilde{y}^2\sigma^2 t + a^2\left(\tilde{z}_0 + mt\right)^2\right) \\
&= \left(a^2 + \sigma^2 t\right)\left[\tilde{u} - \frac{(\tilde{y}\sigma^2 t + a^2(\tilde{z}_0 + mt))}{(a^2 + \sigma^2 t)}\right]^2 - \frac{(\tilde{y}\sigma^2 t + a^2(\tilde{z}_0 + mt))^2}{(a^2 + \sigma^2 t)} + \left(\tilde{y}^2\sigma^2 t + a^2\left(\tilde{z}_0 + mt\right)^2\right)
\end{aligned}
\tag{C.11}
$$

This results in the simplified expression of the second part of the numerator:

$$
\begin{aligned}
&e^{\left(\frac{-\sigma^2 t (\tilde{y} - \tilde{u})^2}{2a^2\sigma^2 t} - \frac{a^2(\tilde{u} - \tilde{z}_0 - mt)^2}{2a^2\sigma^2 t}\right)} \\
&= e^{-\frac{\left(a^2 + \sigma^2 t\right)\left[\tilde{u} - \frac{\left(\tilde{y}\sigma^2 t + a^2(\tilde{z}_0 + mt)\right)}{(a^2 + \sigma^2 t)}\right]^2 - \frac{\left(\tilde{y}\sigma^2 t + a^2(\tilde{z}_0 + mt)\right)^2}{(a^2 + \sigma^2 t)} + \left(\tilde{y}^2\sigma^2 t + a^2(\tilde{z}_0 + mt)^2\right)}{2a^2\sigma^2 t}}
\end{aligned}
\tag{C.12}
$$

We then combine the intermediate step in equation C.12 to obtain the first part of the numerator in equation C.9 (working on the terms in the exponentials):

$$
\begin{aligned}
&4a^2\tilde{z}_0\tilde{u} + \left(a^2 + \sigma^2 t\right)\tilde{u}^2 - 2\tilde{u}\left(\tilde{y}\sigma^2 t + a^2(\tilde{z}_0 + mt)\right) + \left(\tilde{y}^2\sigma^2 t + a^2\left(\tilde{z}_0 + mt\right)^2\right) \\
&= \left(a^2 + \sigma^2 t\right)\tilde{u}^2 - 2\tilde{u}\left(\tilde{y}\sigma^2 t + a^2(\tilde{z}_0 + mt) - 2a^2\tilde{z}_0\right) + \left(\tilde{y}^2\sigma^2 t + a^2\left(\tilde{z}_0 + mt\right)^2\right) \\
&= \left(a^2 + \sigma^2 t\right)\tilde{u}^2 - 2\tilde{u}\left(\tilde{y}\sigma^2 t - a^2(\tilde{z}_0 - mt)\right) + \left(\tilde{y}^2\sigma^2 t + a^2\left(\tilde{z}_0 + mt\right)^2\right)
\end{aligned}
$$

$$= \left(a^2+\sigma^2 t\right)\left[\tilde{u}^2 - 2\tilde{u}\frac{(\tilde{y}\sigma^2 t - a^2(\tilde{z}_0 - mt))}{(a^2+\sigma^2 t)} + \left(\frac{(\tilde{y}\sigma^2 t - a^2(\tilde{z}_0 - mt))}{(a^2+\sigma^2 t)}\right)^2\right] - \frac{(\tilde{y}\sigma^2 t - a^2(\tilde{z}_0 - mt))^2}{(\alpha^2+\sigma^2 t)} + \left(\tilde{y}^2\sigma^2 t + \alpha^2\left(\tilde{z}_0 + mt\right)^2\right)$$

$$= \left(a^2+\sigma^2 t\right)\left[\tilde{u} - \frac{(\tilde{y}\sigma^2 t - a^2(\tilde{z}_0 - mt))}{(a^2+\sigma^2 t)}\right]^2 - \frac{(\tilde{y}\sigma^2 t - a^2(\tilde{z}_0 - mt))^2}{(a^2+\sigma^2 t)} + \left(\tilde{y}^2\sigma^2 t + a^2\left(\tilde{z}_0 + mt\right)^2\right) \tag{C.13}$$

Which results in the following exponential term:

$$= e^{\frac{-\left(\left(a^2+\sigma^2 t\right)\left[\tilde{u}-\frac{\left(\tilde{y}\sigma^2 t-a^2(\tilde{z}_0-mt)\right)}{(a^2+\sigma^2 t)}\right]^2-\frac{\left(\tilde{y}\sigma^2 t-a^2(\tilde{z}_0-mt)\right)^2}{(a^2+\sigma^2 t)}+\left(\tilde{y}^2\sigma^2 t+a^2(\tilde{z}_0+mt)^2\right)\right)}{2a^2\sigma^2 t}} \tag{C.14}$$

We can then re-write the numerator of equation C.9 as the difference between equations C.12 and C.14, hence:

$$\left[1-e^{\left(\frac{-2\tilde{z}_0\tilde{u}}{\sigma^2 t}\right)}\right]e^{\left(\frac{-(\tilde{y}-\tilde{u})^2}{2a^2}\right)}e^{\left(\frac{-(\tilde{u}-\tilde{z}_0-mt)^2}{2\sigma^2 t}\right)} = e^{-\frac{\left(a^2+\sigma^2 t\right)\left[\tilde{u}-\frac{\left(\tilde{y}\sigma^2 t+a^2(\tilde{z}_0+mt)\right)}{(a^2+\sigma^2 t)}\right]^2-\frac{\left(\tilde{y}\sigma^2 t+a^2(\tilde{z}_0+mt)\right)^2}{(a^2+\sigma^2 t)}+\left(\tilde{y}^2\sigma^2 t+a^2(\tilde{z}_0+mt)^2\right)}{2a^2\sigma^2 t}}$$
$$- e^{\frac{-\left(\left(a^2+\sigma^2 t\right)\left[\tilde{u}-\frac{\left(\tilde{y}\sigma^2 t-a^2(\tilde{z}_0-mt)\right)}{(a^2+\sigma^2 t)}\right]^2-\frac{\left(\tilde{y}\sigma^2 t-a^2(\tilde{z}_0-mt)\right)^2}{(a^2+\sigma^2 t)}+\left(\tilde{y}^2\sigma^2 t+a^2(\tilde{z}_0+mt)^2\right)\right)}{2a^2\sigma^2 t}} \tag{C.15}$$

we can then re-write equation C.9 as

$$g\left(u|z,y\right) = \frac{e^{-\frac{\left(\left(a^2+\sigma^2 t\right)\left[\tilde{u}-\frac{\left(\tilde{y}\sigma^2 t+a^2(\tilde{z}_0+mt)\right)}{(a^2+\sigma^2 t)}\right]^2-\frac{\left(\tilde{y}\sigma^2 t+a^2(\tilde{z}_0+mt)\right)^2}{(a^2+\sigma^2 t)}+\left(\tilde{y}^2\sigma^2 t+a^2(\tilde{z}_0+mt)^2\right)\right)}{2a^2\sigma^2 t}}}{\Phi\left(\sqrt{\alpha}\beta\right)-\Phi\left(\sqrt{\alpha}\left(\beta-\frac{2\tilde{z}_0}{\sigma^2 t\alpha}\right)\right)e^{\left(-\frac{\alpha\eta}{2}+\frac{2\tilde{z}_0\left(\tilde{z}_0-\alpha\beta\sigma^2 t\right)}{\sigma^4 t^2\alpha}\right)}} - \frac{e^{\frac{-\left(\left(a^2+\sigma^2 t\right)\left[\tilde{u}-\frac{\left(\tilde{y}\sigma^2 t-a^2(\tilde{z}_0-mt)\right)}{(a^2+\sigma^2 t)}\right]^2-\frac{\left(\tilde{y}\sigma^2 t-a^2(\tilde{z}_0-mt)\right)^2}{(a^2+\sigma^2 t)}+\left(\tilde{y}^2\sigma^2 t+a^2(\tilde{z}_0+mt)^2\right)\right)}{2a^2\sigma^2 t}}}{\Phi\left(\sqrt{\alpha}\beta\right)-\Phi\left(\sqrt{\alpha}\left(\beta-\frac{2\tilde{z}_0}{\sigma^2 t\alpha}\right)\right)e^{\left(-\frac{\alpha\eta}{2}+\frac{2\tilde{z}_0\left(\tilde{z}_0-\alpha\beta\sigma^2 t\right)}{\sigma^4 t^2\alpha}\right)}} \tag{C.16}$$

We then denote the following constants, which are not a function of $\tilde{u}$ to simplify the notation of equation C.16

$$L_0 = \left(1/\sqrt{\alpha/2\pi}\right)\left[\Phi\left(\sqrt{\alpha}\beta\right) - \Phi\left(\sqrt{\alpha}\left(\beta - \frac{2\tilde{z}_0}{\sigma^2 t\alpha}\right)\right) e^{\left(-\frac{\alpha\eta}{2} + \frac{2\tilde{z}_0\left(\tilde{z}_0 - \alpha\beta\sigma^2 t\right)}{\sigma^4 t^2 \alpha}\right)}\right] \tag{C.17}$$

$$L_1 = e^{\frac{-\left(a^2+\sigma^2 t\right)}{2\sigma^2 a^2 t}\left[\frac{-\left(\tilde{y}\sigma^2 t + a^2(\tilde{z}_0+mt)\right)^2}{(a^2+\sigma^2 t)^2} + \frac{\left(\tilde{y}^2\sigma^2 t + a^2(\tilde{z}_0+mt)^2\right)}{(a^2+\sigma^2 t)}\right]} \qquad L_2 = e^{\frac{-\left(a^2+\sigma^2 t\right)}{2\sigma^2 a^2 t}\left[\frac{-\left(\tilde{y}\sigma^2 t - a^2(\tilde{z}_0-mt)\right)^2}{(a^2+\sigma^2 t)^2} + \frac{\left(\tilde{y}^2\sigma^2 t + a^2(\tilde{z}_0+mt)^2\right)}{(a^2+\sigma^2 t)}\right]} \tag{C.18}$$

which results in

$$g\left(u|z,y\right) = \frac{L_1}{L_0} \times e^{\frac{-[u-M_1]^2}{2\psi}} - \frac{L_2}{L_0} \times e^{\frac{-[u-M_2]^2}{2\psi}} \tag{C.19}$$

where $\tilde{u} = u - \underline{\nu}$,

$$\begin{aligned} M_1 &= \underline{\nu} + \frac{(\tilde{y}\sigma^2 t + a^2(\tilde{z}_0 + mt))}{(a^2+\sigma^2 t)} \\ M_2 &= \underline{\nu} + \frac{(\tilde{y}\sigma^2 t - a^2(\tilde{z}_0 - mt))}{(a^2+\sigma^2 t)} \\ \psi &= \frac{\sigma^2 a^2 t}{(a^2+\sigma^2 t)} \end{aligned} \tag{C.20}$$

