Essays on Asset Pricing under Market Frictions

Kazuhiro Hiraki

Submitted in partial fulfillment of the requirements of the Degree of Doctor of Philosophy (PhD) in Finance

School of Economics and Finance

Queen Mary University of London

United Kingdom

August 2019

Statement of originality

I, Kazuhiro Hiraki, confirm that the research included within this thesis is my own work or that where it has been carried out in collaboration with, or supported by others, that this is duly acknowledged below and my contribution indicated.

I attest that I have exercised reasonable care to ensure that the work is original, and does not to the best of my knowledge break any UK law, infringe any third party's copyright or other Intellectual Property Right, or contain any confidential material.

I accept that the College has the right to use plagiarism detection software to check the electronic version of the thesis.

I confirm that this thesis has not been previously submitted for the award of a degree by this or any other university.

Signature:

Date: 11 August, 2019

Details of collaboration

Author contributions and additional collaborators are listed below for each chapter.

Introduction: Kazuhiro Hiraki wrote Introduction.

Chapter 1: This Chapter is based on joint work with George Skiadopoulos.

Chapter 2: This Chapter is based on joint work with George Skiadopoulos.

Chapter 3: This Chapter is based on joint work with Chuanping Sun.

Conclusions: Kazuhiro Hiraki wrote Conclusions.

Acknowledgment

First and foremost, I would like to thank my supervisor, Prof. George Skiadopoulos, for his continued support, precious guidance and passionate encouragement. I would not have been able to complete this piece of work without his enormous kindness and devotion to the supervision and our collaborative research projects since the inception of my study at Queen Mary University of London.

I am also grateful to the former and current faculty members at the School of Economics and Finance, especially Dr Renato Faccini, Dr Elise Gourier, Dr Tatsuro Senga, Prof. Jason Sturgess, Dr Roman Sustek, and Dr Konstantinos Zachariadis, for their support. I would also like to thank the administrative staff, especially Nick Owen, for their assistance.

Part of my thesis is hugely benefited from valuable feedback received at internal seminars at Queen Mary University of London and Hitotsubashi ISC as well as at the European meeting of the Econometric Society 2018, German Finance Association 2018 annual meeting, OptionMetrics Research Conference 2018, The Seventh Public Investors Conference and Paris-December Finance Meeting 2018. I would like to thank the organizers and the participants at these seminars and conferences. Specifically, I would like to thank the discussants, Dr Julian Thimme and Dr Ding Luo, for their valuable comments and suggestions. I would also like to thank Prof. Richard Payne and Prof. Roman Kozhan for being my thesis examiners. The discussion during the viva was enlightening and their thoughtful comments improved this thesis significantly.

My great thanks go to my colleagues for our friendship and stimulating conversations on academic topics. In particular, I would like to thank Sean for his enthusiasm in our joint research projects as well as for sharing enjoyable experiences.

I also wish to acknowledge the support from my employer, the Bank of Japan, which enabled me to concentrate on my PhD study.

Last but not least, I would like to thank my family for their enormous support and encouragement. Specifically, I am hugely indebted to my wife, Shiori, who understands and supports my struggle with research.

Abstract

This thesis comprises three studies on asset pricing under market frictions.

In Chapter One, I derive a model-free formula to estimate the *contribution of frictions to expected returns* (CFER) within a formal asset pricing setting. I show that properly scaled deviations from put-call parity reliably estimate CFER. The estimated CFER is sizable, it predicts stock returns and it has superior properties than previously proposed measures of deviations from put-call parity. I find that its predictive power stems from capturing the effect of market frictions, rather than from omitted factors and informed option trading. My findings suggest that market frictions, especially transaction costs, have a sizable effect even on large optionable stocks.

In Chapter Two, I study the effect of CFER on the *risk-neutral* expected asset return and the return predictive ability of the risk-neutral skewness (RNS). I show that a non-zero CFER is equivalent to the violation of the *martingale restriction* (MR), that is, the deviation of the risk-neutral expected return from the risk-free rate. I theoretically and empirically show that Bakshi et al.'s (2003) formula for RNS incurs a bias when the underlying asset violates MR. I document that the ability of RNS to predict stock returns stems from its estimation bias due to the violation of MR, implying that its predictability is driven by the presence of market frictions.

In Chapter Three, I examine the mean-variance portfolio strategy, which uses the estimated CFER as the mean stock return input. This CFER-based strategy outperforms various strategies including the equally-weighted portfolio, minimum variance portfolios and existing option-based portfolio strategies, especially when constraints on portfolio weights are imposed. My empirical results complement Jagannathan and

Ma's (2002, 2003) theoretical finding that constraints on portfolio weights help to improve the performance of *mean*-variance portfolio by mitigating measurement errors in the expected stock returns.

Contents

List of Figures

List of Tables

Introduction

Understanding the determinants of the expected asset return is one of the central research themes in asset pricing study. In a frictionless market, the expected asset return is determined by the covariance risk premium term, that is, the covariance between the stochastic discount factor and the asset return. Researchers have been devoting enormous effort to investigate the determinants of the covariance risk premium term. On the other hand, a growing number of studies have been documenting that market frictions such as margin constraints, short-sale constraints and transaction costs also affect the expected asset return, in addition to the covariance risk premium component.

The purpose of this thesis is to study the *contribution of frictions to expected returns* (CFER), which is the effect of frictions on the expected return not attributable to the covariance risk premium term. Specifically, I propose an option-based estimation approach for the CFER of stocks, theoretically and empirically examine the property of CFER, and investigate the implications of a non-zero CFER on the estimation of option-implied variables and the portfolio selection.

In Chapter One, I develop the theoretical framework I rely on throughout this thesis. I consider a formal asset pricing setting under market frictions, in which an agent determines her consumption and portfolio allocation on the stock and equity options by maximizing her life-time utility. Market frictions are formulated as constraints on her portfolio allocation choice and transaction costs she incurs for trading financial assets. Under this setting, the expected stock return is expressed as

$$\text{(Expected excess return)} = CFER + \text{(Covariance risk premium term)}. \qquad (1)$$

Then, I theoretically and empirically document that the CFER term can be reliably estimated by properly scaled deviations from put-call parity, that is, the difference between the underlying stock price and the synthetic stock price calculated from option prices.

In the empirical part of Chapter One, I estimate CFER for U.S. common stocks and obtain the following two important findings. First, I find that the estimated CFER has strong return predictive power; the long-short spread portfolio of the CFER-sorted value-weighted decile portfolios earns the average return and risk-adjusted returns of about 20% per year. I confirm that the strong return predictive ability of CFER is robust to the recent data snooping concerns and does not stem from omitted risk factors. Moreover, I show that CFER outperforms the implied volatility spread (IVS) as a predictor of stock returns. I also document that the size of market frictions is a more pertinent explanation to the return predictability of CFER than option trading activity.

Second, market frictions have a non-negligible effect even on optionable stocks (stocks with liquid option trading) which tend to be large and liquidly traded stocks. I reconcile this finding with Hou et al. (2018), who document that almost all previously documented friction-related anomalies vanish once micro-cap stocks are weighted less in the analysis. This is possible thanks to the ability of CFER to identify outperforming and underperforming stocks due to its theoretical foundation as a risk-adjusted signed measure of expected return (alpha) which captures the effect of market frictions; on the contrary, common measures of market frictions do not possess this property. In addition, I theoretically show that the upper bound of the alpha of CFER-sorted spread portfolios is at least twice the round-trip transaction costs. This result accommodates the sizable alpha of the CFER-sorted spread portfolio I have found.

Chapter Two examines the effect of CFER on the risk-neutral expected asset return and the return predictive ability of the risk-neutral skewness (RNS) by relating the *martingale restriction* (MR) to my CFER framework. MR is a property of asset prices in the absence of market friction first examined by Longstaff (1995); if the market is frictionless and arbitrage-free, asset prices discounted by the risk-free rate should be martingales under the risk-neutral measure. I show that equation (1) implies that the expected excess return under the risk-neutral measure equals CFER, that is, a non-zero CFER is equivalent to the violation of MR by the underlying stock.

Next, I show that the Bakshi et al. (2003) (BKM) formulae for risk-neutral moments (RNMs) do not correctly estimate RNMs if MR is violated. This is due to their implicit assumption that the underlying asset satisfies MR and hence the expected risk-neutral return of the underlying equals the risk-free rate. The bias arises because a wrong expected return value is employed to calculate the central moments (i.e., moments around the mean) of the underlying return. To remedy this drawback of the original BKM formulae, I propose the generalized BKM formulae, which account for the possible violation of MR, that is, a non-zero CFER.

My empirical analysis in Chapter Two reveals that the violation of MR has an important effect on the return predictive power of RNS; the estimated RNS based on the original BKM formula (O-RNS) has strong return predictive power as documented in the previous literature, whereas the estimated RNS based on my generalized BKM formula does not predict future stock returns. Furthermore, I provide evidence that the predictive power of O-RNS stems from its estimation bias caused by the violation of MR. The bias component is highly correlated with CFER in line with the fact that CFER measures the degree of the violation of MR. In short, O-RNS predicts future returns because its estimation bias contains predictive power inherited from the strong predictability of CFER.

In Chapter Three, I study the application of my CFER framework to the Markowitz (1952) mean-variance portfolio construction. To this end, I generalize Martin and Wagner's (2018) formula for the expected stock return by allowing the existence of the CFER-type market frictions. This generalized formula expresses the covariance risk premium term in equation (1) using the risk-neutral simple stock variance of Martin (2017). Relying on this theoretical result, I examine three mean-variance portfolio strategies, for which I use CFER, the risk premium component implied by the risk-neutral simple variance, and the sum of these two components, respectively, as the estimate of the expected stock returns.

My empirical analysis shows that the mean-variance portfolio strategy for which I use the estimated CFER as the estimate of the expected stock returns (the Q-CFER strategy; Q stands for "quantitative" estimate of the expected return) outperforms other portfolio strategies which have been documented to have good performance, including the equally-weighted portfolio (DeMiguel et al., 2009b), minimum variance portfolios (e.g., Jagannathan and Ma, 2003), and existing option-based portfolio strategies in DeMiguel et al. (2013), Kempf et al. (2015), and Martin and Wagner (2018). This finding suggests that, in contrast to the well-documented poor performance of the mean-variance portfolio based on the historical sample mean returns, the option-implied expected stock returns enable us to construct a mean-variance portfolio which has superior empirical performance.

The outperformance of the Q-CFER strategy suggests that incorporating the effect of market frictions into the estimation of the expected return is of importance. Tilting portfolio weights toward stocks with a high CFER component is effective to improve the portfolio expected return without much increasing the portfolio variance. This is because the CFER component is not a compensation for risk exposures.

Moreover, I find that imposing constraints on portfolio weights dramatically improves the performance of mean-variance portfolios. This result complements the theoretical result in Jagannathan and Ma (2002, 2003) on the *mean*-variance portfolio construction. Albeit less famous compared to their seminal result on the *minimum* variance portfolio, they also show that portfolio weight constraints on the mean-variance portfolio problem can be interpreted to have a *shrinkage-like* effect on the *expected stock returns*; the weight constraints mitigate measurement errors in the expected stock returns by increasing (decreasing) possibly underestimated (overestimated) expected return elements. Nevertheless, Jagannathan and Ma fail to find the empirical usefulness of this shrinkage-like effect for the mean-variance portfolios based on the historical sample mean returns. My result shows for the first time that this shrinkage-like effect of the weight constraints on the mean-variance portfolio problem is useful for portfolio selection once the option-implied expected return is employed.

Now, I briefly overview the strands of the literature pertaining to this thesis. To begin with, the estimation of CFER is pertaining to a number of theoretical and empirical studies on asset pricing and market frictions. Early studies by He and Modest (1995) and Luttmer (1996) examine whether the equity risk premium puzzle may be solved by taking market frictions into account. A strand of the literature develops asset pricing models by assuming specific frictions such as liquidity risk (Acharya and Pedersen, 2005), market and funding liquidity constraints (Brunnermeier and Pedersen, 2009), margin constraints (Gârleanu and Pedersen, 2011; Chabakauri, 2013), margin and leverage constraints (Frazzini and Pedersen, 2014) and exclusion of strategies with possible unlimited losses (Jarrow, 2016). Brennan and Wang (2010) and Hou et al. (2016) propose a reduced form model of frictions/mispricing and asset pricing, that is, they make no assumption on the type of frictions, yet they make assumptions on the dynamics of mispricing and the specification of the IMRS to estimate the effect of mispricing/frictions on expected returns from historical data.

There are also numerous empirical studies which examine the relation between the cross-section of stock returns and market frictions. Examples include stock-level illiquidity (Amihud, 2002), short-sale constraints (e.g., Chen et al., 2002; Ofek et al., 2004; Asquith et al., 2005; Drechsler and Drechsler, 2014), "betting against beta" effect due to leverage constraints (Frazzini and Pedersen, 2014; Jylhä, 2018), uncertainty about future shorting costs (Engelberg et al., 2018), idiosyncratic volatility (Ang et al., 2006; Stambaugh et al., 2015), delay in the response of prices to information (Hou and Moskowitz, 2005) and intermediaries' liquidity constraints (Nagel, 2012). Related to this strand of literature, Hou et al. (2018) examine more than 100 friction-related anomaly variables.

My approach for the estimation of CFER is distinct from previous approaches that measure the effect of frictions on the expected return in three important ways. First, my CFER measure circumvents assumptions on the dynamics of the effect of frictions and agents' preferences, since it has no free parameters. Second, my approach provides a real-time, forward-looking measure of the effect of frictions on each individual stock. This is in contrast to the strand of the friction-based anomaly literature, which captures the effect of frictions via the alphas of portfolios constructed on backward-looking proxies of market frictions. Third, the quantitative nature of CFER provides richer predictions compared to other characteristic variables and measures of deviations from put-call parity found to predict stock returns. This quantitative nature of CFER plays key roles in all the three studies in this thesis. In Chapter One, my theoretical framework pins down the quantitative relation between the estimated

CFER, the empirically estimated alphas and the size of market frictions. Regarding my analysis on RNS in Chapter Two, the estimated CFER provides the degree of the violation of MR. This information enables me to measure the bias in O-RNS caused by the violation of MR. In Chapter Three, I use the estimated CFER as an input to the mean-variance problem for the formation of the Q-CFER mean-variance portfolio.

My thesis contributes to the literature on the informational content of option prices in three important ways. First, I document that CFER has superior properties compared to previously proposed measures based on deviations from put-call parity such as IVS (e.g., Bali and Hovakimian, 2009; Cremers and Weinbaum, 2010) and the DOTS measure (Goncalves-Pinto et al., 2019). In Chapter One, I document that IVS and DOTS are only an approximation of my CFER estimator, and I find that the estimated CFER outperforms IVS and DOTS as a sorting variable to form portfolios. Moreover, unlike IVS and DOTS, my theoretical framework allows me to quantitatively relate CFER to empirically estimated alphas thanks to the fact that CFER is an estimate of part of the expected return.

Second, Chapter Two contributes to the literature on the return predictive ability of O-RNS. While it has been well-documented that O-RNS has return predictive power, the sign of its predictive relation and the mechanism of the predictability are still open questions. Conrad et al. (2013) document that O-RNS *negatively* predicts future returns, consistent with theoretical models in which risk-averse agents accept lower returns in exchange for positive skewness (e.g., Harvey and Siddique, 2000; Mitton and Vorkink, 2007). On the other hand, a number of recent studies document that RNS *positively* predicts future returns (Rehman and Vilkov, 2012; Stilger et al., 2017; Gkionis et al., 2018; Bali et al., 2018; Borochin and Zhao, 2018; Chordia et al., 2019), yet there is no consensus on the mechanism for the positive predictive relation between O-RNS and stock returns. My analysis sheds light on this ongoing debate by showing that O-RNS does not correctly estimate RNS when the underlying stock violates MR. This finding invalidates the (implicit) assumption in the previous literature that O-RNS correctly estimates the true risk-neutral skewness. More importantly, I document that it is the estimation bias in O-RNS caused by the violation of MR that predicts future stock returns; the estimation bias is highly correlated with CFER in line with the fact that CFER measures the degree of the violation of MR. This implies that O-RNS predicts stock returns because it proxies CFER and hence signals the effect of frictions on the expected stock returns. Analogous to my discussion on RNS, I also find that the predictive power of the Xing et al. (2010) slope measure of the implied volatility curve stems from its correlation with CFER.

Third, my study in Chapter Two reveals that a non-zero CFER and the violation of MR result in the incorrect estimation of option-based measures. This point has to do with the literature which assesses biases in empirically estimated option-based measures. Bliss and Panigirtzoglou (2002) and Hentchel (2003) examine the effect of measurement errors in option prices on the estimation of risk-neutral distributions and implied volatilities, respectively. Dennis and Mayhew (2009) investigate the effect of measurement errors in option prices and discretely observed strikes on the estimation of RNMs based on the BKM formulae. Ammann and Feser (2019) extend Dennis and Mayhew (2009) and investigate robust interpolation methodologies to obtain a continuum of option prices to calculate RNMs in the presence of measurement errors and limited number of option price observations. These studies assume that the underlying satisfies MR and focus on the effect of practical issues such as discretely traded strikes and measurement errors. My study differs from these studies in that the bias in RNMs caused by the violation of MR occurs even under an ideal situation where a continuum of option prices are observable without any measurement errors.

Chapter Three contributes to the literature on the mean-variance portfolio construction. It has been widely documented that the historical sample mean is notoriously noisy and the empirical performance of the mean-variance portfolio based on the historical sample mean is poor (e.g., Michaud, 1989; Best and Grauer, 1991; Chopra and Ziemba, 1993).

The previous literature proposes various ways to reduce measurement errors in the expected stock returns such as employing a Bayesian approach and factor-based expected stock return models (e.g., Black and Litterman, 1991; Garlappi et al., 2007; Lai et al., 2011, among others). Nevertheless, a large proportion of studies have been focusing on the minimum variance portfolio construction to bypass the estimation of the expected returns (e.g., Jagannathan and Ma, 2003). My option-based approach differs from the existing approaches because my option-based expected stock return is forward-looking and it incorporates the effect of market frictions on the expected stock returns. These features enable me to utilize additional information on the expected stock returns. Consequently, the Q-CFER mean-variance portfolio has superior performance compared to other existing portfolio strategies.

My analysis also contributes to the literature on the application of option-implied information for asset allocations and portfolio selection including Aït-Sahalia and Brandt (2008) and Kostakis et al. (2011), who use the option-implied distributional information for asset allocations. My analysis is directly related to DeMiguel et al. (2013), and Martin and Wagner (2018) in that these studies also document that option-implied information on the expected returns is useful to enhance the portfolio performance. My Q-CFER mean-variance portfolio outperforms the option-based strategies proposed in these studies. This is because I utilize richer option-implied information on the expected stock return to conduct the mean-variance optimization, whereas these studies utilizes option-based proxies of future stock returns and/or the cross-sectional ranking information of the expected stock returns.

Chapter 1

The Contribution of Frictions to Expected Returns

1.1 Introduction

In a frictionless market, asset expected returns are determined by the covariance between the stochastic discount factor and the asset's return (covariance risk premium). In the presence of market frictions, such as margin constraints and transaction costs, asset expected returns will also be determined by the *contribution of frictions to expected returns* (CFER). The estimation of CFER is of importance because CFER is inherently part of the expected return. Hence, its estimate can be used to predict asset returns, reveal the impact of market frictions on expected returns and identify the dominant type of market frictions.

In this Chapter, we document that properly scaled deviations from put-call parity calculated from the underlying stock price and the synthetic stock price obtained from option prices reliably estimate stock's CFER. Our approach is distinct from previous approaches that measure the effect of frictions on the expected return in three important ways. First, our CFER measure circumvents assumptions on the dynamics of the effect of frictions and agents' preferences, since it has no free parameters. Second, our approach provides a real-time, forward-looking measure of the effect of frictions on each individual stock. Previous studies typically capture this effect via the alphas of portfolios constructed on backward-looking proxies of market frictions. Third, the quantitative nature of CFER provides richer predictions compared to other characteristic variables and measures of deviations from put-call parity found to predict stock returns. For example, our theoretical framework pins down the relation between the estimated CFER, the empirically estimated alphas and the size of market frictions.

We employ a constrained consumption-portfolio choice setting where a marginal agent trades in both the stock and option market.[1] We model market frictions as constraints on the agent's portfolio allocation and as transaction costs for trading assets; we allow market frictions to affect both the stock and option prices. In this case, the required expected stock return may deviate from the covariance risk premium in an additive manner:

$$\text{(Expected excess return)} = \text{CFER} + \text{(Covariance risk premium term)}. \tag{1.1}$$

For instance, in the case where the agent buys stocks, she may demand a greater than the covariance risk premium expected return to compensate incurred transaction costs and the cost for posting her wealth as margin. CFER is the wedge between the expected excess return and the covariance risk premium. Importantly, CFER is not a compensation for risks, but it solely captures the effect of market frictions which cause limits of arbitrage.[2]

Why do scaled deviations from put-call parity reliably estimate the underlying stock's CFER? Our argument consists of two steps. First, in Section 1.2.2, we show that deviations from put-call parity contain information about the effect of frictions on both the underlying *and* the synthetic stock. Then, in Sections 1.2.3 and 1.3.3, we theoretically and empirically examine whether the effect of frictions on the synthetic stock can explain the size of observed deviations from put-call parity. We use the embedded leverage (Frazzini and Pedersen, 2012) and option margin constraints (Hitzemann et al., 2017) models, separately, to measure the effect of frictions on option prices and hence on the synthetic stock. We find that this effect is negligible compared to the observed deviations from put-call parity. This implies that devi-

[1] The existence of such an agent is realistic and it is in line with the recent literature. For example, a number of studies document that financial intermediaries is a marginal investor who trades in various financial markets simultaneously (e.g., Adrian et al., 2014; He et al., 2017)

[2] There is another strand of literature on the limits of arbitrage, where no constraints on portfolio allocations are imposed nor transaction costs and hence no CFER term appears (see for a survey Gromb and Vayanos, 2010 and references therein). Instead, this strand of literature views the risk-averseness of the agents as a friction. This approach is distinct from ours because it does not render deviations from the law of one price, which is a key ingredient of our option-based estimation formula of CFER.

ations from put-call parity are mainly determined by the effect of frictions on the underlying stock price, leading to our conclusion that scaled deviations from put-call parity reliably estimate CFER. This finding does not contradict the literature which documents that market frictions affect option returns (e.g., Frazzini and Pedersen, 2012; Santa-Clara and Saretto, 2009; Hitzemann et al., 2017) for two reasons. First, the findings in the previous literature imply that the dollar effect of frictions on options are non-negligible compared to the *option price level*. On the other hand, it is the ratio of the dollar effect of frictions on options to the *stock price level* that matters for the effect of frictions on the synthetic stock price; this ratio is much smaller than the former ratio. Second, the effect of frictions on a synthetic stock equals the difference between that on a call and on a put option. We find empirically that the effect of frictions on call and put options tends to have the same sign and similar magnitude so they roughly offset each other.

We estimate CFER for each optionable U.S. common stock from January 1996 to April 2016. Four are our main empirical findings. First, we document that CFER strongly predicts stock returns. A long-short spread portfolio of the CFER-sorted value-weighted decile portfolios yields a positive and statistically significant average return of 164 bps per month (t-stat: 5.76). Risk-adjusted returns with respect to standard asset pricing models are also sizable and statistically significant. For example, the Carhart (1997) four-factor alpha of the spread portfolio is 186 bps per month (t-stat: 6.56). This result is robust to a number of robustness tests and importantly to recent data snooping concerns (e.g., Harvey et al., 2016; Harvey, 2017; Hou et al., 2018) thanks to the sufficiently high t-statistics even for the value-weighted portfolios and the formal theoretical foundation of CFER.

The strong predictive power of CFER might not be surprising in the presence of the literature which documents that other measures based on deviations from put-call parity, such as the implied volatility spread (IVS) (e.g., Bali and Hovakimian, 2009; Cremers and Weinbaum, 2010) and the DOTS measure (Goncalves-Pinto et al., 2019), predict future stock returns. However, our CFER measure has superior properties compared to these measures. We show that IVS and DOTS are only an approximation of our CFER estimator. Consistent with this relation, we find that CFER outperforms IVS and DOTS. Moreover, in contrast to CFER, IVS and DOTS cannot be related quantitatively to their associated empirically estimated alphas because they cannot be interpreted as an estimate of a part of the expected return. Thus, no explanation about the size of their alphas has been provided so far. CFER provides such an explanation thanks to its theoretical properties as we discuss below. Interestingly, our results on the predictive power of CFER contribute to the debate about the source of the predictive power of deviations from put-call parity. We document that the predictive power of CFER is greater for stocks with low option trading volume. Our friction-based view can accommodate this empirical pattern because we find that these stocks face larger market frictions. On the other hand, the theory which attributes the predictive power of IVS to informed option trading yields the opposite prediction. Our finding is largely in line with Goncalves-Pinto et al. (2019), who provide evidence that the predictive power of IVS may not be the result of informed option trading.

Second, we find that the regressions of "CFER-adjusted excess return" (excess return less CFER) of the CFER-sorted portfolios on risk factors yield insignificant intercepts. This type of regressions are motivated by the theoretical prediction of equation (1.1). Our results imply that the predictive power of CFER originates from estimating the effect of market frictions accurately, rather than from omitted risk factors, thus corroborating our first finding. Our approach enable us to bypass the *risk-versus-mispricing* debate (Hou et al., 2018) which recognizes that it is notoriously difficult to conclude whether empirically observed anomalies (alphas) are driven by omitted risk factors ("risk") or "mispricing" (joint hypothesis problem, Fama, 1991); other option-implied measures cannot make use of this approach.

Our third finding is that market frictions matter even for large optionable stocks. The estimated CFER ranges from -1.24% to 0.89% per month in a 5th to 95th percentile range; CFER can become twice as large in magnitude as the average U.S. equity premium (0.5% per month; Mehra, 2012). This result in conjunction with the, stated above, large alpha of the CFER-sorted long-short portfolio (about 20% per year) may seem to be in contrast to the findings in the recent literature. For instance, Hou et al. (2018) document that 95 out of 102 friction-related anomaly variables become insignificant when the effect from microcap stocks is mitigated. This would imply that friction-related anomalies are largely irrelevant to stocks with sufficient option trading which tend to be large stocks.

We reconcile this seemingly contradicting finding by theoretically deriving the range of possible CFER values by considering specific type of frictions. In the case where we consider only transaction costs, we show that the value of CFER lies approximately between -2ρ and 2ρ, where 2ρ stands for the round-trip transaction costs. This result yields two implications which explain why deviations from put-call parity can result in a large alpha even within the universe of optionable stocks, thus accommodating our CFER-sorted spread portfolio's reported alpha. First, the theoretical upper bound of the CFER-sorted spread portfolio's alpha (the expected excess return not explained by the covariance risk premium) implies that the CFER-sorted portfolio can earn alpha as big as 2% per month, given the empirically estimated value of round-trip transaction costs for large stocks (approximately 1%; see e.g., Lesmond et al., 1999; Hasbrouck, 2009). Second, the size of transaction costs does not have a monotonic relation with CFER; larger ρ implies that the stock may have very positive CFER or very negative CFER. This implies that when one sorts stocks based on a proxy of ρ, the highest ρ portfolio will contain both outperforming and underperforming stocks, and the portfolio may show no abnormal return on average. This explains the findings of Hou et al. (2018).

Fourth, even though CFER is a "sufficient statistic" which subsumes the overall effect of any relevant market frictions on expected returns, we find that transaction costs play a key role for its determination. Fama and MacBeth (1973) regressions of CFER on a set of market frictions-related variables reveal that CFER is strongly related to a number of proxies of transaction costs. We also find that the tighter short-sale constraints are, the more negative CFER becomes. This is in accordance with previous literature, which documents that stocks which are subject to short-sale constraints underperform (e.g., Ofek et al., 2004; Drechsler and Drechsler, 2014), yet we find that the effect of short-sale constraints to CFER is of second order importance. This is expected because short-sale constraints are less pronounced among big stocks like our optionable stocks (e.g., D'Avolio, 2002; Asquith et al., 2005).

Our study is related to three strands of literature. First, it draws upon the theoretical literature on asset pricing under market frictions, especially Gârleanu and

Pedersen (2011).[3] We share a key insight that deviations from the law of one price (LoOP) between the underlying and derivatives occur due to limits of arbitrage caused by market frictions; put-call parity is an example of LoOP between the underlying stock and a synthetic stock. We differ from their study in two ways. First, they mainly investigate the bond-CDS basis, whereas we study the deviations between individual stock prices and their corresponding synthetic stock prices. Second, the focus of the two studies is different. Gârleanu and Pedersen (2011) investigate theoretically why violation in the LoOP occurs. We treat the empirically observed deviations from LoOP as given and study their relation with expected stock returns. Our study is also related to Brennan and Wang (2010) and Hou et al. (2016), who propose a reduced form model of frictions/mispricing and asset pricing. These models make no assumption on the type of frictions, yet they make assumptions on the dynamics of mispricing and the specification of the IMRS to estimate the effect of mispricing/frictions on expected returns from historical data. In contrast, the estimation of CFER does not involve any parameter estimations nor historical data.

The study in this Chapter also complements empirical studies which examine the relation between the cross-section of stock returns and market frictions such as stock-level illiquidity (Amihud, 2002), short-sale constraints (e.g., Chen et al., 2002; Ofek et al., 2004; Asquith et al., 2005; Drechsler and Drechsler, 2014), "betting against beta" effect due to leverage constraints (Frazzini and Pedersen, 2014; Jylhä, 2018), uncertainty about future shorting costs (Engelberg et al., 2018), idiosyncratic volatility (Ang et al., 2006; Stambaugh et al., 2015), delay in the response of prices to information (Hou and Moskowitz, 2005) and intermediaries' liquidity constraints (Nagel, 2012). We also document that market frictions affect expected stock returns even for large stocks, and we reconcile our findings with Hou et al.'s (2018) seemingly contradicting evidence that most of friction-related anomalies are spurious.