We then compute, using equation C.16 to compute a closed form expression for an option pricing model of Toft and Pryck [1997] in the spirit of Duffie and Lando [2001] where the firm value is imperfectly observed. We derive the call option pricing model as:

$$\begin{aligned} DL_{CALL} &= \int_{\underline{\nu}}^{\infty} TP_{CALL}\left(e^u, V_B, C, r, \delta, \tau, \sigma_A, T, K, t\right) g\left(u|z,y\right) du \\ &= \int_{\underline{\nu}}^{\infty} e^{u+(-r+m+\sigma^2/2)(T-t)}\left[\Phi\left(-z_1\right) - e^{(2m/\sigma^2+2)(\underline{\nu}-u)}\Phi\left(-z_2\right)\right] g\left(u|z,y\right) du \end{aligned}$$

$$
\begin{aligned}
&+ \int_{\underline{\nu}}^{\infty} Be^{-\gamma(u-\underline{\nu})} \left[\Phi\left(-z_3\right) - e^{(2m/\sigma^2-2\gamma)(\underline{\nu}-u)}\Phi\left(-z_4\right)\right] g\left(u|z,y\right) du \\
&- \int_{\underline{\nu}}^{\infty} e^{-r(T-t)}\left(A+K\right)\left[\Phi\left(-z_5\right) - e^{(2m/\sigma^2)(\underline{\nu}-u)}\Phi\left(-z_6\right)\right] g\left(u|z,y\right) du
\end{aligned}
\tag{C.21}
$$

where

$$
\begin{array}{ll}
-z_1 = \frac{-y^*+\left(m(T-t)+u+\sigma^2(T-t)\right)}{\sigma\sqrt{T-t}} & -z_2 = \frac{-y^*+\left(m(T-t)-u+2\underline{\nu}+\sigma^2(T-t)\right)}{\sigma\sqrt{T-t}} \\
-z_3 = \frac{-y^*+\left(m(T-t)+u-\gamma\sigma^2(T-t)\right)}{\sigma\sqrt{T-t}} & -z_4 = \frac{-y^*+\left(m(T-t)-u+2\underline{\nu}-\gamma\sigma^2(T-t)\right)}{\sigma\sqrt{T-t}} \\
-z_5 = \frac{-y^*+(m(T-t)+u)}{\sigma\sqrt{T-t}} & -z_6 = \frac{-y^*+(m(T-t)-u+2\underline{\nu})}{\sigma\sqrt{T-t}}
\end{array}
\tag{C.22}
$$

To compute the integral in equation C.21, note that equation C.19 is essentially a difference in the Normally distributed random variable probability distribution functions (PDF) being integrated over the TP option pricing formula, which is a function of Normal distributed cumulative distribution functions (CDFS). As such the end product will be an integral, over the range of either $(-\infty, \underline{\nu})$ for a put or $(\underline{\nu}, \infty)$ for a call, of the product of a Normally distributed CDF and the Normally PDF. To compute this integral we make use of the equations $10,010.1$ and $10,010.4$ in Owen [1980] which allows for a closed form expression of the integral. The result itself is then just a function of the Bivariate Normally distributed cumulative probability distribution function. We provide a description and proof in Lemma C.1.3.

Lemma C.1.3. *Let A and B are real valued constants and $Z \sim \mathcal{N}(0,1)$, with probability (cumulative) distribution function denoted $\phi(z)$ $(\Phi(z))$ respectively then*

$$
\int_{-\infty}^{\underline{\nu}} \Phi(A+Bz)\phi(z)dz = \Phi_2\left(\frac{A}{\sqrt{1+B^2}}, \underline{\nu}; \frac{-B}{\sqrt{1+B^2}}\right)
\tag{C.23}
$$

where $\Phi_2\left(z_1, z_2; \rho\right)$ is the cumulative bivariate normal distribution of two joint random variables Z_1 and Z_2 with correlation ρ.

Proof.

$$
\int_{-\infty}^{\underline{\nu}} \Phi(A+Bz)\phi(z)dz = \int_{-\infty}^{\underline{\nu}}\int_{-\infty}^{A+Bz} \phi(x)\phi(z)dxdz
$$

$$
\begin{aligned}
&= \int_{-\infty}^{\underline{\nu}} \int_{-\infty}^{A+Bz} \frac{e^{\frac{-1(x^2+z^2)}{2}}}{\sqrt{2\pi}\sqrt{2\pi}} dydz \\
&= \int_{-\infty}^{\underline{\nu}} \int_{-\infty}^{A/\sqrt{1+B^2}} \frac{\sqrt{1+B^2} e^{\frac{-1\left((y\sqrt{1+B^2}+Bz)^2+z^2\right)}{2}}}{\sqrt{2\pi}\sqrt{2\pi}} dydz \\
&= \int_{-\infty}^{\underline{\nu}} \int_{-\infty}^{A/\sqrt{1+B^2}} \frac{e^{\frac{-1\left(y^2(1+B^2)+2yzB\sqrt{1+B^2}+B^2z^2+z^2\right)}{2}}}{\sqrt{2\pi}\sqrt{2\pi}\sqrt{1/(1+B^2)}} dydz \\
&= \Phi_2\left(\frac{A}{\sqrt{1+B^2}}, \underline{\nu}; \frac{-B}{\sqrt{1+B^2}}\right) \qquad \text{(C.24)}
\end{aligned}
$$

□

where we let $y = (x - Bz)/\sqrt{1+B^2}$ $(x = y\sqrt{1+B^2} + Bz)$ hence $dy = dx/\sqrt{1+B^2}$ and

$$
\begin{aligned}
y^2(1+B^2) + 2yzB\sqrt{1+B^2} + z^2(1+B^2) &= (1+B^2)\left[y^2 + 2yz\frac{B}{\sqrt{1+B^2}} + z^2\right] \\
&= \frac{1}{1+B^2}\left[y^2 - 2yz\left(\frac{-B}{\sqrt{1+B^2}}\right) + z^2\right] \qquad \text{(C.25)}
\end{aligned}
$$

Lemma C.1.4. *Let A and B are real valued constants and $Z \sim \mathcal{N}(0,1)$, with probability (cumulative) distribution function denoted $\phi(z)$ $(\Phi(z))$ respectively then*

$$
\begin{aligned}
\int_h^k \Phi(A+Bx)\phi(x)dx &= \int_{-\infty}^{A/\sqrt{1+B^2}} \phi(x)\Phi\left(k\sqrt{B^2+1} + Bx\right)dx \\
&- \int_{-\infty}^{A/\sqrt{1+B^2}} \phi(x)\Phi\left(h\sqrt{B^2+1} + Bx\right)dx \\
&= \Phi_2\left(\frac{A}{\sqrt{1+B^2}}, k; \frac{-B}{\sqrt{1+B^2}}\right) - \Phi_2\left(\frac{A}{\sqrt{1+B^2}}, h; \frac{-B}{\sqrt{1+B^2}}\right) \qquad \text{(C.26)}
\end{aligned}
$$

Proof.

$$
\int_h^k \Phi(A+Bx)\phi(x)dx = \int_{-\infty}^k \Phi(A+Bx)\phi(x)dx - \int_{-\infty}^h \Phi(A+Bx)\phi(x)dx
$$

$$= \Phi_2\left(\frac{A}{\sqrt{1+B^2}}, k; \frac{-B}{1+B^2}\right) - \Phi_2\left(\frac{A}{\sqrt{1+B^2}}, h; \frac{-B}{1+B^2}\right) \tag{C.27}$$

□

Taking $k \longrightarrow \infty$ in equation C.26 yields:

$$\begin{aligned}
\int_h^\infty \Phi(A+Bx)\phi(x)dx &= \lim_{k\to\infty} \Phi_2\left(\frac{A}{\sqrt{1+B^2}}, k; \frac{-B}{\sqrt{1+B^2}}\right) - \Phi_2\left(\frac{A}{\sqrt{1+B^2}}, h; \frac{-B}{\sqrt{1+B^2}}\right) \\
&= \Phi\left(\frac{A}{\sqrt{1+B^2}}\right) - \Phi_2\left(\frac{A}{\sqrt{1+B^2}}, h; \frac{-B}{\sqrt{1+B^2}}\right)
\end{aligned} \tag{C.28}$$