Finally, as discussed above, our research is pertaining to the literature on the informational content of option prices, especially to studies which document that

[3] Other related studies include early studies by He and Modest (1995) and Luttmer (1996), who examine whether the equity risk premium puzzle may be solved by taking market frictions into account, and a strand of the literature which develops asset pricing models by assuming specific frictions such as liquidity risk (Acharya and Pedersen, 2005), market and funding liquidity constraints (Brunnermeier and Pedersen, 2009), margin constraints (Chabakauri, 2013), margin and leverage constraints (Frazzini and Pedersen, 2014) and exclusion of strategies with possible unlimited losses (Jarrow, 2016).

measures based on deviations from put-call parity (Ofek et al., 2004; Bali and Hovakimian, 2009; Cremers and Weinbaum, 2010; Muravyev et al., 2016; Goncalves-Pinto et al., 2019) predict future stock returns.[4] Our CFER measure is also related to Martin (2017), Martin and Wagner (2018) and Kadan and Tang (2017), who show that the expected stock returns can be estimated quantitatively based on an option-based forward-looking measure. We differ from them in that we estimate the effect of frictions on the expected return, whereas they estimate the covariance risk premium term under the frictionless market assumption.

The rest of this Chapter is organized as follows. In Section 1.2, we provide the option-based formula to estimate CFER within our asset pricing model under market frictions and discuss the testable predictions of the model. Section 1.3 describes the data, the way we implement our formula to estimate CFER, and the summary statistics of the estimated CFER. In Section 1.4, we document the ability of the estimated CFER to predict stock returns. In Section 1.5, we discuss the relation between CFER and market frictions. Section 1.6 provides results on robustness tests. Section 1.7 concludes and discusses the findings.

1.2 Theoretical framework

1.2.1 Asset pricing under market frictions

We assume that the time horizon is finite and discrete, indexed by $t = 0, 1, 2, \dots, T$. Three types of assets, the stock, the risk-free bond, and European call and put options written on the stock, are traded in the market. We denote the stock price by S_t and its dividend payment at time t by D_t. The gross stock return and the gross risk-free rate from time t to $t+1$ are denoted by $R_{t,t+1}$ and $R^0_{t,t+1}$, respectively. We assume that options written on the stock are one-period options (i.e., options traded at time t mature at $t+1$) and traded at a set of strikes $\mathcal{K}_t$. The time t call (put) option price with strike price $K \in \mathcal{K}_t$ is denoted by $C_t(K)$ $(P_t(K))$.

We assume that there exists an agent who participates in *both* the stock market and the option market. She sets her optimal consumption and asset allocations by

[4]There is also a voluminous literature which finds that other option-based measures predict future stock returns; see Giamouridis and Skiadopoulos (2011) and Christoffersen et al. (2013) for reviews.

maximizing her expected lifetime utility, yet her asset allocation is subject to constraints caused by market frictions. The assumption of the existence of such an agent is realistic and it is in line with recent literature; for example, large financial intermediaries trade in multiple financial markets including the stock market and the option market under market frictions, and a growing number of recent studies consider financial intermediaries to be the marginal investors (e.g., Adrian et al., 2014; He et al., 2017).

Let θ_t^0, θ_t^S, $\theta_t^c(K)$ and $\theta_t^p(K)$ be the agent's position on the risk-free bond, the stock, the call and put options, respectively, and let $\boldsymbol{\theta}_t$ be the vector of these thetas. The agent solves the following portfolio-consumption problem,

$$\max_{\{c_j,\boldsymbol{\theta}_j\}} \sum_{j=t}^{T} \beta^{j-t}\mathbb{E}_t^{\mathbb{P}}[u(c_j)], \tag{1.2}$$

where $\mathbb{E}_t^{\mathbb{P}}$ is the conditional expectation under the agent's subjective belief $\mathbb{P}$ given the information up to time t, β is the subjective discount factor, $u(c)$ is the time-separable utility function.[5] The agent chooses a consumption stream $\{c_j\}_{j\geq t}$ and portfolio allocations $\{\boldsymbol{\theta}_t\}_{j\geq t}$ subject to the following conditions. First, the agent's wealth at time t, W_t, changes over time as follows:

$$W_{t+1} = \theta_t^0 R_{t,t+1}^0 + \theta_t^S(S_{t+1}+D_{t+1}) + \sum_{K\in\mathcal{K}_t}\left[\theta_t^c(K)(S_{t+1}-K)^+ + \theta_t^p(K)(K-S_{t+1})^+\right], \tag{1.3}$$

where $(x)^+ = \max(x,0)$. Second, the consumption at time t is given by

$$c_t = W_t - \theta_t^0 - \theta_t^S S_t - \sum_{K\in\mathcal{K}_t}\left[\theta_t^c(K)C_t(K) + \theta_t^p(K)P_t(K)\right] - TC_t(\Delta\boldsymbol{\theta}_t). \tag{1.4}$$

In equations (1.3) and (1.4), we normalize the price of the one-period bond at time t to unity. The last term, $TC_t(\Delta\boldsymbol{\theta}_t)$, is the transaction costs function, which denotes the dollar amount of transaction costs the agent incurs at time t. We formulate it as

$$TC_t(\Delta\boldsymbol{\theta}_t) = \rho_t^S S_t|\Delta\theta_t^S| + \rho_t^0|\theta_t^0| + \sum_{K\in\mathcal{K}_t}\left(\rho_t^{c,K}C_t(K)|\theta_t^c(K)| + \rho_t^{p,K}P_t(K)|\theta_t^p(K)|\right), \tag{1.5}$$

where positive coefficients ρ_t^S, ρ_t^0, $\rho_t^{c,K}$ and $\rho_t^{p,K}$ denote the proportional transaction

[5] The assumption of a time-separable utility function is done for simplicity. Our subsequently developed asset pricing model holds even for recursive utility functions.

costs for trading one unit of the respective assets. Equation (1.5) shows that the transaction costs the agent incurs are given by the product of the transaction cost per unit of asset (e.g., $\rho_t^S S_t$) with the units being traded (e.g., $|\Delta\theta_t^S|$). Note that the transaction costs for trading the stock depends on the *change* in the position $|\Delta\theta_t^S|$ between time $t-1$ and t, whereas these for the other assets depend on the level of the position. This is because we assume that options and the bond are one-period assets and hence their beginning-of-the-period holdings are zero. Therefore, the change in the position of these assets coincides with the level of the position at time t.

Third, we formalize market frictions (other than transaction costs) as constraints on the portfolio allocation of the agent. Specifically, we assume that there are L types of constraints on the portfolio allocation of the agent:

$$g_t^l(\boldsymbol{\theta}_t) \geq 0, \qquad l = 1, 2, \ldots, L. \tag{1.6}$$

Even though the functional form of g_t^l (nor the number of frictions L) is not required for our subsequent theoretical results in this subsection, we provide the margin constraint as an example of the constraint function to facilitate the understanding of our formulation:

$$g_t^{MC}(\boldsymbol{\theta}_t) = W_t - |\theta_t^S|\mu_t^S S_t - \sum_{K\in\mathcal{K}} (|\theta_t^c(K)|\mu_t^c(K)C_t(K) + |\theta_t^p(K)|\mu_t^p(K)P_t(K)), \tag{1.7}$$

where μ_t^S, $\mu_t^c(K)$ and $\mu_t^p(K)$ are the margin rates of the stock, call option and put option, respectively (e.g., $\mu_t^S S_t$ is the margin that traders need to hold when they trade one unit of the stock). The constraint $g_t^{MC}(\boldsymbol{\theta}_t) \geq 0$ imposes that the aggregate margin the agent needs to hold cannot exceed her wealth W_t. The absolute values of asset allocations are taken since typically traders need to hold margins both when they long and short assets.

Let $V_t(W_t, \theta_{t-1}^S)$ be the time-t value function of the constrained maximization problem (1.2) subject to equations (1.3) to (1.6). The allocation on the stock at the previous period θ_{t-1}^S is treated as a state variable because it affects the transaction costs incurred at time t and hence the agent's decision making. On the other hand, allocations on the bond and options at the previous period are not treated as state variables because we consider one-period bonds and options.[6] Then, the Bellman

[6]The value function V_t depends on t because we consider a finite horizon model, and the constraint

equation is given by

$$V_t(W_t, \theta^S_{t-1}) = \max_{c_t, \boldsymbol{\theta}_t} \left\{ u(c_t) + \beta \mathbb{E}^{\mathbb{P}}_t [V_{t+1}(W_{t+1}, \theta^S_t)] \right\} \quad s.t. \quad \text{equations (1.3)–(1.6)}. \tag{1.8}$$

Given equations (1.3) to (1.6), the first-order condition of the Bellman equation (1.8) regarding the allocation to the stock θ^S_t yields

$$S_t = \mathbb{E}^{\mathbb{P}}_t [m^*_{t,t+1}(S_{t+1} + D_{t+1})] + M^S_t - \frac{\partial TC_t}{\partial \theta^S_t} + \frac{\beta}{u'(c_t)} \mathbb{E}^{\mathbb{P}}_t \left[\frac{\partial V_{t+1}}{\partial \theta^S_t} \right], \tag{1.9}$$

where $m^*_{t,t+1} = \beta V'_{t+1}(W_{t+1})/u'(c_t)$ is the intertemporal marginal rate of substitution (IMRS) between time t and $t+1$. The second term in the right-hand side is given by $M^S_t = \left[\sum_{l=1}^{L} \lambda^l_t (\partial g^l_t(\boldsymbol{\theta}_t)/\partial \theta^S_t) \right] / u'(c_t)$, where λ^l_t is the Lagrange multiplier of l-th constraint of equation (1.6). This term captures the effect of market frictions (other than transaction costs) on the market price of the stock. It can be interpreted economically as the nominal shadow cost of frictions.[7] The third term in the right-hand side denotes the contemporaneous effect of transaction costs, that is, the amount of transaction costs the agent needs to incur by trading one more unit of the stock. The last term is pertaining to the future transaction costs. For example, in the case where the agent expects to unwind her position on the stock in the next period, she takes the future transaction costs for unwinding into account when she solves the intertemporal optimization problem. Equation (1.9) shows that the current stock price deviates from the IMRS-discounted expected future cum-dividend stock price due to transaction costs and other binding constraints. Equivalently, equation (1.9) shows that the standard asset pricing formula $1 = \mathbb{E}^{\mathbb{P}}_t [m^*_{t,t+1} R_{t,t+1}]$, which holds in frictionless markets, does not hold. This result echos the findings in He and Modest (1995) and Luttmer (1996).

functions g^l_t and the transaction costs function TC_t may also depend on time-varying parameters.

[7] For instance, in the case where the margin constraint (equation (1.7)) is the only constraint, M^S_t becomes $M^S_t = -(\lambda^{MC}_t / u'(c_t)) \times sgn(\theta^S_t) \mu^S_t S_t$. The former term is the nominal shadow price of one unit of wealth pledgeable as margin and the latter part equals the amount of margin that needs to be posted for trading one unit of the stock. Note that the effect of market frictions on the stock price, M^S_t may depend on the agent's allocations to other assets, as well. For example, in the above margin constraint example, the value of M^S_t depends on whether the margin constraint is binding and hence it depends on the total amount of margins the agent needs to post, which is a function of the vector of allocations $\boldsymbol{\theta}_t$.

Option prices satisfy the following first-order conditions:

$$\begin{aligned} C_t(K) &= \mathbb{E}_t^{\mathbb{P}}[m_{t,t+1}^*(S_{t+1}-K)^+] + M_t^c(K) - \frac{\partial TC_t}{\partial \theta_t^c(K)}, \\ P_t(K) &= \mathbb{E}_t^{\mathbb{P}}[m_{t,t+1}^*(K-S_{t+1})^+] + M_t^p(K) - \frac{\partial TC_t}{\partial \theta_t^p(K)}, \end{aligned} \tag{1.10}$$

where $M_t^o(K) = \left[\sum_{l=1}^{L} \lambda_t^l (\partial g_t^l(\boldsymbol{\theta}_t)/\partial \theta_t^o(K))\right] / u'(c_t)$ $(o \in \{c,p\})$ captures the effect of market frictions on the market call and put option prices in analogy to M_t^S in equation (1.9).[8] Unlike the stock case, the term related to future transaction costs does not appear in the first-order conditions for options. This is because we consider one-period options and hence she will no longer trade them as they expire at $t+1$.

The following Theorem provides the asset pricing model under market frictions.

Theorem 1.2.1. *Under market frictions, the following asset pricing model holds:*

$$\mathbb{E}_t^{\mathbb{P}}[R_{t,t+1}] - R_{t,t+1}^0 = CFER_{t,t+1} - \frac{Cov_t^{\mathbb{P}}(m_{t,t+1}^*, R_{t,t+1})}{\mathbb{E}_t^{\mathbb{P}}[m_{t,t+1}^*]}, \tag{1.11}$$

where $CFER_{t,t+1}$ is the contribution of frictions to the expected return from t to $t+1$, defined as

$$CFER_{t,t+1} = -\frac{R_{t,t+1}^0}{S_t}\left(M_t^S - \frac{\partial TC_t}{\partial \theta_t^S} + \frac{\beta}{u'(c_t)}\mathbb{E}_t^{\mathbb{P}}\left[\frac{\partial V_{t+1}}{\partial \theta_t^S}\right]\right). \tag{1.12}$$

Proof. See Appendix 1.A.1. □

Four remarks are in order at this point. First, equation (1.11) boils down to the standard frictionless market asset pricing model when constraints are not binding and no transaction costs exist, i.e. $CFER_{t,t+1} = 0$ as equation (1.12) shows. Second, CFER is part of the expected excess return, which cannot be explained by the covariance risk premium term (the covariance between the asset return and the IMRS) and hence it does not represent compensation for risk. The CFER term appears due to the binding constraints on asset allocations and the existence of transaction costs. Therefore, CFER cannot be interpreted as a new risk factor. Rather, in the Gârleanu and Pedersen (2011) terminology, one may view $CFER_{t,t+1}$ as the "alpha" of the

[8] Regarding the risk-free bond, to simplify the exposition, we assume that $\mathbb{E}_{t,t+1}^{\mathbb{P}}[m_{t,t+1}^*] = 1/R_{t,t+1}^0$ holds, which is equivalent to assuming that there is no effect of frictions on the risk-free bond market. In Appendix 1.D, we extend the model to allow for a non-zero effect of frictions on the risk-free bond market. Then, we demonstrate that the effect of frictions on the risk-free rate has no impact on the results presented in the main body of this Chapter.

stock, that is, the expected excess return adjusted for risk (Gârleanu and Pedersen, 2011, p.1990). Consistent with this view, in the case where the margin constraints are the only constraints (i.e., $L = 1$), $CFER_{t,t+1}$ coincides with what they call the alpha, which is the product of the Lagrange multiplier (shadow price) of the margin constraints and the margin rate of the stock.

Third, even though we name our new component of the expected return the "contribution of frictions to the expected return", it should be noted that CFER does not subsume the total effect of frictions. This is because the covariance risk premium term is also affected by market frictions; the IMRS depends on the agent's optimal allocation $\boldsymbol{\theta}_t$, which the agent chooses by taking the presence of market frictions into account. Therefore, CFER does not coincide with the difference between the expected return in this model and the expected return in the hypothetical frictionless market model. Instead, as we have pointed out above, the CFER term represents the part of the expected excess return not attributable to the covariance risk premium term, where the covariance risk premium term is calculated with the IMRS and the stock return formed in the presence of frictions.

Fourth, CFER is determined by three channels. The first one is the contemporaneous binding constrains represented by M_t^S. The second type is the contemporaneous effect of transaction costs represented by $\partial TC_t / \partial \theta_t^S$. The third one is pertaining to future transaction costs. For example, in the case where the agent expects to unwind her position on the stock in the next period, she takes future transaction costs for unwinding into account when she solves the intertemporal optimization problem.

The literature argues that uncertainty about future transaction costs affect expected stock returns. For example, Acharya and Pedersen (2005) show that the stochastic "liquidity cost," which they model as future transaction costs for selling stocks, affects stock returns. Engelberg et al. (2018) find that the uncertainty about future short-selling cost affects prices among the cross-section of stocks. We briefly argue that the future transaction cost component of CFER at least conceptually incorporates these mechanisms. To see this, the last term of CFER equals

$$\frac{\beta}{u'(c_t)} \mathbb{E}_t^{\mathbb{P}} \left[\frac{1}{S_t} \frac{\partial V_{t+1}}{\partial \theta_t^S} \right] = \mathbb{E}_t^{\mathbb{P}} \left[\beta \frac{u'(c_{t+1})}{u'(c_t)} \rho_{t+1}^S \frac{S_{t+1}}{S_t} sgn(\Delta \theta_{t+1}^S) \right]$$

due to the envelop theorem. For simplicity, let us assume that the stock pays no dividends and no $g_t^l(\boldsymbol{\theta}_t)$ functions in equation (1.6) which depend on the total wealth (e.g., the margin constraints function) are not binding. Then, this equation equals

$$\begin{aligned}\mathbb{E}_t^{\mathbb{P}}\left[m_{t,t+1}^* R_{t,t+1}\rho_{t+1}^S sgn(\Delta\theta_{t+1}^S)\right] = \mathbb{E}_t^{\mathbb{P}}[m_{t,t+1}^* R_{t,t+1}]\mathbb{E}_t^{\mathbb{P}}[\rho_{t+1}^S sgn(\Delta\theta_{t+1}^S)] \\ + Cov_t^{\mathbb{P}}(m_{t,t+1}^* R_{t,t+1}, \rho_{t+1}^S sgn(\Delta_{t+1}^S)).\end{aligned}$$

This equation shows that uncertainty about future transaction costs (including future short-selling costs) will affect CFER through the covariance between the stochastic future transaction costs and the future stock return.

1.2.2 Estimation of CFER: The formula

We assume that the dividend payment at $t+1$, D_{t+1}, is deterministic given the information up to time t. This assumption is plausible when the time length between t and $t+1$ is short (e.g., one-month as it will be the case in the subsequent empirical analysis) because near future dividend payments are usually pre-announced. We define the synthetic stock price $\widetilde{S}_t(K)$ as

$$\widetilde{S}_t(K) = C_t(K) - P_t(K) + \frac{K + D_{t+1}}{R_{t,t+1}^0}. \tag{1.13}$$

This combination of long call, short put and risk-free bond position is called a synthetic stock because it has the same payoff at time $t+1$ as the underlying stock, $S_{t+1} + D_{t+1}$.

Theorem 1.2.2. *Assume that D_{t+1} is deterministic given the information up to time t. Then, for any strike K, CFER is decomposed as*

$$CFER_{t,t+1} = CFER_{t,t+1}^{MF}(K) + U_{t,t+1}(K) + T_{t,t+1}(K), \quad where \tag{1.14}$$

$$CFER_{t,t+1}^{MF}(K) = \frac{R_{t,t+1}^0}{S_t}(\widetilde{S}_t(K) - S_t), \tag{1.15}$$

$$U_{t,t+1}(K) = -\frac{R_{t,t+1}^0}{S_t}[M_t^c(K) - M_t^p(K)], \tag{1.16}$$

$$T_{t,t+1}(K) = \frac{R_{t,t+1}^0}{S_t}\left(\frac{\partial TC_t}{\partial \theta_t^c(K)} - \frac{\partial TC_t}{\partial \theta_t^p(K)}\right). \tag{1.17}$$

Proof. The proof of Theorem 1.2.2 relies on the idea that deviations from put-call parity, $S_t - \widetilde{S}_t(K)$, are determined by the difference between the effect of market frictions on the underlying stock and that on the synthetic stock. To see this, under the conditionally deterministic dividend payment assumption, substitution of equation (1.10) into equation (1.13) yields the synthetic stock price as

$$\widetilde{S}_t(K) = \mathbb{E}_t^{\mathbb{P}}[m^*_{t,t+1}(S_{t+1}+D_{t+1})]+(M_t^c(K)-M_t^p(K))-\left(\frac{\partial TC_t}{\partial \theta_t^c(K)} - \frac{\partial TC_t}{\partial \theta_t^p(K)}\right). \quad (1.18)$$

Since the first term in the right-hand side of equations (1.9) and (1.18) are the same, deviations from put-call parity equal

$$\begin{aligned} S_t - \widetilde{S}_t(K) = &\left(M_t^S - \frac{\partial TC_t}{\partial \theta_t^S} + \beta \mathbb{E}_t^{\mathbb{P}}\left[\frac{\partial V_{t+1}}{\partial \theta_t^S}\right]\right) \\ &- \left((M_t^c(K) - M_t^p(K)) - \left(\frac{\partial TC_t}{\partial \theta_t^c(K)} - \frac{\partial TC_t}{\partial \theta_t^p(K)}\right)\right). \end{aligned} \quad (1.19)$$

Then, by scaling equation (1.19) by $-R^0_{t,t+1}/S_t$, we obtain

$$CFER^{MF}_{t,t+1}(K) = CFER_{t,t+1} - (U_{t,t+1}(K) + T_{t,t+1}(K)). \quad (1.20)$$

given equations (1.12), (1.15), (1.16) and (1.17). This proves equation (1.14). □

Theorem 1.2.2 shows that the stock's CFER equals the sum of the scaled deviation from put-call parity, $CFER^{MF}_{t,t+1}$, and the scaled effect of frictions on the synthetic stock, $U_{t,t+1} + T_{t,t+1}$, where $U_{t,t+1}$ denotes the effect of constraints and $T_{t,t+1}$ denotes the effect of transaction costs. This is a direct consequence of equation (1.20), which shows that once scaling by $R^0_{t,t+1}/S_t$, the deviation from put-call parity equals the difference between the effect of frictions on the underlying and that on the synthetic stock (equation (1.18)). Equations (1.19) and (1.20) echo equation (29) of Gârleanu and Pedersen (2011), which shows that deviations from LoOP are determined by market frictions, and they do not depend on preferences nor on the subjective beliefs. As a result, the calculation of $CFER^{MF}_{t,t+1}$ requires no assumptions on the preferences (IMRS m^*) and subjective beliefs ($\mathbb{P}$), neither estimation of any parameters and hence we call it the "model-free" (MF) part of CFER. It can be computed from observable option prices as long as a pair of European call and put options with the same maturity and strike is available. On the other hand, the effect of frictions on the synthetic stock (options) is not observable. We discuss its estimation in the next subsection.

1.2.3 Estimation of CFER: Modeling the effect of frictions on options

1.2.3.1 Estimation of $U_{t,t+1}$

To estimate $U_{t,t+1}(K)$, we need to model the effect of frictions on option prices, $M_t^c(K)$ and $M_t^p(K)$ (equation (1.16)). To this end, we employ two alternative models of frictions. The first is the *embedded leverage* effect documented by Frazzini and Pedersen (2012). They theoretically and empirically show that options with high embedded leverage attract investors who are subject to leverage constraints and hence these options have lower returns. Their finding suggests that leverage constraints affect option prices (i.e., non-zero $M_t^c(K)$ and $M_t^p(K)$ arise) and hence this may yield a non-zero $U_{t,t+1}(K)$. Based on the Frazzini and Pedersen (2014) study on embedded leverage (EL), we compute

$$U_{t,t+1}^{EL}(K) = k(|\Delta_p(K)| - \Delta_c(K)). \tag{1.21}$$

where $\Delta_p(K)$ and $\Delta_c(K)$ are the put and call options' delta, respectively, and the coefficient k is the sensitivity of option returns to the level of embedded leverage. Appendix 1.B.1.1 provides the proof of equation (1.21).

We consider option margins as the second type of market friction in the option market. Santa-Clara and Saretto (2009) and Hitzemann et al. (2017) document that margin constraints affect option returns and hence they may also yield a non-zero $U_{t,t+1}(K)$. Under margin constraints (equation (1.7)), the calculation of $U_{t,t+1}(K)$ yields

$$U_{t,t+1}^{MC}(K) = -R_{t,t+1}^0 \frac{\lambda_t^{MC}}{u'(c_t)} \left[\mu_t^c(K) \frac{C_t(K)}{S_t} sgn(\theta_t^c(K)) - \mu_t^p(K) \frac{P_t(K)}{S_t} sgn(\theta_t^p(K)) \right], \tag{1.22}$$

where the superscript MC stands for "margin constraints." To calculate $U_{t,t+1}^{MC}(K)$ in equation (1.22), we need to specify the option margin rates, $\mu_t^c(K)$ and $\mu_t^p(K)$. These are determined by the option exchange and depend on the strike price as well as on whether options are bought or sold. To this end, we follow Hitzemann et al. (2017), who compute the option margin rates by using the CBOE option margin rules:[9]

$$\mu_t^i(K) = 1, \qquad \text{when } \theta_t^i(K) > 0,\ i \in \{c, p\}, \tag{1.23}$$

$$\mu_t^c(K) = \frac{\max(0.2S_t - (K - S_t)^+, 0.1S_t)}{C_t(K)}, \qquad \text{when } \theta_t^c(K) < 0, \tag{1.24}$$

$$\mu_t^p(K) = \frac{\max(0.2S_t - (S_t - K)^+, 0.1K)}{P_t(K)}, \qquad \text{when } \theta_t^p(K) < 0. \tag{1.25}$$

To simplify the calculation in the latter two cases, we consider options with strikes which satisfy $8/9 \leq K/S_t \leq 1.1$. This range of strikes is not restrictive because we only use options with strikes which satisfy this range in our subsequent empirical analysis. Then, given the above margin rules, $U_{t,t+1}^{MC}(K)$ can be calculated as:

$$U_{t,t+1}^{MC}(K) = E_t(K) \times R_{t,t+1}^0 \lambda_t^{MC} / u'(c_t), \quad where \tag{1.26}$$

$$E_t(K) = \begin{cases} (C_t(K) - P_t(K))/S_t & \text{when } \theta_t^c(K) > 0 \text{ and } \theta_t^p(K) > 0 \\ -(S_t - K)/S_t & \text{when } \theta_t^c(K) < 0 \text{ and } \theta_t^p(K) < 0 \\ (0.2 + [C_t(K) - (S_t - K)^+]/S_t) & \text{when } \theta_t^c(K) > 0 \text{ and } \theta_t^p(K) < 0 \\ -(0.2 + [P_t(K) - (K - S_t)^+]/S_t) & \text{when } \theta_t^c(K) < 0 \text{ and } \theta_t^p(K) > 0. \end{cases} \tag{1.27}$$

Appendix 1.B.1.2 provides the proof of equations (1.26) and (1.27).

In Section 1.3.3, we will empirically examine the size of $U_{t,t+1}$ based on equations (1.21), (1.26), and (1.27). Appendix 1.B.1 provides supplementary materials on the estimation of $U_{t,t+1}$.

1.2.3.2 $T_{t,t+1}$ once mid-option prices are used

Under the standard specification of the transaction cost function, equation (1.5), we can see from equation (1.12) that the true *theoretical* value of the stock's CFER ($CFER_{t,t+1}$) does not depend on the transaction costs for trading options, because $\partial TC_t / \partial \theta_t^S$ does not depend on them. On the other hand, our model-free *estimator* of CFER ($CFER_{t,t+1}^{MF}$) deviates from the true CFER value when the transaction costs for trading options exist; the $T_{t,t+1}$ term in equation (1.14) represents this difference.

[9] Even though each option exchange can have a different margin rule, Hitzemann et al. (2017) document that the CBOE margin rule is the de facto standard margin rule in the U.S. option exchanges.

Since the relative transaction costs for trading options are known to be much bigger than those for trading stocks, this would imply that the $T_{t,t+1}$ can be non-negligible and hence using $CFER^{MF}_{t,t+1}(K)$ as an estimate of $CFER_{t,t+1}$ may suffer from non-negligible bias *if* traded option prices are used to calculate $CFER^{MF}_{t,t+1}$.

Contrary, in the subsequent empirical analysis, we employ *mid*-option prices (the average of bid and ask option prices) because the option dataset we will employ (the OptionMetrics database) provides option bid and ask quotes but does not provide traded option prices. This results in $T_{t,t+1}$ not being required to estimate CFER because we find that option transaction costs have negligible effect on the mid option prices. We show this as follows. First, in line with Muravyev (2016), we model the bid and ask option prices as the sum of the ex-transaction costs option price (e.g., $C_t^{exTC}(K) = \mathbb{E}_t^{\mathbb{P}}[m^*_{t,t+1}(S_{t+1} - K)^+] + M_t^c(K)$), plus the three components of the bid-ask spread (fixed costs, adverse selection part, inventory risk part) and a zero-mean market microstructure noise term. Under standard assumptions employed in the market microstructure literature, we prove that the mid-option price equals approximately the ex-transaction costs option price; Appendix 1.B.2.1 provides the details. Then, in Appendix 1.B.2.2, we show via simulation that the error of the approximation is negligible compared to the magnitude of observed deviations from put-call parity. This implies that the $T_{t,t+1}$ term does not affect option mid prices and hence it does not need to be known to estimate CFER.

Intuitively, our result that $T_{t,t+1}$ is not required to estimate CFER when mid-option prices are used, comes at no surprise. Our theoretical approach effectively decomposes the effect of transaction costs to the option price (last term in the right hand side of equation (1.10)) to the standard components of the bid-ask spread described above. Given the evidence that a major component of the option bid-ask spread is fixed costs (e.g., Huang and Stoll, 1997; Engle and Neri, 2010), which does not affect the mid-price (Muravyev, 2016, p.678), the mid-option price proxies the ex-transaction costs option price.

1.2.4 Relation to measures of deviations from put-call parity

Measures of deviations from put-call parity such as the implied volatility spread (IVS) defined as the difference between the implied volatility (IV) of the call and put options with the same strike and maturity (e.g., Bali and Hovakimian, 2009; Cremers and Weinbaum, 2010) and the DOTS measure of Goncalves-Pinto et al. (2019) have been documented to predict stock returns. We show that (i) $CFER^{MF}_{t,t+1}(K)$ may contain richer (and certainly no less) information than IVS and DOTS regarding future stock returns, and (ii) neither $CFER^{MF}_{t,t+1}(K)$ nor other measures of deviations from put-call parity should be interpreted as a measure of whether the stock is mispriced.

Proposition 1.2.1. *Let $IV^c_t(K)$ and $IV^p_t(K)$ be the Black-Scholes call and put implied volatilities (BS-IVs), respectively. Then, the following approximate equation holds.*

$$CFER^{MF}_{t,t+1}(K) \approx \frac{R^0_{t,t+1}\mathcal{V}_t(K)}{S_t}(IV^c_t(K) - IV^p_t(K)), \tag{1.28}$$

where $\mathcal{V}_t(K)$ is the Black-Scholes vega evaluated at a strike K and a volatility equal to $(IV^c_t(K) + IV^p_t(K))/2$.