We next derive the following transformation to evaluate an integral

$$\begin{aligned}
&\int_{\underline{\nu}}^\infty e^{\theta u}\Phi\left(A+Bu\right)\phi(u;m,s^2)du \\
&= \int_{\underline{\nu}}^\infty e^{\theta u}\Phi\left(A+Bu\right)\frac{e^{\frac{-(u-m)^2}{2s^2}}}{\sqrt{2\pi s^2}}du \\
&= \int_{\underline{\nu}}^\infty \Phi\left(A+Bu\right)\frac{e^{\frac{-(u^2-2um+m^2-2s^2u\theta)}{2s^2}}}{\sqrt{2\pi s^2}}du \\
&= \int_{\underline{\nu}}^\infty \Phi\left(A+Bu\right)\frac{e^{\frac{-(u^2-2um+m^2-2s^2u\theta)}{2s^2}}}{\sqrt{2\pi s^2}}du \\
&= e^{\frac{-m^2+(m+s^2\theta)^2}{2s^2}}\int_{\underline{\nu}}^\infty \Phi\left(A+Bu\right)\frac{e^{\frac{-(u^2-2u(m+s^2\theta)+(m+s^2\theta)^2)}{2s^2}}}{\sqrt{2\pi s^2}}du \\
&= e^{\frac{-m^2+(m+s^2\theta)^2}{2s^2}}\int_{\underline{\nu}}^\infty \Phi\left(A+Bu\right)\frac{e^{\frac{-(u-(m+s^2\theta))^2}{2s^2}}}{\sqrt{2\pi s^2}}du \\
&= e^{\frac{-m^2+(m+s^2\theta)^2}{2s^2}}\int_{\underline{\nu}'}^\infty \Phi\left(A+B(sz+(m+s^2\theta))\right)\phi(z)dz
\end{aligned}$$

$$= e^{\theta m+(s\theta)^2/2}\left[\Phi\left(\frac{A'}{\sqrt{1+(Bs)^2}}\right) - \Phi_2\left(\frac{A'}{\sqrt{1+(Bs)^2}}, \underline{\nu}'; \frac{-Bs}{\sqrt{1+(Bs)^2}}\right)\right] \tag{C.29}$$

where we let $z = (u-(m+s^2\theta))/s$ which results in $A' = A + B(m+s^2\theta)$ and $\underline{\nu}' = (\underline{\nu}-(m+s^2\theta))/s$. In the resulting equation C.29 we then set $\theta = \{1, -2m/\sigma^2-1, -\gamma, -2m/\sigma^2+\gamma, -2m/\sigma^2\}$ to evaluate some of the integrals in equation C.30.

$$\begin{aligned} DL_{CALL} &= \int_{\underline{\nu}}^{\infty} e^{u-\delta(T-t)}\left[\Phi\left(-z_1\right) - e^{(2m/\sigma^2+2)(\underline{\nu}-u)}\Phi\left(-z_2\right)\right] g\left(u|z,y\right) du \\ &+ \int_{\underline{\nu}}^{\infty} Be^{-\gamma(u-\underline{\nu})}\left[\Phi\left(-z_3\right) - e^{(2m/\sigma^2-2\gamma)(\underline{\nu}-u)}\Phi\left(-z_4\right)\right] g\left(u|z,y\right) du \\ &- \int_{\underline{\nu}}^{\infty} e^{-r(T-t)}\left(A+K\right)\left[\Phi\left(-z_5\right) - e^{(2m/\sigma^2)(\underline{\nu}-u)}\Phi\left(-z_6\right)\right] g\left(u|z,y\right) du \end{aligned} \tag{C.30}$$

where

$$\begin{aligned} -z_1 &= \frac{-y^*+\left(m(T-t)+u+\sigma^2(T-t)\right)}{\sigma\sqrt{T-t}} & -z_2 &= \frac{-y^*+\left(m(T-t)-u+2\underline{\nu}+\sigma^2(T-t)\right)}{\sigma\sqrt{T-t}} \\ -z_3 &= \frac{-y^*+\left(m(T-t)+u-\gamma\sigma^2(T-t)\right)}{\sigma\sqrt{T-t}} & -z_4 &= \frac{-y^*+\left(m(T-t)-u+2\underline{\nu}-\gamma\sigma^2(T-t)\right)}{\sigma\sqrt{T-t}} \\ -z_5 &= \frac{-y^*+(m(T-t)+u)}{\sigma\sqrt{T-t}} & -z_6 &= \frac{-y^*+(m(T-t)-u+2\underline{\nu})}{\sigma\sqrt{T-t}} \end{aligned} \tag{C.31}$$

For evaluating the first integral in equation C.30 we make use of the result from equation C.29

$$\begin{aligned} &\int_{\underline{\nu}}^{\infty} e^{u-\delta(T-t)}\left[\Phi\left(-z_1\right) - e^{(2m/\sigma^2+2)(\underline{\nu}-u)}\Phi\left(-z_2\right)\right] g\left(u|z,y\right) du \\ &= \int_{\underline{\nu}}^{\infty} e^{u-\delta(T-t)}\left[\Phi\left(-z_1\right) - e^{(2m/\sigma^2+2)(\underline{\nu}-u)}\Phi\left(-z_2\right)\right]\left[\frac{L_1}{L_0}\times e^{-\frac{(u-M_1)^2}{2\psi}} - \frac{L_2}{L_0}\times e^{-\frac{(u-M_2)^2}{2\psi}}\right] du \\ &= e^{-\delta(T-t)}\int_{\underline{\nu}}^{\infty} e^{u}\Phi\left(-z_1\right)\left[\frac{L_1}{L_0}\times e^{-\frac{(u-M_1)^2}{2\psi}} - \frac{L_2}{L_0}\times e^{-\frac{(u-M_2)^2}{2\psi}}\right] du \\ &- e^{-\delta(T-t)}\int_{\underline{\nu}}^{\infty} e^{u}e^{(2m/\sigma^2+2)(\underline{\nu}-u)}\Phi\left(-z_2\right)\left[\frac{L_1}{L_0}\times e^{-\frac{(u-M_1)^2}{2\psi}} - \frac{L_2}{L_0}\times e^{-\frac{(u-M_2)^2}{2\psi}}\right] du \end{aligned}$$

$$
\begin{aligned}
&= e^{-\delta(T-t)} \int_{\underline{\nu}}^{\infty} e^{u} \Phi\left(-z_1\right) \frac{L_1}{L_0} \times e^{-\frac{(u-M_1)^2}{2\psi}} du - e^{-\delta(T-t)} \int_{\underline{\nu}}^{\infty} e^{u} \Phi\left(-z_1\right) \frac{L_2}{L_0} \times e^{-\frac{(u-M_2)^2}{2\psi}} du \\
&- e^{-\delta(T-t)} \int_{\underline{\nu}}^{\infty} e^{u} e^{(2m/\sigma^2+2)(\underline{\nu}-u)} \Phi\left(-z_2\right) \frac{L_1}{L_0} \times e^{-\frac{(u-M_1)^2}{2\psi}} du \\
&+ e^{-\delta(T-t)} \int_{\underline{\nu}}^{\infty} e^{u} e^{(2m/\sigma^2+2)(\underline{\nu}-u)} \Phi\left(-z_2\right) \frac{L_2}{L_0} \times e^{-\frac{(u-M_2)^2}{2\psi}} du \\
&= e^{-\delta(T-t)} \frac{L_1}{L_0} \times \int_{\underline{\nu}}^{\infty} e^{u} \Phi\left(-z_1\right) e^{-\frac{(u-M_1)^2}{2\psi}} du - e^{-\delta(T-t)} \frac{L_2}{L_0} \times \int_{\underline{\nu}}^{\infty} e^{u} \Phi\left(-z_1\right) e^{-\frac{(u-M_2)^2}{2\psi}} du \\
&- e^{-\delta(T-t)} e^{(2m/\sigma^2+2)(\underline{\nu})} \frac{L_1}{L_0} \times \int_{\underline{\nu}}^{\infty} e^{u(-2m/\sigma^2-1)} \Phi\left(-z_2\right) e^{-\frac{(u-M_1)^2}{2\psi}} du \\
&+ e^{-\delta(T-t)} e^{(2m/\sigma^2+2)\underline{\nu}} \frac{L_2}{L_0} \times \int_{\underline{\nu}}^{\infty} e^{u(-2m/\sigma^2-1)} \Phi\left(-z_2\right) e^{-\frac{(u-M_2)^2}{2\psi}} du
\end{aligned}
\tag{C.32}
$$

$$
\begin{aligned}
&= e^{-\delta(T-t)} \frac{L_1}{L_0} \sqrt{2\pi\psi} e^{M_1+\psi/2} \left[\Phi \left(\frac{\frac{-y^* + (M_1+\psi) + \left(m(T-t)+\sigma^2(T-t)\right)}{\sigma\sqrt{T-t}}}{\sqrt{1+\left(\sqrt{\psi}/(\sigma\sqrt{T-t})\right)^2}} \right) \right] \\
&- e^{-\delta(T-t)} \frac{L_1}{L_0} \sqrt{2\pi\psi} e^{M_1+\psi/2} \Phi_2 \left(\frac{\frac{-y^* + (M_1+\psi) + \left(m(T-t)+\sigma^2(T-t)\right)}{\sigma\sqrt{T-t}}}{\sqrt{1+\left(\sqrt{\psi}/(\sigma\sqrt{T-t})\right)^2}}, \frac{\underline{\nu} - (M_1+\psi)}{\sqrt{\psi}}; \left(\frac{-(\sqrt{\psi})/(\sigma\sqrt{T-t})}{\sqrt{1+\left(\sqrt{\psi}/(\sigma\sqrt{T-t})\right)^2}} \right) \right) \\
&- e^{-\delta(T-t)} \frac{L_2}{L_0} \sqrt{2\pi\psi} e^{M_2+\psi/2} \left[\Phi \left(\frac{\frac{-y^* + (M_2+\psi) + \left(m(T-t)+\sigma^2(T-t)\right)}{\sigma\sqrt{T-t}}}{\sqrt{1+\left(\sqrt{\psi}/(\sigma\sqrt{T-t})\right)^2}} \right) \right] \\
&+ e^{-\delta(T-t)} \frac{L_2}{L_0} \sqrt{2\pi\psi} e^{M_2+\psi/2} \Phi_2 \left(\frac{\frac{-y^* + (M_2+\psi) + \left(m(T-t)+\sigma^2(T-t)\right)}{\sigma\sqrt{T-t}}}{\sqrt{1+\left(\sqrt{\psi}/(\sigma\sqrt{T-t})\right)^2}}, \frac{\underline{\nu} - (M_2+\psi)}{\sqrt{\psi}}; \left(\frac{-\sqrt{\psi}/(\sigma\sqrt{T-t})}{\sqrt{1+\left(\sqrt{\psi}/(\sigma\sqrt{T-t})\right)^2}} \right) \right) \\
&- e^{-\delta(T-t)} e^{(2m/\sigma^2+2)\underline{\nu}} \frac{L_1}{L_0} \sqrt{2\pi\psi} e^{M_1(-2m/\sigma^2-1)+\psi(-2m/\sigma^2-1)^2/2} \times \left[\Phi \left(\frac{\frac{-y^* + (M_1+\psi(-2m/\sigma^2-1)) + \left(m(T-t)+2\underline{\nu}+\sigma^2(T-t)\right)}{-\sigma\sqrt{T-t}}}{\sqrt{1+\left(\sqrt{\psi}/(-\sigma\sqrt{T-t})\right)^2}} \right) \right] \\
&+ e^{-\delta(T-t)} e^{(2m/\sigma^2+2)\underline{\nu}} \frac{L_1}{L_0} \sqrt{2\pi\psi} e^{M_1(-2m/\sigma^2-1)+\psi(-2m/\sigma^2-1)^2/2} \\
&\times \Phi_2 \left(\frac{\frac{-y^* + (M_1+\psi(-2m/\sigma^2-1)) + \left(m(T-t)+2\underline{\nu}+\sigma^2(T-t)\right)}{-\sigma\sqrt{T-t}}}{\sqrt{1+\left(\sqrt{\psi}/(-\sigma\sqrt{T-t})\right)^2}}, \frac{\underline{\nu} - (M_1+\psi(-2m/\sigma^2-1))}{\sqrt{\psi}}; - \left(\frac{\sqrt{\psi}/(-\sigma\sqrt{T-t})}{\sqrt{1+\left(\sqrt{\psi}/(-\sigma\sqrt{T-t})\right)^2}} \right) \right) \\
&+ e^{-\delta(T-t)} e^{(2m/\sigma^2+2)(\underline{\nu})} \frac{L_2}{L_0} \sqrt{2\pi\psi} e^{M_2(-2m/\sigma^2-1)+\psi(-2m/\sigma^2-1)^2/2} \times \left[\Phi \left(\frac{\frac{-y^* + (M_2+\psi(-2m/\sigma^2-1)) + \left(m(T-t)+2\underline{\nu}+\sigma^2(T-t)\right)}{-\sigma\sqrt{T-t}}}{\sqrt{1+\left(\sqrt{\psi}/(-\sigma\sqrt{T-t})\right)^2}} \right) \right]
\end{aligned}
$$