Proof. See Appendix 1.A.2. □

Proposition 1.2.2. *Let η^c_t and η^p_t be the early exercise premium of the American call and put option, respectively. Then, the following relation holds:*

$$DOTS_t(K) = \frac{CFER^{MF}_{t,t+1}(K)}{R^0_{t,t+1}} + u_t,\ u_t = \frac{1}{S_t}\left[\eta^c_t - \frac{D_{t+1}}{2R^0_{t,t+1}} - \left(\eta^p_t - \frac{K(R^0_{t,t+1}-1)}{2R^0_{t,t+1}}\right)\right]. \tag{1.29}$$

where

$$DOTS_t(K) := \frac{1}{S_t}\left(\frac{S^U_t(K) + S^L_t(K)}{2} - S_t\right), \tag{1.30}$$

and $S^U_t(K) = C^{ask}_t(K) - P^{bid}_t(K) + K + D_{t+1}/R^0_{t,t+1}$ and $S^L_t(K) = C^{bid}_t(K) - P^{ask}_t(K) + K/R^0_{t,t+1}$ are the no-arbitrage bounds for the stock price (i.e., $S^L_t \le S_t \le S^U_t$) calculated from the bid and ask prices of American call and put options (C^{bid}_t, P^{bid}_t, C^{ask}_t, and P^{ask}_t) with strike K.

Proof. See Appendix 1.A.3. □

Propositions 1.2.1 and 1.2.2 show that IVS and DOTS are approximately proportional to $CFER^{MF}_{t,t+1}$. Therefore, they provide a theoretical friction-based explanation for the empirically documented ability of IVS and DOTS to predict future stock returns. Moreover, these results explain formally Goncalves-Pinto et al.'s (2019) finding that DOTS and IVS are highly correlated. However, $CFER^{MF}_{t,t+1}$ may outperform (and certainly not underperform) IVS and DOTS as a cross-sectional predictor of stock returns. This is because IVS and DOTS are only approximations of $CFER^{MF}_{t,t+1}$. Our theoretical model suggests that $CFER^{MF}_{t,t+1}$ is the most appropriate way to utilize the informational content embedded in deviations from put-call parity. Intuitively, the "scaling coefficient" $R^0_{t,t+1}/S_t$ of $CFER^{MF}_{t,t+1}$ converts the deviation in *prices* $(\widetilde{S}_t(K) - S_t)$ to a return metric, hence making the model-free CFER a *quantitative* measure of the effect of market frictions on expected stock *returns*.[10] This is not the case with the two empirical measures of deviations from put-call parity, because IVS is denominated in (difference in) volatility and Goncalves-Pinto et al. (2019) *interpret* DOTS as the level of mispricing (relative to the stock price). The strength of the predictive power of IVS will depend on the size of the approximation error in equation (1.28) and on the impact of omitting the vega scaling factor, and that of DOTS will depend on the size of u_t which is a function of the early exercise premium of the American call and put options.

The distinction between $CFER^{MF}_{t,t+1}$ and previous measures of deviations from put-call parity reveals a common misconception and highlights the interpretation of $CFER^{MF}_{t,t+1}$. Deviations from put-call parity should not be interpreted as a measure of the current level of mispricing in the underlying stock price, that is, the synthetic stock price should not be interpreted as a measure of the "fundamental" price. This can be seen by the fact that the synthetic stock price is "contaminated" by the effect of frictions to stock prices even when option prices equal their IMRS-discounted payoff value. In this case, the synthetic stock price (equation (1.18)) equals the expected IMRS-discounted value of the payoff at time $t+1$, $S_{t+1}+D_{t+1}$, where S_{t+1} is affected by the market frictions (this can be seen by considering (1.9) for time $t+1$). Equivalently stated, a positive (negative) $CFER^{MF}_{t,t+1}$ should not be interpreted as evidence that the stock is underpriced (overpriced). By definition, $CFER^{MF}_{t,t+1}$, proxies the part of expected *return*, which is not explained by the covariance risk premium and hence it has an alpha interpretation, that is, positive (negative) $CFER^{MF}_{t,t+1}$ signifies that the

[10] $CFER^{MF}_{t,t+1}(K)$ can be rewritten as $CFER^{MF}_{t,t+1}(K) = R^0_{t,t+1}(\widetilde{S}_t(K)/S_t - 1)$. Then, dividing equation (1.18) by S_t yields the relation $\widetilde{S}_t(K)/S_t = \mathbb{E}^{\mathbb{P}}_t[m^*_{t,t+1}R_{t,t+1}] - (U_{t,t+1}(K) + T_{t,t+1}(K))/R^0_{t,t+1}$, that is, the ratio of the two stock prices $\widetilde{S}_t(K)/S_t$ is a function of the underlying stock return.

stock will outperform (underperform) in the future.[11]

1.2.5 Properties of CFER and testable hypotheses

Our CFER formula allows us to test the theoretical properties of CFER demonstrated by the following three hypotheses.

1.2.5.1 Ability of CFER to predict stock returns

HYPOTHESIS 1. *The expected asset return is increasing with CFER.*

Equation (1.11) shows that the greater CFER is, the greater the asset's expected return. Hence, we expect that when we sort stocks in portfolios based on their respective estimated CFER, (i) the post-ranking portfolios' average return and CFER will be positively related, and (ii) a long-short spread portfolio where we go long in the portfolio which contains the stocks with the highest CFER and short in the portfolio which contains the stocks with the lowest CFER should earn a positive average return and alpha. Moreover, our discussion in Section 1.2.4 suggests that CFER is expected to have superior (or at least as good) predictive power compared to the IVS and DOTS measures of deviations from put-call parity (equations (1.28) and (1.29)).

HYPOTHESIS 2. *In the case where the IMRS m^* is described by a linear combination of risk factors, then the regression of the* CFER-adjusted excess return*,* $R_{t,t+1} - R^0_{t,t+1} - CFER_{t,t+1}$ *on the risk factors should yield a zero intercept.*

It is valid to subtract CFER from excess returns to consider the *CFER-adjusted excess return*, $R_{t,t+1} - R^0_{t,t} - CFER_{t,t+1}$ since CFER is a quantitative measure of part of the expected return. In the case where the IMRS m^* is described by a linear combination of risk factors $\boldsymbol{f}$, then our asset pricing equation (1.11) implies that

$$\mathbb{E}^{\mathbb{P}}_t[R_{t,t+1}] - R^0_{t,t+1} - CFER_{t,t+1} = \boldsymbol{\beta}'_R \mathbb{E}^{\mathbb{P}}_t[\boldsymbol{f}_{t+1}], \tag{1.31}$$

where $\boldsymbol{\beta}_R$ is the vector of the factor betas. Next, consider the following regression model:

$$R_{t,t+1} - R^0_{t,t+1} - CFER_{t,t+1} = \alpha + \boldsymbol{\beta}' \boldsymbol{f}_{t+1} + \varepsilon_{t+1}. \tag{1.32}$$

[11] A stock which is underpriced (overpriced) does not necessarily mean that it will outperform (underperform) in the future unless the friction which causes the mispricing is relaxed.

Then, the intercept α of equation (1.32) should be zero under the second hypothesis. This theoretical result is a generalization of the expected return-beta representation theorem from the case of frictionless markets (Cochrane, 2005, Chapter 12.1) to the case where market frictions and hence a non-zero CFER exist. Hypothesis 2 provides a guide to detect the source of (empirically estimated) alphas of the CFER-sorted long-short portfolio by circumventing the joint hypothesis problem (Fama, 1991).[12] In the case where a CFER-sorted long-short portfolio earns a significant positive alpha (Hypothesis 1), yet one fails to reject Hypothesis 2, this would suggest that the significant alpha of CFER-*non*-adjusted return stems from the presence of market frictions rather than from omitted risk factors in the model employed to calculate the alpha.

1.2.5.2 Relation of CFER with transaction costs and margin constraints

Up to this point, our CFER measure has a "black-box" property; it subsumes the effect of all relevant market frictions. Now, we introduce additional structure in our model to establish the theoretical relation between CFER and specific type of frictions. We consider three types of market frictions: transaction costs (equation (1.5)), margin constraints (equation (1.7)), and short-sale constraints. For simplicity, we assume that the stock's proportional transaction costs, ρ_t^S, takes one of the following two values depending on whether the agent short-sells the stock (i.e., $\theta_t^S < 0$ and $\Delta\theta_t^S < 0$) or not:

$$\rho_t^S = \begin{cases} \rho + \varsigma & \text{if } \theta_t^S < 0 \text{ and } \Delta\theta_t^S < 0, \\ \rho & \text{otherwise,} \end{cases} \tag{1.33}$$

where $\varsigma > 0$ reflects the severity of short-sale constraints. The finite value of ς expresses the fact that the transaction costs for short-selling a stock is greater compared to buying a stock due to stock lending fees (e.g., D'Avolio, 2002) and search costs (e.g., Duffie et al., 2002). The short-sale ban corresponds to the case $\varsigma = \infty$. The following Proposition provides the relation between the range of CFER value, and the margin constraints, transaction costs and short-sale constraints.

[12] Note that a zero intercept is a necessary but not a sufficient condition for asset pricing models to be valid. The latter would require testing whether the factors are priced something which is beyond the scope of this study since our focus is not on testing asset pricing models.

Proposition 1.2.3. *Suppose that three types of frictions, transaction costs, margin constraints, and short-sale constraints are present and modeled by equations (1.5), (1.7), and (1.33), respectively. Then, CFER satisfies the following approximate inequality range:*

$$-2\rho - \varsigma - \frac{\lambda_t^{MC}}{u'(c_t)}\mu_t^S \lesssim CFER_{t,t+1} \lesssim 2\rho + \frac{\lambda_t^{MC}}{u'(c_t)}\mu_t^S, \tag{1.34}$$

The approximate width of the range is given by

$$2\frac{\lambda_t^{MC}}{u'(c_t)}\mu_t^S + 4\rho + \varsigma. \tag{1.35}$$

Proof. See Appendix 1.A.4. □

Proposition 1.2.3 yields four predictions formulated under Hypothesis 3. First, stocks with more extreme, ether positive or negative, (smaller) CFER values are subject to greater (smaller) market frictions. This suggests a U-shape relation between CFER and proxies of market frictions. Second, equation (1.34) suggests that, higher transaction costs, tighter margin constraints and severer short-sale constraints imply larger variations in the cross-section of CFER. Third, the more severe the short selling constraint is, the more negatively skewed the distribution of CFER values will be. Therefore, at least over negative CFER sub-samples (which imply the agent sells the stock since she demands a more negative expected return), severe short-sale constraints imply a lower CFER. Finally, expression (1.35) is the theoretical approximate upper bound of the "alpha" (expected excess return adjusted for risk premium) of the CFER-sorted spread portfolio; the maximum alpha would be attained when one buys the stocks with the highest CFER and shorts the stocks with the lowest CFER. The upper bound of the alpha of the spread portfolios is approximately given by the sum of twice the round-trip transaction costs, twice the effect of margin constraints and the short-sale cost.

HYPOTHESIS 3. *CFER and market frictions are related as follows:*

3-1: *An extreme CFER value implies the presence of large frictions.*

3-2: *Larger frictions imply greater cross-sectional variation in CFER.*

3-3: *The more severe the short-sale constraint, the more negative CFER becomes.*

3-4: *Expression (1.35) provides the upper bound of CFER-sorted long-short portfolio's alpha.*

We examine the conjectures of Hypothesis 3 in Sections 1.3.4 and 1.5.

Note that in contrast to CFER, other return predicting variables such as IVS and DOTS, cannot yield Hypothesis 2 and 3 because they are not a *quantitative* measure of a part of expected returns.

1.3 Data and CFER estimation

1.3.1 Data sources

We obtain end-of-month U.S. equity option prices and implied volatilities (IVs) from the OptionMetrics Ivy DB database (OM) via the Wharton Research Data Services. Our dataset spans January 1996 to April 2016 (244 months). Options written on the U.S. individual equities are American style. OM calculates IVs via the Cox et al. (1979) binomial tree model, which takes the early exercise premium of American options into account. Given that our formula to estimate CFER (Theorem 1.2.2) relies on European option prices, we convert OM-IVs to the corresponding European option prices via the Black and Scholes (1973) option pricing formula.[13] We also obtain the risk-free rate and dividend payment data from the OM database to calculate the present value of dividend payments over the option's life time. We remove IVs if the recorded corresponding option bid price is non-positive, the IV is missing, and the option's open interest is non-positive. We discard data with time to maturity shorter than 8 days or longer than 270 days. We keep option data only when the moneyness K/S_t is between 0.9 and 1.1 to ensure that the most liquid option contracts are considered.

We obtain stock returns from the Center for Research in Security Prices (CRSP). In line with the literature, our stock universe consists of all U.S. common stocks (CRSP share codes 10 and 11). We obtain the time-series of risk factors in the CAPM, Fama and French (1993) 3-factor model (FF3), Carhart (1997) 4-factor model (FFC), and Fama and French (2015) 5-factor model (FF5) from Kenneth French's website.

[13]This approach is often taken in the literature (see e.g., Martin and Wagner, 2018).

We obtain the factors in the Stambaugh and Yuan (2017) mispricing factor model (SY) from Yu Yuan's website. We construct various firms' and stocks' characteristics variables (e.g., size, book-to-market, bid-ask spread) based on CRSP and Compustat database. Appendix 1.C provides the definition and the data source of the various characteristics variables.

1.3.2 Estimation of CFER: Choice of strikes and maturities

We estimate both the "model-free" CFER (MF-CFER) and "fully-estimated" CFER (full-CFER). MF-CFER is estimated as the model-free scaled deviations from put-call parity term, whereas we estimate full-CFER as the sum of MF-CFER and the estimated U term (the second term in the right-hand side of equation (1.14)). In the empirical estimation, both MF-CFER and the U term depend on the choice of strike K and maturity T.[14] For simplicity, we describe how we choose K and T to estimate MF-CFER; we apply the same procedure to calculate the U term and hence the full-CFER.

Regarding the choice of K, on each end-of-month date t and for each traded option maturity T for which the call and put IVs are available, we take the weighted average of the scaled deviations from put-call parity across strikes. We set the weight to be the open interest of the corresponding options in line with Cremers and Weinbaum (2010) to reduce possible measurement errors in the options data. As a robustness check, we also examine the MF-CFER of the traded option with strike price closest to the forward price $f_{t,T} = R^0_{t,T}(S_t - PVD_{t,T})$, where $PVD_{t,T}$ is the present value of dividend payments during $[t+1, T]$.

Regarding the choice of the options' maturity, we proceed as follows. The horizon of the estimated CFER should correspond to the horizon of expected returns. Given that in the subsequent analysis we will be considering monthly returns, we construct a 30-day constant maturity MF-CFER as follows. First we multiply each average MF-CFER by $30/dtm$, where dtm denotes days-to-maturity. Then, we linearly interpolate the average MF-CFER obtained from the two traded maturities surrounding the 30-day maturity to construct the 30-day constant maturity MF-CFER. Note that a

[14] Even though the left-hand side of equation (1.14) does not depend on K, the empirically estimated sum of MF-CFER plus U may vary across K because the estimated U term may contain estimation errors.

similar interpolation is employed in the CBOE's calculation of the VIX index, which represents the model-free implied volatility over the next 30 days. The estimated CFER is treated as missing if the 30-day maturity is not bracketed by two traded maturities. As a robustness check, we also use the MF-CFER closest to 30-days to-maturity options. We calculate this proxy only when the closest options' maturity is between 15-day and 45-day, otherwise we treat the estimated CFER as missing.

In sum, we have two ways on the choice for K, averaged across strikes (AVE) versus closest to forward-ATM (ATM), and two ways on the choice for T, linearly interpolated 30-day constant maturity CFER (CM) versus closest to 30-day (CLS). Thus, there are in total four choices for the pair of K and T labeled, AVE-CM, ATM-CM, AVE-CLS, and ATM-CLS. Among these four choices, we use the AVE-CM as the baseline for the purposes of our subsequent analysis, yet its estimation result is highly correlated with the estimates based on other three ways to estimate CFER.

1.3.3 Estimation of CFER: MF-CFER versus full-CFER

In this subsection, we estimate the full-CFER by alternative models and compute its correlation with the estimated MF-CFER to shed light on whether CFER can be estimated accurately by relying only on the model-free scaled deviations from put-call parity. We estimate the former by estimating the U term separately, for the embedded leverage and margin constraint models described in Section 1.2.3 for each pair of call and put options. Then, we construct the AVE-CM (average across strikes, constant maturity) embedded leverage-based full-CFER (EL-CFER) and the margin constraints-based full-CFER (MC-CFER) as discussed in Section 1.3.2.

To estimate $U^{EL}_{t,t+1}(K)$, we use equation (1.21) and set the coefficient on the embedded leverage k to -1.25% per month by following Frazzini and Pedersen (2012). For option deltas, we use the deltas provided by the OM database.

To estimate $U^{MC}_{t,t+1}(K)$, we use equations (1.26) and (1.27). We separately estimate $R^0_{t,t+1}\lambda^{MC}_t/u'(c_t)$ and $E_t(K)$, then take their product. Regarding the term $R^0_{t,t+1}\lambda^{MC}_t/u'(c_t)$, note that it corresponds to the shadow cost of capital, which is shown to be equal to the spread between the uncollateralized and collateralized risk-free bond rates (see Gârleanu and Pedersen, 2011).[15] Gârleanu and Pedersen (2011)

[15]While Gârleanu and Pedersen's (2011) model is a continuous-time one, we formally show this

find that the shadow cost of capital is time-varying and becomes higher during market distress periods reaching about 10% per year during the recent financial crisis (p.1982 and Figure 1). As a result, we examine three alternative values for $R^0_{t,t+1}\lambda^{MC}_t/u'(c_t)$: 10% per year, 5% per year (as a proxy of the time-series average of the shadow price of capital), and the scaled TED spread (TED spread multiplied by 3.3) to take into account the time-varying nature of the shadow cost of capital and its reported maximum value of 10%.[16]

Regarding the calculation of $E_t(K)$, equation (1.27) shows that it depends on the signs of $\theta^c_t(K)$ and $\theta^p_t(K)$. We infer the signs of these option positions as follows. In line with Gârleanu et al. (2009) and Hitzemann et al. (2017), we assume that there are two types of agents, the market-maker and the end-user and we also assume that the market-maker is the marginal investor considered in Section 1.2.1. Let $d^c_t(K)$ and $d^p_t(K)$ be the end-user's demand for the call and put option, respectively. Then, at equilibrium, $sgn(\theta^c_t) = -sgn(d^c_t)$ and $sgn(\theta^p_t) = -sgn(d^p_t)$ hold because options are in zero net supply. We infer the signs of the end-user's demand instead of the market-maker's position and take the opposite signs. To infer the signs of the end-user's demand, we follow Gârleanu et al. (2009) and Hitzemann et al. (2017) and use the options' *expensiveness*, which they proxy by the difference between the historical volatility and the implied volatility. In particular, we assume that the options' expensiveness is above (below) the reference point s if and only if the end-user's demand $d_t(K)$ is positive (negative):

$$expensiveness^i_t(K) < s \;\Leftrightarrow\; d^i_t(K) < 0 \;\Leftrightarrow\; \theta^i_t(K) > 0, \quad i \in \{c, p\}. \tag{1.36}$$

We estimate the expensiveness measure (i.e., the left hand side of equation (1.36)) as the difference between the Black and Scholes (1973) implied volatility and the one-year historical volatility, both of which are provided by the OM database. To select the value of threshold s, we rely on Gârleanu and Pedersen's (2011) and Christoffersen et al.'s (2018) finding that the end-user is a net seller of equity options, which implies

result in our discrete time model by introducing the uncollateralized bond in Appendix 1.B.1.2.

[16] Note that Gârleanu and Pedersen (2011) regress the estimated shadow price of capital on the TED spread to obtain the coefficient on the TED spread about 1.8, whereas we multiply the TED spread by the factor of about 3.3 to match the 10% maximum value. Our choice of the scaling factor is conservative for our robustness check purposes, because it results in a greater (absolute) value for $U^{MC}_{t,t+1}$.

that there are more "cheap" options than "expensive" options. This suggests that the proportion of options whose expensiveness is below s should be greater than 50%. Given this finding, we examine three values for $s = 0$, $s = 0.01$ and $s = 0.02$ which, under the estimation rule given by equation (1.36), yield that roughly 50%, 55%, and 62% of the options in our sample are "cheap," respectively.

In sum, we examine three alternative values for $R^0_{t,t+1}\lambda^{MC}_t/u'(c_t)$ and three alternative values for $E_t(K)$ (depending on three reference point values s to determine the signs of $\theta^c_t(K)$ and $\theta^p_t(K)$). This yields nine estimates of $U^{MC}_{t,t+1}(K)$ and hence nine estimates of fully estimated MC-CFER in total. Once we estimate $U_{t,t+1}(K)$ for each strike K and maturity T, we calculate the average U term, $\overline{U}_{t,t+1}$, under the AVE-CM specification described in Section 1.3.2. $\overline{U}_{t,t+1}$ is the difference between AVE-CM MF-CFER and AVE-CM full-CFER. Two are the main findings regarding the question whether CFER can be accurately proxied by the estimated MF-CFER. First, the distribution of $\overline{U}_{t,t+1}$ is fairly symmetric around zero. This holds regardless of how we model frictions (embedded leverage or option margins) and it suggests that the model-free proxy of CFER is on average accurate. Second, the magnitude of $\overline{U}_{t,t+1}$ per se is also small. This is because we find that the AVE-CM MF-CFER and AVE-CM full-CFER are almost perfectly correlated (correlation above 0.96) for both the embedded leverage and the margin constraints models. In the latter model, the pairwise correlations between the AVE-CM MF-CFER and each one of the nine AVE-CM MC-CFER are at least 0.998.

The above results suggest that the effect of frictions on the synthetic stock via the U term is negligible; in Section 1.4.2, we provide further supporting evidence. Moreover, our argument in Section 1.2.3.2 and the results in Appendix 1.B.2 suggest that the $T_{t,t+1}$ term in equation (1.20) does not affect the estimation of CFER which relies on option mid-prices either. Therefore, we can reliably estimate CFER by using only the model-free part of the CFER formula and in the remaining of this Chapter we will report results based on the MF-CFER unless stated otherwise.

Note that the fact that $\overline{U}_{t,t+1}$ is negligible suggests that the theoretical CFER value, $CFER_{t,t+1}$, takes a non-negligible value if and only if the model-free estimate of CFER, $CFER^{MF}_{t,t+1}$, takes a non-negligible value. Therefore, our findings suggest that it is not likely that the true CFER value is zero (i.e., the expected stock returns are driven purely by risk premia) yet deviations from put-call parity are non-zero because of a non-zero $\overline{U}_{t,t+1}$ term (i.e., option prices are affected by frictions). Similarly, our findings suggest that it is unlikely that deviations from put-call parity may be zero but $CFER_{t,t+1}$ is large (i.e., frictions have a large effect on the expected stock return), because $\overline{U}_{t,t+1}$ needs to take a large and similar value to $CFER_{t,t+1}$ for this case to occur.

A final remark is in order. Our findings suggest that market frictions have a negligible effect on $\overline{U}_{t,t+1}$. This should not be interpreted as evidence that the option market is frictionless, though. We reconcile our finding with the previous empirical evidence that market frictions affect option returns (Santa-Clara and Saretto, 2009; Frazzini and Pedersen, 2012; Hitzemann et al., 2017) based on the following two clarifications. First, the previous literature implies that the ratio of the effect of frictions on options over the respective option prices ($M_t^c(K)/C_t(K)$ and $M_t^p(K)/P_t(K)$) are non-negligible. On the other hand, the size of $U_{t,t+1}(K)$ is determined by the ratios $M_t^c(K)/S_t$ and $M_t^p(K)/S_t$, which are much smaller than the former ratios because the denominator (the stock price) is much greater than option prices. Second, the effect of frictions on the synthetic stock price $M_t^c(K) - M_t^p(K)$ can be smaller than that on individual options (i.e., $M_t^c(K)$ and $M_t^p(K)$). This would occur if $M_t^c(K)$ and $M_t^p(K)$ have the same sign and are of similar size. Then, they would offset each other when their aggregate effect on the synthetic stock is calculated. This is always the case for the embedded leverage model and mostly true for the margin constraints model. In Appendix 1.B.1.3, we discuss these points in detail.

1.3.4 CFER: Summary statistics

Table 1.1, Panel A, reports the summary statistics of the estimated MF-CFER at the end of each month for each one of the four ways of estimating CFER. We can see that there are about 333,000 stock-month CFER observations for the case of the AVE-CM and ATM-CM CFER, whereas this number increases to about 347,000 observations for the case of the AVE-CLS and ATM-CLS CFER. This yields on average about 1,370 (1,420) stocks in each month in the case of AVE-/ATM-CM CFER (AVE-/ATM-CLS CFER). This is a sufficient number to form well diversified decile portfolios in the subsequent analysis. The mean and the median of the estimated CFER are about

-0.1% and -0.04% per month (30-day), respectively. Results are similar across the four construction methods of CFER. The distribution of CFER is skewed to the left and it is highly leptokurtic. The estimated CFER is sizable; it takes both positive and negative values, ranging from -1.24% to 0.89 % per month (-14% to 11% per year) in a 5th to 95th percentile range of AVE-CM CFER. Note that this range is consistent with the theoretically derived CFER bounds, equation (1.34). Lesmond et al. (1999) and Hasbrouck (2009) document that the round-trip transaction costs (i.e., 2ρ in equation (1.34)) for large stocks are in the order of 1.0%. Moreover, the fact that the distribution of the estimated CFER has longer left tail is consistent with the existence of short-sale constraints (equation (1.34)).

CFER also has fairly large variations; the standard deviation is about 1% and the interquartile range (IQR, the difference between 75th and 25th percentile points) is between 47–60 bps on average across stocks over time depending on the CFER construction method. This magnitude of variation is relatively large compared to the long-run average U.S. equity risk premium, which is about 50 bps per month (or 6% per year, see e.g., Mehra, 2012). The percentage of the negative observations of CFER is about 55% in any of the four construction ways of CFER; CFER takes more often negative than positive values over the full sample period. Table 1.1, Panel B, reports that the four ways of computing CFER are almost perfectly correlated. Therefore, the subsequent analysis is expected to be robust to the choice of the method to estimate the 30-day CFER.

[Table 1.1 about here.]

Our theoretical considerations in Section 1.2 suggest that the summary statistics of the estimated CFER should be related to various proxies of market frictions as well as other option-based measures of deviations from put-call parity. Regarding proxies of market frictions, we will extensively investigate their relation with the estimated CFER in Section 1.5. Here, we discuss the relation between the estimated CFER, DOTS, and IVS. The summary statistics of the estimated CFER are in line with those of these alternative measures based on deviations from put-call parity. For example, Cremers and Weinbaum (2010) report that the median of IVS is negative (see equation (1.28)). They also report that the IVS are highly volatile and exhibit substantial cross-sectional variation. Goncalves-Pinto et al. (2019) document that their DOTS measure has a negative median value and exhibits large variations.

Our theoretical results in Section 1.2.4 yield more specific predictions regarding the estimated values of CFER, IVS, and DOTS. We show that our model-free CFER is (approximately) proportional to IVS and DOTS and hence the range of the estimated CFER and that of IVS and DOTS should also be proportional. We start with comparing the range of the estimated CFER and DOTS. Proposition 1.2.2 shows DOTS and MF-CFER are approximately proportional by the factor of the gross risk-free rate, which is very close to one. Therefore, we expect that the ranges of the estimated CFER and DOTS are similar.

Goncalves-Pinto et al. (2019) report the average DOTS of the DOTS-sorted decile portfolios. They report that the average DOTS of the bottom and the top DOTS-sorted decile portfolios are -0.92% and 0.79%, respectively. We repeat the same exercise by using AVE-CM CFER as a sorting variable. The average AVE-CM CFER of the bottom and the top decile portfolios are -1.31% and 0.93%, respectively. Therefore, the ranges of the average DOTS and AVE-CM CFER are comparable, although the range of DOTS is slightly narrower. Note that the narrower range of DOTS originates from the fact that Goncalves-Pinto et al. (2019) do not scale DOTS (equation (1.30)) by time-to-maturity. Since they use option data whose maturity are no longer than one month (hence the average maturity across their dataset is shorter than one month), the unscaled DOTS generally takes a smaller value in magnitude compared to AVE-CM CFER, which we scale to denote 30-day return.

Next, we compare IVS and MF-CFER. Goncalves-Pinto et al. (2019) also document the average IVS of the DOTS-sorted decile portfolios and the that of the bottom and the top decile portfolios are -12.2% and 9.5%, respectively. This means that the range of AVE-CM CFER is approximately ten times narrower than that of IVS. This is consistent with Proposition 1.2.1, which shows that MF-CFER is proportional to IVS by the factor of $R^0_{t,t+1}\mathcal{V}_t(K)/S_t$; for one-month at-the-money options, this scaling coefficient is approximately equal to 0.1. These results suggest that the distribution of the estimated CFER is aligned with the existing measures of deviations from put-call parity in a theoretically predicted manner.

[Figure 1.1 about here.]

Next, we look at the time-series evolution of the estimated CFER. Figure 1.1a shows the time-series evolution of the monthly median, 25th percentile point and 75th percentile point of AVE-CM CFER. The monthly median had been mostly negative until around 2007. Accordingly, the proportion of negatively estimated CFER in each month had been about 60% to 70% until around 2007, which is higher than that of the whole observations from 1996 to 2016 (about 55%). The proportion of negatively estimated CFER before 2007 is in line with Ofek et al. (2004), who use the data between 1999 and 2001, in that they also report that the underlying stock price is greater than the synthetic stock price for about two thirds of their observations.

The monthly median CFER experienced a huge drop during the height of the financial crisis (September to November 2008). This is in line with the temporary short-sale ban during the market meltdown. The short-sale ban implies that CFER may take a very negative value, implied by equation (1.34) with $\varsigma = \infty$. This finding is in line with Grundy et al. (2012) in that their empirical results show that underlying stock prices became significantly more expensive compared to synthetic stock prices (i.e., negative model-free CFER) more frequent for banned stocks during the ban period. After the financial crisis, the monthly median of the estimated CFER clearly changed its property; the median value fluctuates around zero and it frequently takes positive value. Unfortunately, as our model does not yield predictions on the average value of CFER, investigating reasons behind this change is out of the scope of this study.