$$
\begin{aligned}
& - e^{-\delta(T-t)} e^{(2m/\sigma^2+2)(\underline{\nu})} \frac{L_2}{L_0} \sqrt{2\pi\psi} e^{M_2(-2m/\sigma^2-1)+\psi(-2m/\sigma^2-1)^2/2} \\
& \times \left[\Phi_2 \left(\frac{\frac{-y^* + (M_2+\psi(-2m/\sigma^2-1)) + \left(m(T-t)+2\underline{\nu}+\sigma^2(T-t)\right)}{-\sigma\sqrt{T-t}}}{\sqrt{1+\left(\sqrt{\psi}/(-\sigma\sqrt{T-t})\right)^2}}, \frac{\underline{\nu} - (M_2+\psi(-2m/\sigma^2-1)^2)}{\sqrt{\psi}}; - \left(\frac{-\sqrt{\psi}/(\sigma\sqrt{T-t})}{\sqrt{1+\left(\sqrt{\psi}/(-\sigma\sqrt{T-t})\right)^2}} \right) \right) \right]
\end{aligned} \tag{C.33}
$$

For evaluating the second integral in equation C.30 we make use of the result from equation C.29

$$
\begin{aligned}
&\int_{\underline{\nu}}^{\infty} Be^{-\gamma(u-\underline{\nu})} \left[\Phi\left(-z_3\right) - e^{(2m/\sigma^2-2\gamma)(\underline{\nu}-u)}\Phi\left(-z_4\right)\right] g\left(u|z,y\right) du \\
&= \int_{\underline{\nu}}^{\infty} Be^{-\gamma(u-\underline{\nu})} \left[\Phi\left(-z_3\right) - e^{(2m/\sigma^2-2\gamma)(\underline{\nu}-u)}\Phi\left(-z_4\right)\right] \left[\frac{L_1}{L_0} \times e^{\frac{-(u-M_1)^2}{2\psi}} - \frac{L_2}{L_0} \times e^{\frac{-(u-M_2)^2}{2\psi}}\right] du \\
&= \int_{\underline{\nu}}^{\infty} Be^{-\gamma(u-\underline{\nu})}\Phi\left(-z_3\right) \left[\frac{L_1}{L_0} \times e^{\frac{-(u-M_1)^2}{2\psi}} - \frac{L_2}{L_0} \times e^{\frac{-(u-M_2)^2}{2\psi}}\right] du \\
&- \int_{\underline{\nu}}^{\infty} Be^{-\gamma(u-\underline{\nu})}e^{(2m/\sigma^2-2\gamma)(\underline{\nu}-u)}\Phi\left(-z_4\right) \left[\frac{L_1}{L_0} \times e^{\frac{-(u-M_1)^2}{2\psi}} - \frac{L_2}{L_0} \times e^{\frac{-(u-M_2)^2}{2\psi}}\right] du \\
&= \int_{\underline{\nu}}^{\infty} Be^{-\gamma(u-\underline{\nu})}\Phi\left(-z_3\right) \frac{L_1}{L_0} \times e^{\frac{-(u-M_1)^2}{2\psi}} du - \int_{\underline{\nu}}^{\infty} Be^{-\gamma(u-\underline{\nu})}\Phi\left(-z_3\right) \frac{L_2}{L_0} \times e^{\frac{-(u-M_2)^2}{2\psi}} du \\
&- \frac{L_1}{L_0} \times \int_{\underline{\nu}}^{\infty} Be^{-\gamma(u-\underline{\nu})}e^{(2m/\sigma^2-2\gamma)(\underline{\nu}-u)}\Phi\left(-z_4\right) e^{\frac{-(u-M_1)^2}{2\psi}} du \\
&+ \int_{\underline{\nu}}^{\infty} Be^{-\gamma(u-\underline{\nu})}e^{(2m/\sigma^2-2\gamma)(\underline{\nu}-u)}\Phi\left(-z_4\right) \frac{L_2}{L_0} \times e^{\frac{-(u-M_2)^2}{2\psi}} du
\end{aligned}
\tag{C.34}
$$

$$
\begin{aligned}
&= \frac{L_1}{L_0} Be^{\gamma(\underline{\nu})} \int_{\underline{\nu}}^{\infty} e^{-\gamma(u)}\Phi\left(-z_3\right) e^{\frac{-(u-M_1)^2}{2\psi}} du - \frac{L_2}{L_0} Be^{\gamma(\underline{\nu})} \int_{\underline{\nu}}^{\infty} e^{-\gamma(u)}\Phi\left(-z_3\right) e^{\frac{-(u-M_2)^2}{2\psi}} du \\
&- \frac{L_1}{L_0} Be^{\gamma(\underline{\nu})} e^{(2m/\sigma^2-2\gamma)(\underline{\nu})} \int_{\underline{\nu}}^{\infty} e^{-(2m/\sigma^2-\gamma)(u)}\Phi\left(-z_4\right) e^{\frac{-(u-M_1)^2}{2\psi}} du \\
&+ \frac{L_2}{L_0} Be^{\gamma(\underline{\nu})} e^{-(2m/\sigma^2-2\gamma)\underline{\nu}} \int_{\underline{\nu}}^{\infty} e^{-(2m/\sigma^2-\gamma)(u)}\Phi\left(-z_4\right) e^{\frac{-(u-M_2)^2}{2\psi}} du \\
&+ \frac{L_1}{L_0} Be^{\gamma(\underline{\nu})} \sqrt{2\pi\psi} e^{(-\gamma)M_1+(-\gamma)^2\psi/2} \times \left[\Phi\left(\frac{\frac{-y^*+(M_1+\psi(-\gamma))+\left(m(T-t)-\gamma\sigma^2(T-t)\right)}{\sigma\sqrt{T-t}}}{\sqrt{1+\left(\sqrt{\psi}/(\sigma\sqrt{T-t})\right)^2}}\right)\right] \\
&- \frac{L_1}{L_0} Be^{\gamma(\underline{\nu})} \sqrt{2\pi\psi} e^{(-\gamma)M_1+(-\gamma)^2\psi/2} \\
&\times \left[\Phi_2\left(\frac{\frac{-y^*+(M_1+\psi(-\gamma))+\left(m(T-t)-\gamma\sigma^2(T-t)\right)}{\sigma\sqrt{T-t}}}{\sqrt{1+\left(\sqrt{\psi}/(\sigma\sqrt{T-t})\right)^2}}, \frac{\underline{\nu}-(M_1+\psi(-\gamma))}{\sqrt{\psi}}; \left(\frac{-\sqrt{\psi}/(\sigma\sqrt{T-t})}{\sqrt{1+\left(\sqrt{\psi}/(\sigma\sqrt{T-t})\right)^2}}\right)\right)\right] \\
&- \frac{L_2}{L_0} Be^{\gamma(\underline{\nu})} \sqrt{2\pi\psi} e^{M_2(-\gamma)+(-\gamma)^2\psi/2} \times \Phi\left(\frac{\frac{-y^*+M_2+(\psi(-\gamma))+\left(m(T-t)-\gamma\sigma^2(T-t)\right)}{\sigma\sqrt{T-t}}}{\sqrt{1+\left(\sqrt{\psi}/(\sigma\sqrt{T-t})\right)^2}}\right)
\end{aligned}
$$

$$
\begin{aligned}
&+\frac{L_2}{L_0}Be^{\gamma(\underline{\nu})}\sqrt{2\pi\psi}e^{M_2(-\gamma)+(-\gamma)^2\psi/2}\\
&\times\left[\Phi_2\left(\frac{\frac{-y^*+(M_2+\psi(-\gamma))+\left(m(T-t)-\gamma\sigma^2(T-t)\right)}{\sigma\sqrt{T-t}}}{\sqrt{1+\left(\sqrt{\psi}/(\sigma\sqrt{T-t})\right)^2}},\frac{\underline{\nu}-(M_2+\psi(-\gamma))}{\sqrt{\psi}};\left(\frac{-(\sqrt{\psi})/(\sigma\sqrt{T-t})}{\sqrt{1+\left(\sqrt{\psi}/(\sigma\sqrt{T-t})\right)^2}}\right)\right)\right]\\
&-\frac{L_1}{L_0}Be^{\gamma(\underline{\nu})}e^{(2m/\sigma^2-2\gamma)(\underline{\nu})}\sqrt{2\pi\psi}e^{M_1(-2m/\sigma^2+\gamma)+\psi(-2m/\sigma^2+\gamma)^2/2}\\
&\times\left[\Phi\left(\frac{\frac{-y^*+(M_1+(\psi(-2m/\sigma^2+\gamma)))+\left(m(T-t)+2\underline{\nu}-\gamma\sigma^2(T-t)\right)}{-\sigma\sqrt{T-t}}}{\sqrt{1+\left(\sqrt{\psi}/(-\sigma\sqrt{T-t})\right)^2}}\right)\right]\\
&+\frac{L_1}{L_0}Be^{\gamma(\underline{\nu})}e^{(2m/\sigma^2-2\gamma)(\underline{\nu})}\sqrt{2\pi\psi}e^{M_1(-2m/\sigma^2+\gamma)+\psi(-2m/\sigma^2+\gamma)^2/2}\\
&\times\Phi_2\left(\frac{\frac{-y^*+(M_1+(\psi(-2m/\sigma^2+\gamma))+\left(m(T-t)+2\underline{\nu}-\gamma\sigma^2(T-t)\right)}{-\sigma\sqrt{T-t}}}{\sqrt{1+\left(\sqrt{\psi}/(-\sigma\sqrt{T-t})\right)^2}},\frac{\underline{\nu}-(M_1+\psi(-2m/\sigma^2+\gamma))}{\sqrt{\psi}};-\left(\frac{-\sqrt{\psi}/(\sigma\sqrt{T-t})}{\sqrt{1+\left(\sqrt{\psi}/(-\sigma\sqrt{T-t})\right)^2}}\right)\right)\\
&+\frac{L_2}{L_0}Be^{\gamma(\underline{\nu})}e^{-(2m/\sigma^2-2\gamma)\underline{\nu}}\sqrt{2\pi\psi}e^{M_2(-2m/\sigma^2+\gamma)+\psi(-2m/\sigma^2+\gamma)^2/2}\\
&\times\left[\Phi\left(\frac{\frac{-y^*+(M_2+(\psi(-2m/\sigma^2+\gamma)))+\left(m(T-t)+2\underline{\nu}-\gamma\sigma^2(T-t)\right)}{-\sigma\sqrt{T-t}}}{\sqrt{1+\left(\sqrt{\psi}/(-\sigma\sqrt{T-t})\right)^2}}\right)-\right]\\
&-\frac{L_2}{L_0}Be^{\gamma(\underline{\nu})}e^{-(2m/\sigma^2-2\gamma)\underline{\nu}}\sqrt{2\pi\psi}e^{M_2(-2m/\sigma^2+\gamma)+\psi(-2m/\sigma^2+\gamma)^2/2}\\
&\times\Phi_2\left(\frac{\frac{-y^*+(M_2+\psi(-2m/\sigma^2+\gamma))+\left(m(T-t)+2\underline{\nu}-\gamma\sigma^2(T-t)\right)}{-\sigma\sqrt{T-t}}}{\sqrt{1+\left(\sqrt{\psi}/(-\sigma\sqrt{T-t})\right)^2}},\frac{\underline{\nu}-(M_2+\psi(-2m/\sigma^2+\gamma))}{\sqrt{\psi}};-\left(\frac{-\sqrt{\psi}/(-\sigma\sqrt{T-t})}{\sqrt{1+\left(\sqrt{\psi}/(-\sigma\sqrt{T-t})\right)^2}}\right)\right)
\end{aligned}
\tag{C.35}
$$

For evaluating the third integral in equation C.30 the result of equation C.29

$$
\begin{aligned}
&-\int_{\underline{\nu}}^{\infty}e^{-r(T-t)}\left(A+K\right)\left[\Phi\left(-z_5\right)-e^{(2m/\sigma^2)(\underline{\nu}-u)}\Phi\left(-z_6\right)\right]g\left(u|z,y\right)du\\
&=-\int_{\underline{\nu}}^{\infty}e^{-r(T-t)}\left(A+K\right)\left[\Phi\left(-z_5\right)-e^{(2m/\sigma^2)(\underline{\nu}-u)}\Phi\left(-z_6\right)\right]\left[\frac{L_1}{L_0}\times e^{\frac{-(u-M_1)^2}{2\psi}}-\frac{L_2}{L_0}\times e^{\frac{-(u-M_2)^2}{2\psi}}\right]du\\
&=-\int_{\underline{\nu}}^{\infty}e^{-r(T-t)}\left(A+K\right)\Phi\left(-z_5\right)\left[\frac{L_1}{L_0}\times e^{\frac{-(u-M_1)^2}{2\psi}}-\frac{L_2}{L_0}\times e^{\frac{-(u-M_2)^2}{2\psi}}\right]du\\
&+\int_{\underline{\nu}}^{\infty}e^{-r(T-t)}\left(A+K\right)e^{(2m/\sigma^2)(\underline{\nu}-u)}\Phi\left(-z_6\right)\left[\frac{L_1}{L_0}\times e^{\frac{-(u-M_1)^2}{2\psi}}-\frac{L_2}{L_0}\times e^{\frac{-(u-M_1)^2}{2\psi}}\right]du\\
&=-\int_{\underline{\nu}}^{\infty}e^{-r(T-t)}\left(A+K\right)\Phi\left(-z_5\right)\frac{L_1}{L_0}e^{\frac{-(u-M_1)^2}{2\psi}}du+\int_{\underline{\nu}}^{\infty}e^{-r(T-t)}\left(A+K\right)\Phi\left(-z_5\right)\frac{L_2}{L_0}e^{\frac{-(u-M_2)^2}{2\psi}}du
\end{aligned}
$$

$$
\begin{aligned}
&+\int_{\underline{\nu}}^{\infty} e^{-r(T-t)}\left(A+K\right) e^{(2m/\sigma^2)(\underline{\nu}-u)}\Phi\left(-z_6\right)\frac{L_1}{L_0} e^{\frac{-(u-M_1)^2}{2\psi}}\,du-\int_{\underline{\nu}}^{\infty} e^{-r(T-t)}\left(A+K\right) e^{(2m/\sigma^2)(\underline{\nu}-u)}\Phi\left(-z_6\right)\frac{L_2}{L_0} e^{\frac{-(u-M_2)^2}{2\psi}}\,du \\
&=-e^{-r(T-t)}\left(A+K\right)\frac{L_1}{L_0}\int_{\underline{\nu}}^{\infty}\Phi\left(-z_5\right) e^{\frac{-(u-M_1)^2}{2\psi}}\,du+e^{-r(T-t)}\left(A+K\right)\frac{L_2}{L_0}\int_{\underline{\nu}}^{\infty}\Phi\left(-z_5\right) e^{\frac{-(u-M_1)^2}{2\psi}}\,du \\
&+e^{-r(T-t)}\left(A+K\right)\frac{L_1}{L_0} e^{(2m/\sigma^2)(\underline{\nu})}\int_{\underline{\nu}}^{\infty} e^{-(2m/\sigma^2)(u)}\Phi\left(-z_6\right) e^{\frac{-(u-M_1)^2}{2\psi}}\,du \\
&-e^{-r(T-t)}\left(A+K\right) e^{(2m/\sigma^2)(\underline{\nu})}\frac{L_2}{L_0}\int_{\underline{\nu}}^{\infty} e^{-(2m/\sigma^2)(u)}\Phi\left(-z_6\right) e^{\frac{-(u-M_2)^2}{2\psi}}\,du
\end{aligned}
\tag{C.36}
$$

$$
\begin{aligned}
&= -e^{-r(T-t)}\,(A+K)\,\frac{L_1}{L_0}\sqrt{2\pi\psi}\left[\Phi\left(\frac{\frac{-y^*+(m(T-t)+M_1)}{\sigma\sqrt{T-t}}}{\sqrt{1+\left(\sqrt{\psi}/(\sigma\sqrt{T-t})\right)^2}}\right)-\Phi_2\left(\frac{\frac{-y^*+(m(T-t)+M_1)}{\sigma\sqrt{T-t}}}{\sqrt{1+\left(\sqrt{\psi}/(\sigma\sqrt{T-t})\right)^2}},\frac{\underline{\nu}-M_1}{\sqrt{\psi}};\left(\frac{-\sqrt{\psi}/(\sigma\sqrt{T-t})}{\sqrt{1+\left(\sqrt{\psi}/(\sigma\sqrt{T-t})\right)^2}}\right)\right)\right]\\
&+ e^{-r(T-t)}\,(A+K)\,\frac{L_2}{L_0}\sqrt{2\pi\psi}\left[\Phi\left(\frac{\frac{-y^*+(m(T-t)+M_2)}{\sigma\sqrt{T-t}}}{\sqrt{1+\left(\sqrt{\psi}/(\sigma\sqrt{T-t})\right)^2}}\right)-\Phi_2\left(\frac{\frac{-y^*+(m(T-t)+M_2)}{\sigma\sqrt{T-t}}}{\sqrt{1+\left(\sqrt{\psi}/(\sigma\sqrt{T-t})\right)^2}},\frac{\underline{\nu}-M_2}{\sqrt{\psi}};\left(\frac{-\sqrt{\psi}/(\sigma\sqrt{T-t})}{\sqrt{1+\left(\sqrt{\psi}/(\sigma\sqrt{T-t})\right)^2}}\right)\right)\right]\\
&+ e^{-r(T-t)}\,(A+K)\,\frac{L_1}{L_0}e^{(2m/\sigma^2)(\underline{\nu})}\sqrt{2\pi\psi}e^{M_1(-2m/\sigma^2)+\psi(-2m/\sigma^2)^2/2}\\
&\times\left[\Phi\left(\frac{\frac{-y^*+(M_1+(\psi(-2m/\sigma^2)))+(m(T-t)+2\underline{\nu})}{-\sigma\sqrt{T-t}}}{\sqrt{1+\left(\sqrt{\psi}/(-\sigma\sqrt{T-t})\right)^2}}\right)-\Phi_2\left(\frac{\frac{-y^*+(M_1+\psi(-2m/\sigma^2)^2)+(m(T-t)+2\underline{\nu})}{-\sigma\sqrt{T-t}}}{\sqrt{1+\left(\sqrt{\psi}/(-\sigma\sqrt{T-t})\right)^2}},\frac{\underline{\nu}-(M_1+(\psi(-2m/\sigma^2)))}{\sqrt{\psi}};-\left(\frac{-\sqrt{\psi}/(-\sigma\sqrt{T-t})}{\sqrt{1+\left(\sqrt{\psi}/(-\sigma\sqrt{T-t})\right)^2}}\right)\right)\right]\\
&- e^{-r(T-t)}\,(A+K)\,\frac{L_2}{L_0}e^{(2m/\sigma^2)(\underline{\nu})}\sqrt{2\pi\psi}e^{M_2(-2m/\sigma^2)+\psi(-2m/\sigma^2)^2/2}\\
&\times\left[\Phi\left(\frac{\frac{-y^*+(M_2+(\psi(-2m/\sigma^2)))+(m(T-t)+2\underline{\nu})}{-\sigma\sqrt{T-t}}}{\sqrt{1+\left(\sqrt{\psi}/(-\sigma\sqrt{T-t})\right)^2}}\right)-\Phi_2\left(\frac{\frac{-y^*+(M_2+\psi(-2m/\sigma^2))+(m(T-t)+2\underline{\nu})}{-\sigma\sqrt{T-t}}}{\sqrt{1+\left(\sqrt{\psi}/(\sigma\sqrt{T-t})\right)^2}},\frac{\underline{\nu}-(M_2+(\psi(-2m/\sigma^2)))}{\sqrt{\psi}};-\left(\frac{-\sqrt{\psi}/(\sigma\sqrt{T-t})}{\sqrt{1+\left(\sqrt{\psi}/(-\sigma\sqrt{T-t})\right)^2}}\right)\right)\right]
\end{aligned}
\tag{C.37}
$$