Figure 1.1b shows the time-series evolution of the monthly IQR of AVE-CM CFER. We measure the dispersion of the estimated CFER by using this statistic rather than the standard deviation because the distribution of CFER is skewed and leptokurtic. As we have discussed in Section 1.2.5.2, the degree of the dispersion in CFER is determined by the size of transaction costs, margin constraints and short-sale constraints among other frictions. The time-series fluctuations in the IQR are in line with this prediction: most of the spikes in the IQR correspond to market turmoils, such as Russian default and LTCM crisis (August to September 1998), the collapse of Lehman Brothers and ensuing market meltdown (September to November 2008), European debt crisis (November 2011, uncertainty was the highest around the general election in Greece), and the Chinese stock market turmoils (June 2015 to January 2016). It is well documented that the margin (and liquidity) constraints become tighter during market turmoil periods (Gârleanu and Pedersen, 2011; Nagel, 2012). Hou et al. (2016) also find that their microstructural friction measure takes greater values during recessions and market distress periods. On the other hand, except these distressed periods, we can see a secular decline in the size of the IQR. This may reflect the decrease in transaction costs. For example, Green et al. (2017) document that the decimalization of quotes in 2001 and the introduction of autoquoting software by the NYSE in 2003 dramatically reduced transaction costs. According to Proposition 1.2.3, a reduction in transaction costs implies that both the upper bound and lower bound of CFER shrink toward zero. The behavior of the 25th and 75th percentile points of AVE-CM CFER in Figure 1.1a supports this conjecture; both the monthly 25th and 75th percentile points time-series generally shrunk toward zero from the late 1990s to the early 2000s.

Finally, note that the calculation of CFER requires the existence of market option prices. As a result, our universe of stocks is confined to the optionable stocks (i.e., stocks which have options written on them). However, this should not be viewed as a shortcoming of this study. In line with the results of Cremers and Weinbaum (2010), our optionable stocks are big stocks; the average market capitalization of stocks with (without) AVE-CM CFER is about 9.1 billion (0.5 billion) U.S. dollars over our sample period from 1996 to 2016. Relatedly, albeit we can estimate AVE-CM CFER for about 27% of stocks (about 1,350 optionable stocks out of 5,000 all common stocks in each month), these stocks on average account for about 90% of the aggregate market capitalization of U.S. common stocks over our sample period. In addition, even though our cross-section of U.S. optionable stocks is subject to smaller frictions compared to the non-optionable stock universe, still the effect of market frictions on their expected returns (i.e. CFER) is sizable as reported and we will further demonstrate in Section 1.4.[17]

[17]For example, the average Amihud's (2002) illiquidity measure of the optionable (non-optionable) stocks is 0.01 (5.28), and the average relative bid-ask spread of optionable (non-optionable) stocks is 0.48% (2.50%) over our sample period.

1.4 CFER as a cross-sectional predictor: Evidence

In this Section, we examine whether CFER predicts equity returns cross-sectionally by taking a portfolio sorting approach (Hypothesis 1 discussed in Section 1.2.5.1). We sort stocks in decile portfolios by using the estimated CFER as a sorting criterion. Portfolio 1 (10) contains the stocks with the lowest (highest) CFER. We form portfolios at the end of each month. Then, we calculate the post-ranking monthly return of each portfolio and the zero-cost long-short spread portfolio, where we go long in Portfolio 10 and short in Portfolio 1. Our testable hypothesis suggests that this zero-cost long-short portfolio will earn a positive average return. We also compare the predictive power of MF-CFER and full-CFER. Finally, we test the second hypothesis discussed in Section 1.2.5.1 and we compare the cross-sectional predictive ability of the MF-CFER with that of IVS and DOTS.

1.4.1 Predictive power of model-free CFER

Table 1.2 reports the results for both the value-weighted and equally-weighted decile portfolios cases, where we use the AVE-CM MF-CFER as a sorting variable. In line with the model's prediction, we can see that there is a monotonically increasing relation between the portfolios' average returns and CFER. Moreover, the average return of the long-short value-weighted spread portfolio is 1.64% per month. This value is sizable and statistically significant (t-stat: 5.77). We also calculate the risk-adjusted returns, in terms of alpha with respect to the CAPM and FFC model.[18] Both α_{CAPM} and α_{FFC} are sizable and statistically significant; α_{CAPM} is 1.70% and α_{FFC} is 1.86% per month and their t-statistics are above five which is above the threshold proposed by Harvey (2017) for the purposes of addressing data snooping concerns. These results show that the estimated CFER predicts future stock returns over and above other well known risk factors.

The equally-weighted portfolio earns an even more significant average return compared to the value-weighted portfolio; the average return is 1.73% per month, α_{CAPM} and α_{FFC} are 1.76% and 1.81% per month, respectively and t-statistics are above nine.

[18] For all portfolio sort exercises in this Section, we also estimate alphas with respect to FF3, FF5, and SY models. Results are qualitatively similar and hence we do not report them due to space limitations.

Even though the equally-weighted result is stronger than the value-weighted result, in the subsequent analysis, we focus on the value-weighted results. This is because, as the value-weighted construction tends to result in lower alphas and t-statistics, our judgment on the predictive ability of CFER will be more conservative and hence more credible.

Interestingly, our results suggest that the use of the estimated MF-CFER as a predictor of the cross-section of stock returns yields economically and statistically significant alphas even though our universe consists of only large optionable stocks. This seems to be in contrast to Hou et al. (2018), who document that the alphas found in a number of asset pricing studies, especially those based on friction-related variables, become insignificant once small stocks are weighted less in the universe of test portfolios. We revisit this issue in Section 1.5 where we argue that we can reconcile this seemingly contradicting finding based on Proposition 1.2.3 developed in Section 1.2.5.2.

[Table 1.2 about here.]

1.4.2 Predictive power of the fully-estimated CFER

Next, we compare MF-CFER to the fully-estimated CFER in terms of their ability to predict stock returns cross-sectionally. In the case where results are similar, this would provide further evidence that CFER can be estimated accurately by using MF-CFER and that there is no loss of information from omitting the U term. We compare AVE-CM MF-CFER versus AVE-CM full-CFER. The latter is computed by the embedded leverage (AVE-CM EL-CFER) and margin constraint (AVE-CM MC-CFER) models, separately. We sort stocks in value-weighted portfolios by using any given method to estimate CFER and we report the average returns and alphas for the respective spread portfolios. Table 1.3 reports the results. Panel A (Panel B), shows the average return and $\alpha_{CAPM}, \alpha_{FF3}, \alpha_{FFC}, \alpha_{FF5}, \alpha_{SY}$ of the value-weighted decile spread portfolio where we sort stocks based on the AVE-CM EL-CFER (AVE-CM MC-CFER). t-statistics of the differences between the MF-CFER results and each one of the AVE-CM full-CFER results are reported in square brackets.

We can see that the full-CFER predicts future stock returns cross-sectionally just as it was the case for MF-CFER, which ignores the U term; the average return and alphas are sizable and statistically significant for both the embedded leverage and margin constraints model and they hold for all values of the parameters which appear in the margin constraint model. However, the differences in the average returns and alphas of the portfolios constructed under the MF- and full-CFER are statistically indistinguishable. These results suggest that the effect of leverage and margin constraints has a negligible effect on deviations from put-call parity and hence on the estimation of stocks' CFER. This result extends the evidence on the almost perfect correlations of MF- and full-CFER provided in Section 1.3.3 and confirms that CFER can be estimated accurately by the model-free part of the CFER formula, that is the scaled deviations from put-call parity.

[Table 1.3 about here.]

1.4.3 Alpha of the CFER-adjusted excess returns

Next, we examine the second hypothesis discussed in Section 1.2.5.1. This serves as a complementary test of the predictive ability of CFER documented in the Section 1.4.1. Failing to reject the second hypothesis will suggest that the significantly positive alphas documented in Section 1.4.1 are a result of the (correctly estimated) impact of market frictions rather than a result of omitted factors.

Table 1.4, Panel A, reports the intercepts of the CFER-adjusted regressions, equations (1.32), where we regress the CFER-adjusted excess returns of the CFER-sorted value-weighted decile portfolios and the spread portfolio on a set of factors $\boldsymbol{f}$. We examine five models (CAPM, FF3, FFC, FF5, SY). We can see that the intercept of the regression is statistically insignificant at a 5% significance level whenever the return of a decile portfolio is used as a dependent variable in almost all cases. An exception occurs only in the case where the spread portfolio return is regressed on the market factor (i.e., the CAPM). Moreover, Gibbons et al. (1989) test (untabulated) does not reject the null hypothesis that all eleven alphas are jointly insignificant for any one of the five models examined even at a 10% significance level.

We repeat our analysis by discarding CFER values below 1st percentile point or above 99th percentile point of the CFER distribution across all stocks. Then, we sort stocks by the estimated CFER. This approach removes possible outliers in the estimated CFER; the possible outliers of the estimated CFER may affect the value of the CFER-adjusted returns and hence the estimated alphas. This data filtering is similar to the standard convention in the Fama and MacBeth (1973) regressions (FM, see e.g., Bali et al., 2016). Table 1.4, Panel B, reports the results. All intercepts now become insignificant at a 5% level. In Panel C, we conduct a further robustness test and report the result from a quintile portfolio sort analysis (without the previous truncation of the most extreme CFER values). A quintile portfolio sort is expected to be more robust to outliers because the formed portfolios contain more stocks and thus they are more diversified. We can see that the intercept is not statistically different from zero in all cases.

Our result does not completely eliminate the possibility that the predictability of CFER is driven by omitted factors, because the factor models we examined above are approximations of the true unobservable IMRS. However, if these factor models approximate the true IMRS well, our results suggest that MF-CFER estimates the additional term to the covariance risk premium term fairly correctly and imply that the predictability of CFER originates from capturing the effect of market frictions, rather than from omitted risk factors.

[Table 1.4 about here.]

1.4.4 Model-free CFER versus IVS and DOTS: Evidence

Finally, we compare the cross-sectional predictive ability of AVE-CM MF-CFER to that of IVS and DOTS. The former is expected to perform better (and certainly not worse) than the other two measures for the purposes of predicting stock returns (Propositions 1.2.1 and 1.2.2). We follow Bali and Hovakimian (2009) and calculate IVS by taking the average of the IVS of available pairs of call and put options across different strikes and maturities. We construct DOTS in line with Goncalves-Pinto et al. (2019). Appendix 1.C describes the construction of IVS and DOTS in detail.

Table 1.5 reports the average returns, and alphas for the spread portfolios formed on AVE-CM MF-CFER, IVS and DOTS. For any given sorting criterion, t-statistics for each average return and alpha are reported within parentheses. Within square brackets, we report the t-statistics of the difference between the performance measures of the CFER-sorted and IVS- (DOTS)-sorted portfolios. For this comparison, the null hypothesis is that each pair of sorting criteria yields equal results and the alternative hypothesis is that the CFER-sorted portfolio outperforms. Regarding the IVS-sorted spread portfolio, we can see that the CFER-sorted spread portfolio earns a greater average return by 47 bps and greater alphas by 45–49 bps. Moreover,these differences are statistically significant. When compared with the DOTS-sorted portfolio, again the CFER-sorted spread portfolio earns a greater average return by 19 bps and greater alphas by 27–42 bps, although the economic and statistical differences are generally smaller than the CFER versus IVS case. This result is expected because DOTS is not subject to the vega scaling point encountered in IVS, and the additional term u_t over our sample period is small.[19] In sum, in line with the previous literature, both IVS and DOTS predict stock returns, yet MF-CFER outperforms IVS. This confirms Propositions 1.2.1 and 1.2.2.

[Table 1.5 about here.]

Next, we explore whether the economic mechanism for the predictability of CFER (compensation for market frictions) is distinct from the one employed to explain the predictability of IVS. Cremers and Weinbaum (2010) argue that the predictive power of IVS is due to informed option trading. In the case where informed traders receive optimistic (pessimistic) information on the future stock performance, they prefer to buy (sell) call options and/or sell (buy) put options to utilize their information rather than trading the underlying stock. As a result, new information is incorporated into option prices prior to the underlying stock price causing deviations from put-call parity. However, the informed option trading view and our friction-based view can be complementary because the existence of market frictions is a necessary condition for the informed trading story to hold; if the market is frictionless, deviations from put-call parity and hence a non-zero CFER would not occur.

To investigate further whether the predictive power of MF-CFER reflects informed option trading, we test the following conjecture. If the predictability of MF-CFER is due to informed option trading, then it should be stronger in stocks which underlie the more liquid options; Easley et al. (1998) model suggests that informed traders will trade in options which have higher liquidity. To test our conjecture, we conduct

[19] We calculate u_t as $DOTS_t(K) - CFER^{MF}_{t,t+1}(K)/R^0_{t,t+1}$ and find that the monthly time series of the median of u_t is close to half of the net risk-free rate, which is close to zero over our sample period.

a dependent bivariate sort, where we first sort stocks into five groups based on their respective aggregate option trading volume on the sorting date, and then within each group we sort stocks based on the estimated CFER into quintile portfolios. If the predictability of MF-CFER is due to informed option trading, then we expect that the CFER-spread portfolios of the higher option trading bins should earn more significant returns (i.e., predictive power) compared to those of lower option trading bins.

Table 1.6 reports the results as well as the average volume and liquidity characteristics of the stocks in the individual option trading volume bins. We can see that the greatest average return and alpha are obtained for the smallest option trading volume bin, which is not consistent with the informed option trading view. The average return (alpha) of the MF-CFER spread portfolio in the lowest trading volume bin is 1.32 (1.31) versus 0.91 (1.04) in the biggest trading volume bin; the t-statistics also decrease from about 6 to about 3.2 as we move from the lowest to the highest trading volume bin. We can also see that stocks in smaller option trading volume bins are smaller, have wider bid-ask spread, and lower liquidity, that is, these stocks are subject to larger market frictions. Therefore, our friction-based view is more consistent with our findings; the CFER-spread portfolio of the smallest option trading volume bin earns the highest return because its constituent stocks are subject to larger market frictions. This implies that the degree of market frictions for trading stocks are more pertinent to the return predictability than option trading activity.[20]

[Table 1.6 about here.]

Our findings above are in accordance with Goncalves-Pinto et al. (2019) in two ways. First, they find that the predictive power of IVS and DOTS does not differ between stocks which have non-zero option trading volume and those which have zero option trading volume. This finding and our empirical result above suggest that deviations from put-call parity convey information on future stock returns even when there is small or even no option trading volume. Second, they document that the predictive power of deviations from put-call parity is a manifestation of price

[20]We repeat a similar bivariate sorting exercise using IVS instead of MF-CFER. We obtain the same pattern, that is, the predictability of IVS is stronger in lower option trading bins. This is consistent with our theoretical result that IVS is an approximation (monotonic transformation) of MF-CFER.

pressure in the underlying stock market than of informed option trading. They find that temporary buying (selling) pressure in the stock market tends to mean lower (higher) DOTS. This mechanism can be accommodated in our theoretical framework by viewing the agent in our model as a stock market-maker. For example, when the market-maker faces selling pressure, she provides liquidity by buying the stock. In this case, she demands higher return (positive CFER) to be compensated for transaction costs and other friction-related costs (Proposition 1.2.3).

1.5 Relation between CFER and friction variables

In this Section, we test the predictions about the relation between CFER, its associated spread portfolio alpha, and market frictions we derived in Section 1.2.5.2 (Hypothesis 3). This will allow us to explain why CFER as a measure of the effect of frictions has strong predictive power even among large stocks.

1.5.1 Degree of frictions and variations in the estimated CFER

First, we examine the first prediction in Hypothesis 3, that is, an extreme CFER value implies the presence of large frictions. To this end, we sort stocks in portfolios according to their CFER values and examine the relation between the estimated CFER and various firms' and stocks' characteristics employed as proxies of market frictions. We expect that the portfolios in which CFER takes extreme values are subject to greater frictions.