References

F. Akbas, C. Jiang, and P. D. Koch. The trend in firm profitability and the cross-section of stock returns. *The Accounting Review*, 92(5):1–32, 2017.

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

B. J. E. An, A. Ang, T. G. Bali, and N. Cakici. The joint cross section of stocks and options. *Journal of Finance*, 69(5):2279–2337, 2010.

T. G. Andersen, N. Fusari, and V. Todorov. Parametric inference and dynamic state recovery from option panels. *Econometrica*, 83(3):1081–1145, 2015.

M. J. Anderson, B. J. Henderson, and N. D. Pearson. Bond lending and bond returns. *SSRN Working Paper*, 1:1–58, 2018.

A. Andrikogiannopoulou and F. Papakonstantinou. Reassessing false discoveries in mutual fund performance: Skill, luck, or lack of power? *Journal of Finance*, 74(5):2667–2688, 2019.

A. Ang. *Asset Management: A Systematic Approach to Factor Investing.* Oxford University Press, New York, NY, 2014.

A. Ang, R. J. Hodrick, Y. Xing, and X. Zhang. The crosssection of volatility and expected returns. *Journal of Finance*, 61(1):259–299, 2006.

C. Asness, T. Moskowitz, and L. H. Pedersen. Value and momentum everywhere. *Journal of Finance*, 68(3):929–985, 2013.

P. Asquith and E. Kim. The impact of merger bids on the participating firms security holders. *Journal of Finance*, 37(2):1209–1228, 1982.

P. Asquith and T. Wizman. Event risk, covenants, and bondholder returns in leveraged buyouts. *Journal of Financial Economics*, 27(1):195–213, 1990.

D. Avramov, T. Chordia, G. Jostova, and A. Philipov. Bonds, stocks, and sources of mispricing. *Working Paper*, 1:1–78, 2019.

J. Bai, Subrahmanyam. A., and Q. Wen. Long-term reversals in the corporate bond market. *Journal of Financial Economics (FORTHCOMING)*, 0(0):1–72, 2019a.

J. Bai, T. G. Bali, and Q. Wen. Common risk factors in the cross-section of corporate bond returns. *Journal of Financial Economics*, 131(3):619–642, 2019b.

J. Bai, Subrahmanyam. A., and Q. Wen. The macroeconomic uncertainty premium in the corporate bond market. *Journal of Financial and Quantitative Analysis (FORTHCOMING)*, 0(0):1–51, 2020a.

J. Bai, T. G. Bali, and Q. Wen. Do the distributional characteristics of corporate bonds predict their future returns? *SSRN Working Paper*, 1:1–67, 2020b.

J. Bai, T. G. Bali, and Q. Wen. In search of the idiosyncratic volatility puzzle in the corporate bond market. *SSRN Working Paper*, 1:1–61, 2020c.

S. R. Baker, N. Bloom, S. Davis, and M. Sammon. What triggers stock market jumps. *Working Paper Stanford University*, 1:1–79, 2020.

T. G. Bali, A. Goyal, D. Huang, F. Jiang, and Q. Wen. The cross-sectional pricing of corporate bonds using big data and machine learning. *Working paper, Georgetown University*, 1:1–62, 2020.

J. Bao and K. Hou. De facto seniority, credit risk, and corporate bond prices. *Review of Financial Studies*, 30(11):4038–4080, 2017.

J. Bao, J. Pan, and J. Wang. The illiquidity of corporate bonds. *Journal of Finance*, 66 (3):911–946, 2011.

L. Barras, O. Scaillet, and Wermers. False discoveries in mutual fund performance: Measuring luck in estimated alphas. *Journal of Finance*, 65(1):179–216, 2010.

S. Bartram, M. Grinblatt, and Y. Nozawa. Book-to-market, mispricing, and the cross-section of corporate bond returns. *NBER Working Paper*, 1:1–58, 2020.

D. S. Bates. Jumps and stochastic volatility: Exchange rate processes implicit in the deutsch mark options. *Review of Financial Studies*, 9(1):69–107, 1996.

R. Battalio and P. Schultz. Regulatory uncertainty and market liquidity: The 2008 short sale ban's impact on equity option markets. *Journal of Finance*, 66(6):2013–2053, 2011.

B. Becker and V. Ivanisha. Reaching for yield in the bond market. *Journal of Finance*, 70 (1):1863–1901, 2015.

Y. Benjamini and Y. Hochberg. Controlling the false discovery rate: A practical and powerful approach to multiple testing. *Journal of the Royal Statistical Society Series B (Methodological)*, 57(1):289–300, 1995.

Y. Benjamini and D. Yekutieli. The control of the false discovery rate in multiple testing under dependency. *Annals of Statistics*, 29(1):1165–1188, 2001.

H. Bessembinder, K. M. Kahle, W. F. Maxwell, and D. Xu. Measuring abnormal bond performance. *Review of Financial Studies*, 22(10):4219–4258, 2009.

S. T. Bharath and T. Shumway. Forecasting default with the Merton distance to default model. *Review of Financial Studies*, 21(3):1339–1369, 2008.

S. Bhojraj and B. Swaminathan. How does the corporate bond market value capital investments and accruals? *Review of Accounting Studies*, 14(1):31–62, 2009.

M. Billet, D. King, and D. Mauer. Bondholder wealth effects in mergers and acquisitions: New evidence from the 1980s and 1990s. *Journal of Finance*, 59(1):107–135, 2004.

F. Black. Fact and fantasy in the use of options. *Financial Analysts Journal*, 31(4):36–72, 1975.

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

N. Bollen and R. Whaley. Does net buying pressure affect the shape of implied volatility functions? *Journal of Finance*, 59(2):711–753, 2004.

C. E. Bonferroni. Teoria statistica delle classi e calcolo delle probabilita. *Libreria Internazionale Seeber*, 1(1):0–0, 1936.

J. Boudoukh, R. Michaely, M. Richardson, and M. R. Roberts. On the importance of measuring payout yield: Implications for asset pricing. *Journal of Finance*, 62(2):877–915, 2007.

M. J. Brennan, T. Chordia, and A. Subrahmanyam. Alternative factor specifications, security characteristics, and the cross-section of expected stock returns. *Journal of Financial Economics*, 49(3):345–373, 1998.

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

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

A. Buraschi and A. Jiltsov. Model uncertainty and option markets with heterogeneous beliefs. *Journal of Finance*, 61(1):2841–2897, 2006.

M. Cahart. On persistence in mutual fund performance. *Journal of Finance*, 52(1):57–82, 1997.

J. Y. Campbell and R. J. Schiller. Stock prices, earnings, and expected dividends. *Journal of Finance*, 43(3):661–676, 1988.

J. Y. Campbell and S. B. Thompson. Predicting excess stock returns out of sample: Can anything beat the historical average? *Review of Financial Studies*, 21(4):1509–1531, 2008.

M. Campello, L. Chen, and L. Zhang. Expected returns, yield spreads, and asset pricing tests. *Review of Financial Studies*, 21(3):1297–1338, 2008.

J. Cao, A. Goyal, X. Xiao, and X. Zhang. Implied volatility changes and corporate bond returns. *SSRN Working Paper*, 1:1–61, 2020.

U. Cetin, R. Jarrow, R. Protter, and M. Warachka. Pricing options in an extended black scholes economy with illiquidity: Theory and empirical evidence. *Review of Financial Studies*, 19(2):493–529, 2006.

S. H. Chakravarty, H. Gulen, and S. Mayhew. Informed trading in stock and option markets. *Journal of Finance*, 59(3):1235–1258, 2005.

K. E. Chandia. *Option Prices and Accounting Choices*. PhD thesis, Rutgers the State University of New Jersey, 2014.

D. Chen, B. Guo, and G. Zhou. Firm fundamentals and the cross section of implied volatility shapes. *Working Paper*, 1:1–45, 2020.

L. Chen, Z. Da, and X. Zhao. What drives stock price movements? *Review of Financial Studies*, 26(4):841–876, 2013.

D. Chichernea, A. Holder, and A. Petkevich. Decomposing the accrual premium: The evidence from two markets. *Journal of Business Finance and Accounting*, 46(7):879–912, 2019.

J. Choi and C. Kim. Anomalies and market disintegration. *Journal of Monetary Economics*, 100(0):16–34, 2018.

T. Chordia, A. Goyal, Y. Nozawa, A. Subrahmanyam, and Q. Tong. Are capital market anomalies common to equity and corporate bond markets? an empirical investigation. *Journal of Financial and Quantitative Analysis*, 52(4):1301–1342, 2017.

T. Chordia, A. Goyal, and A. Saretto. Anomalies and false rejection. *Review of Financial Studies*, 33(5):2134–2179, 2020.

P. Christoffersen, K. Jacobs, and C. Ornthanalai. Dynamic jump intensities and risk premia: Evidence from sp500 returns and options. *Journal of Financial Economics*, 106 (3):447–472, 2012.

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

K. H. Chung, J. Wang, and C. Wu. Volatility and the cross-section of corporate bond returns. *Journal of Financial Economics*, 113(1):397–417, 2019.

T. E. Clark and K. D. West. Approximately normal tests for equal predictive accuracy in nested models. *Journal of Econometrics*, 38(4):291–311, 2007.

J. H. Cochrane. Presidential address: Discount rates. *Journal of Finance*, 66(4):1047–1108, 2011.

J. Conrad, M. Cooper, and G. Kaul. Value versus glamour. *Journal of Finance*, 58(5): 1969–1996, 2003.

D. Cook and J. Easterwood. Poison put bonds: An analysis of their economic role. *Journal of Finance*, 49(5):1905–1920, 1994.

M. J. Cooper and H. Gulen. Is time-series-based predictibility evident in real time? *Journal of Business*, 79(3):1263–1292, 2006.

J. Cox and S. Ross. The valuation of options for alternative stochastic processes. *Journal of Financial Economics*, 3(0):145–166, 1976.

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

L. Crabbe. Event risk: An analysis of losses to bondholders and ”super poison put” bond covenants. *Journal of Finance*, 46(2):689–706, 1991.

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

M. Cremers, R. Y. Goyenko, P. Schultz, and S. Szaura. Informed trading of options, expiration risk, and future stock returns. *Notre Dame and McGill University Working Paper*, 1:1–54, 2019.

L. Dann. Common stock repurchases: An analysis of returns to bondholders and stockholders. *Journal of Financial Economics*, 9(2):113–138, 1981.

M. L. Defond and J. Zhang. The timeliness of the bond market reaction to bad earnings news. *Contemporary Accounting Research*, 31(3):911–936, 2014.

R. Defusco, R. Johnson, and T. Zorn. The effect of executive stock option plans on stockholders and bondholders. *Journal of Finance*, 45(2):617–627, 1990.

D. Dennis and J. McConnell. Corporate mergers and security returns. *Journal of Financial Economics*, 16(2):143–187, 1986.

P. Dennis and S. Mayhew. Risk-neutral skewness: Evidence from stock options. *Journal of Financial and Quantitative Analysis*, 37(3):471–493, 2002.

E. Derman and I. Kani. Riding on the smile. *Risk*, 7(0):32–39, 1994.

U. Dhillon and H. Johnson. The effect of dividend changes on stock and bond prices. *Journal of Finance*, 49(1):281–289, 1994.

J. Dick-Nielsen. Liquidity biases in TRACE. *Journal of Fixed Income*, 19(2):1–28, 2009.

J. Dick-Nielsen. How to clean Enhanced TRACE data. *Working Paper, Copenhagen Business School*, 1:1–28, 2012.

F. X. Diebold and R. S. Mariano. Comparing predictive accuracy. *Journal of Business and Economic Statistics*, 13(405):134–144, 1995.

K. B. Diether, C. J. Malloy, and A. Scherbina. Differences of opinion and the cross section of stock returns. *Journal of Finance*, 57(5):2113–2141, 2002.

E. Dimson. Risk measurement when shares are subject to infrequent trading. *Journal of Financial Economics*, 7(2):197–226, 1979.

H. Doshi, J. Ericsson, S. Szaura, and F. Yu. Accounting transparency and the implied volatility skew. *Working Paper McGill University*, 1:1–41, 2020.

D. Du, R. Elkhami, and J. Ericsson. Time varying volatility and the credit spread puzzle. *Journal of Finance (FORTHCOMING)*, 0(0):1–81, 2018.

D. Du, R. Elkhami, and J. Ericsson. Time varying asset volatility and the credit spread puzzle. *Journal of Finance*, 74(4):1841–1885, 2019.

A. Dubinsky, M. Johannes, A. Kaeck, and N. J. Seeger. Option pricing of earnings announcement risks. *Review of Financial Studies*, 32(2):646–687, 2019.

D. Duffie and D. Lando. Term structure of credit spreads with incomplete accounting information. *Econometrica*, 69(3):633–664, 2001.

D. Duffie, J. Pan, and K. Singleton. Transform analysis and asset pricing for affine jump diffusions. *Econometrica*, 68(6):1343–1376, 2000.

B. Dupire. Pricing with a smile. *Risk*, 7(0):18–20, 1994.

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

P. D. Easton, S. J. Monahan, and F. P. Vasvari. Initial evidence on the role of accounting earnings in the bond market. *Journal of Accounting Research*, 47(3):721–766, 2009.

A. Eberhart and A. Siddique. The long-term performance of corporate bonds (and stocks) following seasoned equity offerings. *Review of Financial Studies*, 15(5):1385–1406, 2002.

C. Eger. An empirical test of the redistribution effect in pure exchange mergers. *Journal of Financial and Quantitative Analysis*, 18(4):547–572, 1983.

R. Elkamhi and J. Ericsson. Time varying risk premia in corporate bond market. *Working Paper McGill University*, 1:1–50, 2008.

E. J. Elton, M. J. Gruber, and C. R. Blake. Fundamental economic variables, expected returns, and bond fund performance. *Journal of Finance*, 50(4):1229–1256, 1995.

B. Eraker, M. Johannes, and N. Polson. The impact of jumps in volatility and returns. *Journal of Finance*, 58(3):1269–1300, 2003.

E. F. Fama and K. R. French. Business conditions and expected returns on stock and bonds. *Journal of Financial Economics*, 25(1):23–49, 1989.

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

E. F. Fama and K. R. French. Dissecting anomalies. *Journal of Finance*, 63(4):1653–1678, 2008.

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

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

W. E. Ferson and C. R. Harvey. The variation of economic risk premiums. *Journal of Political Economy*, 99(2):385–415, 1991.

S. Figlewski and G. P. Webb. Options, short sales, and market completeness. *Journal of Finance*, 48(2):761–777, 1993.

D. Foster, T. Smith, and R. E. Whaley. Assessing goodness of fit of asset pricing models: The distribution of the maximal r2. *Journal of Finance*, 52(2):591–607, 1997.

M. Fournier and K. Jacobs. A tractable framework for option pricing with dynamic market maker inventory and wealth, with k. jacobs. *Journal of Finance and Quantitative Analysis (FORTHCOMING)*, 0(0):0–0, 2020.

A. Gamba and A. Saretto. Endogenous option pricing. *Working Paper*, 1:1–38, 2020.

N. Garleanu, L. Pedersen, and A. Poteshman. Demand-based option pricing. *Review of Financial Studies*, 22(10):4259–4299, 2009.

L. Ge, T.-c. Lin, and N. D. Pearson. Why does the option to stock volume ratio predict stock returns? *Journal of Financial Economics*, 120(3):601–622, 2016.

W. R. Gebhardt, S. Hvidkjaer, and B. Swaminathan. The cross-section of expected corporate bond returns: Betas or characteristics? *Journal of Financial Economics*, 75(1): 85–114, 2005a.

W. R. Gebhardt, S. Hvidkjaer, and B. Swaminathan. Stock and bond market interaction: Does momentum spill over? *Journal of Financial Economics*, 75(1):651–690, 2005b.

R. Geske. The valuation of compound options. *Journal of Financial Economics*, 7(1): 63–81, 1979.

R. Geske, A. Subrahmanyam, and Y. Zhou. Capital structure effects on the prices of equity call options. *Journal of Financial Economics*, 121(2):231–253, 2016.

E. Gettleman and J. M. Marks. Acceleration strategies. *Working paper, University of Illinois at. Urbana-Champaign*, 1:1–50, 2006.

L. R. Glosten, R. Jagannathan, and D. E. Runkle. On the relation between the expected value and the volatility of the nominal excess return on stocks. *Journal of Finance*, 48 (5):1779–1801, 1993.

J. Goldberg and Y. Nozawa. Liquidity supply in the corporate bond market. *Journal of Finance (Forthcoming)*, 0(0):1–76, 2019.

J. Green, J. R. M. Hand, and X. F. Zhang. The characteristics that provide independent information about average u.s. monthly stock returns. *Review of Financial Studies*, 30 (12):4389–4436, 2017.

S. Gu, B. Kelly, and D. Xiu. Empirical asset pricing via machine learning. *Review of Financial Studies*, 0(0):1–51, 2020.

J. Hand, R. Holthausen, and R. Leftwich. The effect of bond rating agency announcements on bond and stock prices. *Journal of Finance*, 47(2):733–752, 1992.

G. Handjinicolaou and A. Kalay. Wealth redistributions or changes in firm value: An analysis of returns to bondholders and stockholders around dividend announcements. *Journal of Financial Economics*, 13(1):35–63, 1984.

C. R. Harvey and Y. Liu. A census of the factor zoo. *Working Paper*, 1:1–7, 2019.

C. R. Harvey and Y. Liu. False (and missed) discoveries in financial economics. *Journal of Finance (FORTHCOMING)*, 0(0):1–82, 2020.

C. R. Harvey, Y. Liu, and H. Zhu. ... and the cross section of expected returns. *Review of Financial Studies*, 29(1):5–68, 2016.

C. R. Harvey, Y. Liu, and A. Saretto. An evaluation of alternative multiple testing methods for finance applications. *Review of Asset Pricing Studies*, 10(2):199–248, 2020.

E. H. Hawkins, S. C. Chamberlin, and W. E. Daniel. Earnings expectations and security prices. *Financial Analyst Journal*, 40(5):24–48, 1984.

Z. He, B. Kelly, and A. Manella. Intermediary asset pricing: New evidence from many asset classes. *Journal of Financial Economics*, 126(1):1–35, 2017.

D. Heath, M. Ringgenberg, M. Samadi, and I. M. Werner. Reusing natural experiments. *Working Paper*, 1:1–55, 2020.

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

G. Hite and J. Owers. Security prices, reactions around corporate spin-off announcements. *Journal of Financial Economics*, 12(4):409–436, 1983.

A. E. Hoerl and R. W. Kennard. Ridge regression: Biased estimation for nonorthogonal problems. *Technometrics*, 12(1):55–67, 1970.

B. Holland, S. Basu, and F. Sun. Neglect of multiplicity when testing families of related hypotheses. *Working Paper*, 1:1–46, 2010.

S. Holm. A simple sequentially rejective multiple test procedure. *Scandinavian Journal of Statistics*, 6(1):65–70, 1979.

K. Hou, C. Xue, and L. Zhang. Replicating anomalies. *Review of Financial Studies*, 33(5): 2019–2133, 2020.

J. Hu. Does option trading convey stock information? *Journal of Financial Economics*, 111(3):625–645, 2014.

H. Jacobs. What explains the dynamics of 100 anomalies? *Journal of Banking and Finance*, 57(1):65–85, 2015.

Y. Jeon, T. H. McCurdy, and X. Zhao. News as sources of jumps in stock returns: Evidence from 21 million news articles for 9000 companies. *Working Paper Rotman School of Management University of Toronto*, 1:1–64, 2021.

T. L. Johnson and E. C. So. The option to stock volume ratio and future returns. *Journal of Financial Economics*, 106(2):262–286, 2012.

G. Jostova, S. Nikolova, A. Philipova, and C. W. Stahel. Momentum in corporate bond returns. *Review of Financial Studies*, 26(7):1649–1693, 2013.

K. Jurado, S. C. Ludvigson, and S. Ng. Measuring uncertainty. *American Economic Review*, 105(3):1177–1216, 2015.

M. Kacperczyk and E. S. Pagnotta. Chasing private information. *Review of Financial Studies (FORTHCOMING)*, 0(0):1–51, 2019.

D. B. Keim and R. F. Stambaugh. Predicting returns in the stock and bond markets. *Journal of Financial Economics*, 17(1):357–390, 1986.

E. Kim and J. McConnell. Corporate mergers and the co-insurance of corporate debt. *Journal of Finance*, 32(2):349–365, 1977.

S. S. Lee. Jumps and information flow in financial markets. *Review of Financial Studies*, 25(2):439–479, 2012.

H. Leland. Corporate debt value, bond covenants, and optimal capital structure. *Journal of Finance*, 49(4):1213–1252, 1994.

H. Leland and K. B. Toft. Optimal capital structure, endogenous bankruptcy, and the term structure of credit spreads. *Journal of Finance*, 51(3):987–1019, 1996.

H. Lin, J. Wang, and C. Wu. Liquidity risk and the cross-section of corporate bond returns. *Journal of Financial Economics*, 99(2):628–650, 2011.

H. Lin, Wu. C., and G. Zhou. Cross-sectional predictability of corporate bond returns. *Working Paper*, 1:1–83, 2019.

H. Lin, Wu. C., and G. Zhou. Extracting information from the corporate yield curve: A machine learning approach. *Working Paper*, 1:1–58, 2020.

A. W. Lo and A. C. MacKinlay. Data-snooping biases in tests of financial asset pricing models. *Review of Financial Studies*, 3(3):431–467, 1990.

L. A. Lochstoer and P. C. Tetlock. What drives anomaly returns? *Journal of Finance*, 75 (3):1417–1455, 2020.

A. Lynch and T. Vital Ahuja. Can subsample evidence alleviate the data-snooping problem? a comparison to the maximal rsquared cutoff test. *Working Paper*, 1:1–32, 2012.

J. M. Maheu and T. McCurdy. News arrival, jump dynamics, and volatility components for individual stock returns. *Journal of Finance*, 59(2):755–793, 2004.

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

S. Mansi and D. Reeb. Corporate diversification: What gets discounted? *Journal of Finance*, 57(5):2167–2183, 2002.

L. Marais, K. Schipper, and A. Smith. The wealth effects of going private for senior securities. *Journal of Financial Economics*, 58(2):895–919, 2003.

W. Maxwell and R. Rao. Do spin-offs expropriate wealth of bondholders. *Journal of Finance*, 58(5):2087–2108, 2003.

W. Maxwell and C. Stephens. The wealth effects of repurchases on bondholders. *Journal of Finance*, 58(2):895–919, 2003.

R. C. Merton. On the pricing of corporate debt: The risk structure of interest rates. *Journal of Financial Economics*, 29(2):449–470, 1974.

R. C. Merton. Option pricing when underlying stock returns are discontinuous. *Journal of Financial Economics*, 3(1):125–144, 1976.

T. Mitton. Methodological variation in empirical corporate finance. *Working Paper*, 1: 1–49, 2019.

E. Morellec and A. Zhdanov. Product market competition and option prices. *Review of Financial Studies*, 32(11):4343–4386, 2019.

T. Moskowitz, H.Y. Ooi, and L. H. Pedersen. Time series momentum. *Journal of Financial Economics*, 104(2):228–250, 2013.

H. Mulherin, J. Netter, and A. Poulsen. Observations on research and publishing from nineteen years as editors of the journal of corporate finance. *Journal of Corporate Finance*, 49(1):120–124, 2018.

K. Muller and S. N. M. Schmickler. Interacting anomalies. *Working Paper*, 1:1–55, 2020.

D. Muravyev and N. D. Pearson. Option trading costs are lower than you think. *Review of Financial Studies (FORTHCOMING)*, 0(0):0–0, 2020.

D. Muravyev, N. Pearson, and J. Broussard. Is there price discovery in the equity options market? *Journal of Financial Economics*, 107(2):259–283, 2013.

D. Muravyev, N. D. Pearson, and J. M. Pollet. Understanding returns to short selling using option-implied stock borrowing fees. *Working Paper*, 1:1–54, 2020.

S. Murray and S. Nikolova. The bond pricing implications of rating-based capital requirements. *SSRN Working Paper*, 1:1–117, 2019.

J. Narayanan and K. Shastri. The valuation impacts of specially designated dividends. *Journal of Financial and Quantitative Analysis*, 23(3):301–312, 1988.

W. K. Newey and K. D. West. A simple, positive semi-definite, heteroskedasticity and autocorrelation consistent covariance matrix. *Econometrica*, 55(3):703–708, 1987.

Y. Nozawa. What drives the cross-section of credit spreads?: A variance decomposition approach. *Journal of Finance*, 72(5):2045–2072, 2017.

D. B. Owen. A table of normal integrals. *Communications in Statistics - Simulation and Computation*, 9(4):389–419, 1980.

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

R. Parrino. Spinoffs and wealth transfers: The marriott case. *Journal of Financial Economics*, 43(2):271–274, 1997.

L. Pastor and R. F. Staumbaugh. Liquidity risk and expected stock returns. *Journal of Political Economy*, 111(3):642–685, 2003.

A. J. Patton and A. Timmermann. Monotonicity in asset returns: New tests with applications to the term structure, the CAPM, and portfolio sorts. *Journal of Financial Economics*, 98(3):605–625, 2010.

L. H. Pedersen. *Efficiently Inefficient: How Smart Money Invests and Market Prices are Determined*. Princeton University Press, Princeton, NJ, 2015.

R. Roll. A simple implicit measure of the effective bidask spread in an efficient market. *Journal of Finance*, 39(4):1127–1139, 1984.

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

M. Rubinstein. Implied binomial trees. *Journal of Finance*, 49(3):771–818, 1994.

R. Sadka. Momentum and post-earnings-announcement drift anomalies: The role of liquidity risk. *Journal of Financial Economics*, 80(2):309–349, 2006.

M. Sandulescu. How integrated are corporate bond and stock markets? *Working paper, University of Michigan*, 1:1–41, 2020.

J. Shanken. Intertemporal asset pricing: An empirical investigation. *Journal of Econometrics*, 45(1):99–120, 1990.

K. C. Smith. Option prices and disclosure: Theory and measurement. *Working Paper Stanford University Graduate School of Business*, 1:1–50, 2018.

K. C. Smith. Financial markets with trade on risk and return. *Review of Financial Studies*, 32(10):4042–4078, 2019.

R. F. Stambaugh, J. Yu, and Y. Yuan. Arbitrage asymmetry and the idiosyncratic volatility puzzle. *Journal of Finance*, 70(1):1903–1948, 2015.

R. Sullivan, A. Timmermann, and H. White. Dangers of data mining: The case of calendar effects in stock returns. *Journal of Financial Economics*, 105(1):249–286, 2001.

S. Szaura. Are stock and corporate bond markets integrated? a big data approach. *Working Paper McGill University*, 1:1–81, 2020.

M. Tian and L. Wu. Cross-sectional variation of option implied volatility skew. *Working Paper Zicklin School of Business, Baruch College, The City University of New York*, 1: 1–58, 2020.

R. Tibshirani. Regression shrinkage and selection via the LASSO. *Journal of the Royal Statistical Society: Series B (Methodological)*, 58(1):267–288, 1996.

K. B. Toft and B. Pryck. Options on leveraged equity: Theory and empirical tests. *Journal of Finance*, 52(3):1151–1180, 1997.

J. M. Vanden. Information quality and options. *Review of Financial Studies*, 21(6):2635–2676, 2008.

M. Vassalou and Y. Xing. Default risk in equity returns. *Journal of Finance*, 59(2):831–869, 2004.

A. Warga and Welch. Bondholder losses in leveraged buyouts. *Review of Financial Studies*, 6(4):959–982, 1993.

R. Woolridge. Dividend changes and security prices. *Journal of Finance*, 38(5):1607–1615, 1983.

X. Xiao and A. Vasquez. Default risk and option returns. *Working Paper*, 1:1–62, 2020.

Y. Xing, X. Zhang, and R. Zhao. What does the individual option volatility smirk tell us about future equity returns? *Journal of Financial and Quantitative Analysis*, 45(3): 641–662, 2010.

X. S. Yan and L. Zheng. Fundamental analysis and the cross-section of stock returns: A data-mining approach. *Review of Financial Studies*, 30(4):1382–1423, 2017.

H. Zhou and T. Hastie. Regularization and variable selection via the elastic net. *Journal of the Royal Statistical Society: Series B (Methodological)*, 67(2):301–320, 2005.

J. V. Zundert and J. Driessen. Are stock and corporate bond markets integrated? evidence from expected returns. *Working Paper*, 1:1–60, 2017.