Table 1.7 reports these characteristics for the CFER-sorted value-weighted decile portfolios, where we use the AVE-CM MF-CFER as a sorting variable. The results confirm Hypothesis 3. We can see an inverse U-shaped relation between the estimated CFER and the SIZE (the logarithm of the market equity) and U-shape relations between CFER and the relative bid-ask spread (BAS), Amihud's (2002) illiquidity measure (Amihud), stock price level, idiosyncratic volatility (IVOL), that is, stocks with extreme estimated CFER values tend to be smaller, have a wider bid-ask spread, lower liquidity and lower stock prices. This is in line with Proposition 1.2.3 because BAS, SIZE and IVOL are natural proxies of transaction costs (e.g., Novy-Marx and

Velikov, 2016). We also see that there is a U-shape relation between CFER and variables which proxy short-sale constraints, that is, the relative short interest (RSI) (see Asquith et al. (2005)), and the estimated shorting fee (ESF) of Boehme et al. (2006). This may seem to be at odds with the literature, which documents that tight short-sale constraints are related to future underperformance (i.e., negative CFER); this implies that CFER should be monotonically (negatively) related to the size of short-sale frictions. However, our result may be driven by the correlation between RSI/ESF and other characteristics, which univariate sorts cannot take into account. Subsequently, we shed light on this relation by running FM regressions, where we control for other characteristic variables such as firm size and liquidity. Finally, we also find a U-shape relation between the market beta and book-to-market ratio (B/M), and the estimated CFER.

The documented U-shape relation between CFER and proxies of transaction costs ρ such as BAS and IVOL, confirms the prediction of Hypothesis 3-1 and reconciles the strong ability of CFER to predict stock returns even for optionable stocks with Hou et al. (2018) results, who document that most of the previously documented friction-related anomalies become insignificant when the effect from microcap stocks is mitigated. This is because the U-shape relation implies that when one sorts stocks based on a proxy of ρ, the highest ρ portfolio contains both outperforming and underperforming stocks since the larger the market friction, the more extreme the CFER values are, just as Proposition 1.2.3 predicts. As a result, the average portfolio abnormal performance may show not to exist. On the other hand, CFER is a *signed* measure of the effect of frictions on the expected return that specifies which stock will outperform and underperform.

[Table 1.7 about here.]

Next, we examine Hypothesis 3-2 by conducting dependent bivariate sorts. We first sort stocks in portfolios by their respective proxy for transaction costs. Then, within any given portfolio, we sort stocks in portfolios based on their respective CFER and calculate the spread portfolios' returns. We expect that the average return of the spread portfolios will increase as a function of the transaction costs proxy. This is because the variation of CFER will be greater within a group of stocks that is subject to larger transaction costs. Table 1.8, Panel A, reports the bivariate dependent sort, first by BAS, then by CFER. The result verifies our conjecture. The average CFER of the CFER-sorted spread portfolio increases with the level of the bid-ask spread. The average return and α_{FFC} of the CFER-sorted spread portfolios also increase with the level of the bid-ask spread. We find a similar pattern in the CFER-sorted portfolios in Table 1.8, Panel B, where we use the SIZE as an alternative sorting proxy for transaction costs. In general, the average CFER, average return, and α_{FFC} of the CFER-sorted spread portfolios decrease in the level of SIZE.[21]

[Table 1.8 about here.]

1.5.2 CFER and frictions: Fama-MacBeth regressions

In addition to the portfolio sort approach, we also examine the relation between CFER and friction-related variables by conducting FM regressions. This constitutes a reverse engineering approach to identify the dominant market frictions for optionable stocks. We run univariate as well as multivariate regressions of CFER on SIZE, BAS, IVOL, Amihud and RSI which are popular proxies of market frictions. For each month t ($t = 1, 2, ..., T$), we estimate the following cross-sectional regression across the i individual stocks

$$CFER_{i,t,t+1} = \alpha + \beta' X_{i,t}, \tag{1.37}$$

where we use AVE-CM MF-CFER as the left-hand side variable and $X_{i,t}$ is a vector that contains the characteristics variables of individual stocks. Then, we calculate the time-series average and the t-statistics of the estimated T cross-sectional intercept α and the β coefficients. To ensure that our estimates are not driven by extreme values, we truncate AVE-CM MF-CFER and variables in $X_{i,t}$ at a 1% threshold level.

Given that the previous analysis has documented a non-linear relation between CFER and firm characteristics, we conduct the regressions twice by splitting our CFER sample to positive and negative values. For the positive CFER subsample, our third hypothesis suggests that more extreme (i.e., higher) CFER corresponds to stocks which are subject to larger transaction costs and frictions. Therefore, we expect a negative sign for the coefficient of SIZE, and a positive sign for the coefficient

[21] We obtain similar results when we use IVOL or Amihud as a proxy for transaction costs.

of BAS, Amihud and IVOL (smaller firms, wider bid-ask spread, lower liquidity and higher IVOL correspond to more extreme CFER). For the negative CFER subsample, we expect the opposite signs for the coefficients of these four variables. Regarding the coefficient of RSI, we expect the negative coefficient for the negative CFER subsamples; a negative CFER is consistent with the agent short-selling the stock because the agent demands the short-sale to be more "profitable." In this case, a higher RSI implies severer short-sale constraints and hence CFER should become more negative. On the other hand, in the case of stocks with positive CFER, Proposition 1.2.3 makes no prediction for the sign of the variables which capture short-sale constraints; this is because short-sale constraints do not matter when the agent buys the stock.

Table 1.9 reports the results. We can see that the coefficients of the SIZE, BAS, IVOL and Amihud variables have the expected signs both in the univariate and multivariate regressions. Moreover, they are statistically significant; the only exception is Amihud for the positive CFER subsample. These findings corroborate that CFER is related to various liquidity- and transaction costs-related variables in the theoretically predicted manner. Regarding RSI, we obtain the expected negative sign in the negative subsample. This shows that, RSI and CFER are negatively related when the agent sells stocks as theory predicts. The result for the positive subsample is mixed. Note that the adjusted R^2 of the univariate regression of CFER on RSI is lower than those of the other friction-related variables, especially for the positive subsample. This may suggest that short-sale constraints are of second-order importance to explain CFER variations compared to the other friction-related variables for our universe of optionable stocks. This is expected because optionable stocks correspond to big stocks for which short-sale constraints are not pronounced (see e.g., D'Avolio, 2002; Drechsler and Drechsler, 2014). Most importantly, the documented relation between CFER and proxies of *various* market frictions suggests that CFER is a "sufficient statistic" which subsumes the overall effect of any relevant market frictions on expected stock returns. The FM regression results also confirm that the variations in CFER is strongly related to transaction costs as suggested by Proposition 1.2.3.

[Table 1.9 about here.]

1.5.3 Size of alphas and market frictions

The magnitude of alpha of the CFER-sorted spread portfolios reported in Section 1.4.1, is in accordance with the upper bound of the alpha predicted by Proposition 1.2.3; the estimated alpha is close yet below the upper bound implied by equation (1.35) (Hypothesis 3-4) after substituting empirical estimates of the variables into equation (1.35). The upper bound is at least as large as twice the round-trip transaction costs (4ρ) and it is approximately 2.0% given the empirically estimated round-trip transaction costs for large stocks (Lesmond et al., 1999; Hasbrouck, 2009). This numerical value is largely in line with the convention in the literature as well. For example, the stylized transaction costs assumption in Brandt et al. (2009) and DeMiguel et al. (2019) translates to approximately $4\rho = 1.8\%$ in our universe of stocks over our sample period. Moreover, the upper bound may be even higher in the presence of other frictions (e.g., short-sale constraints and margin constraints). In sum, the limits of arbitrage for trading (even big optionable) stocks are large enough to generate a 2% alpha per month. Therefore, the alphas reported in Table 1.2 are not excessive given the degree of market frictions investors face.

1.6 Predictive power of CFER: Robustness tests

In this Section, we report a number of further robustness tests on the cross-sectional predictability of CFER. We use portfolio sorts to examine whether our baseline results may differ across the four possible ways of constructing CFER. We also investigate whether results are driven by outliers, stock reversals, non-synchronous trading in the option and underlying market. We also test whether results are stable over time by performing a sub-sample analysis. Finally, we conduct FM regression tests to confirm that CFER is positively related to stock returns.

1.6.1 Robustness tests based on portfolio sorts

First, we examine whether the average return and α_{FFC} of the decile spread portfolio differ across the four CFER proxies. Table 1.10, Panel A, reports the results. We can see that the average return and α_{FFC} are sizable and statistically significant in both the value-weighted and equally-weighted portfolios for any of the four ways of computing CFER. In addition, we can see that CFER computed by the AVE-CM method delivers the highest average return and alpha. This may be due to the fact that AVE-CM CFER reduces any measurement errors in CFER at each strike by averaging them and hence the signal to sort stocks in portfolios has greater predictive power. It may also be the case that the 30-day constant maturity CFER gives cleaner signals for future outperformance or underperformance in the succeeding month than the CFER extracted from options with maturity closest to 30-day does.

[Table 1.10 about here.]

Second, regarding the effect of extreme CFER values, we check whether the predictive power of CFER is driven by few stocks that have extreme CFER value. We perform two alternative robustness tests based on two respective ways of forming portfolios. First, we remove stocks whose CFER is below the 1st percentile point or above the 99th percentile point. Second, we form quintile rather than decile portfolios; each quintile portfolio has twice as many stocks compared to decile portfolios, portfolio returns are more robust to the effect of outliers. Table 1.10, Panel B, reports the average returns and alphas of the long-short portfolio. The first two columns report the average return and the risk-adjusted return of the decile spread portfolio, where we remove stocks whose CFER is below 1st percentile point or above 99th percentile point. We form two spread portfolios as the difference of Portfolio 10 minus Portfolio 1, and Portfolio 9 minus Portfolio 2, respectively. By construction, the latter spread portfolio contains stocks which have less extreme CFER values. We can see that albeit the average return and alpha decrease compared to the full sample results, results are still economically and statistically significant. The third and fourth column of Table 1.10 show the analogous results for the CFER-sorted quintile portfolios. Again, the average and risk-adjusted returns are economically and statistically significant. The results suggest that the predictive power of CFER remains even after we discard 40% of our initial sample with the most extreme CFER observations.

Third, the predictability of CFER may be a manifestation of the short-term reversal effect of Jegadeesh and Titman (1993), which is typically attributed to mispricing due to microstructural frictions (see Chapter 12 of Bali et al., 2016). To examine this conjecture, we conduct a 5×5 dependent bivariate sort, where we first sort stocks according to the previous month return $R_{t-1,t}$, and then sort by the AVE-CM CFER. Hence, we can test whether CFER has predictive power after the previous month return is controlled. We report results in Table 1.11. The first five columns of Table 1.11 report the average returns of the 25 bivariate-sorted portfolios. The sixth to last columns report the average returns, α_{FFC}, and the average CFER of the long-short portfolios of CFER, respectively, after controlling the previous month returns. Overall, the spread portfolios' risk-adjusted returns (seventh column) are still statistically and economically significant after controlling the previous month return. This suggests that the predictive power of CFER is not subsumed by the short-term reversal phenomenon.

[Table 1.11 about here.]

Fourth, we examine whether our results on the documented predictive ability of CFER are of use to real time investors in the presence of non-synchronous trading in the option and the underlying stock market (Battalio and Schultz, 2006). The CBOE option market closes after the underlying stock market. Consequently, in real time, the CFER value computed from option closing prices may not be available to investors on the close of the stock market. As a result, in real time it may be the case that investors cannot exploit the CFER signal since the stock market has closed and hence they cannot trade stocks.[22] In this case, inevitably, investors will trade stocks at the open of the next day. To examine whether the calculated at the end-of-day CFER may be of use to an investor, we calculate post-ranking returns using the open-to-close monthly stock return, where the open stock price is that of the first trading day after the day on which CFER is estimated.

Table 1.12, Panel A, reports the portfolio analysis results, where the open-to-close return is used. The average return of the spread portfolio is 1.60% per month and it is almost the same as the average return obtained from the baseline analysis using close-to-close returns, 1.64%. The FFC alpha is 1.83%, which is again almost the

[22]The underlying market closes at 4:00 p.m. (EST). Prior to June 23, 1997, the closing time for CBOE options on individual stocks was 4:10 p.m. On June 23, 1997, CBOE changed the closing time to 4:02 p.m. (i.e., only two minutes after the closing of the underlying stock market). Since March 5, 2008, OM reports option prices at 15:59 p.m. These changes minimize the potential non-synchronicity bias during our sample period. Nevertheless, in the absence of intra-day option prices, it is not known whether the CFER estimates were available in real time before the stock market close prior to March 5, 2008.

same as the corresponding alpha in the close-to-close return case, 1.86%. This result implies that the predictive power of the estimated CFER prevails even in the presence of non-synchronous trading in the stock and option market; the predictive power of CFER does not change overnight.

[Table 1.12 about here.]

Fifth, we examine whether results are robust in the case where we exclude stocks with low prices. Table 1.12, Panel B, reports the results, where we exclude stocks whose price level is lower than $10. This filtering criteria removes about 10% of stocks compared to the baseline analysis. We can see that the average return and the alphas of the spread portfolio decrease when we remove the low priced stocks. However, the returns are still highly statistically and economically significant. This is in contrast with the literature on the predictability of friction-related variables, where the predictability mainly stems from small, low priced stocks which are more susceptible to market frictions.

Sixth, we examine whether the predictive cross-sectional power of CFER prevails in the case where we use different breakpoints to form the decile portfolios. Table 1.12, Panel C, reports the portfolio sort result, where we form decile portfolios based on the NYSE breakpoints. Hou et al. (2018) recommend using only NYSE stocks to compute breakpoints rather than using all stocks. This is because the latter method allows smaller and more volatile NASDAQ stocks to have a greater relative importance in the extreme decile portfolios and amplifies asset pricing anomalies. We can see that the predictive ability of CFER is robust irrespective to the breakpoint method. The average return and alpha of the spread portfolio are still significant, albeit smaller compared to these obtained in the baseline analysis.

A remark is in order at this point regarding the validity of our findings in the light of the growing concerns on data snooping among the asset pricing literature (e.g., Harvey et al., 2016; Hou et al., 2018; Harvey, 2017). Our results are reassuring because the CFER-sorted decile spread portfolio earns significant alphas even when we follow the construction method recommended by Hou et al. (2018) (i.e., value-weighted and NYSE breakpoints) and the t-statistics are above five, which is above the thresholds based on the Bayesianized t-statistics proposed by Harvey (2017). Furthermore, these recent studies emphasize the importance of relying on a sound theoretical model to explain why a certain variable should predict asset returns. Our approach satisfies this criterion since the predictive power of CFER is justified based on a formal asset pricing model. Moreover, these studies also emphasize that the design and practical specifications of empirical studies should not allow ad-hoc flexibility as possible. The computation of CFER allows little flexibility because it does not require any parameter estimations nor historical data. The only flexibility is in the choice of strike prices and maturities, yet we have shown that the predictive power of CFER is robust to the CFER construction methods (Table 1.10, Panel A).

Seventh, we examine whether the predictive power of CFER still exists over alternative sub-periods. We divide our initial sample period into January 1996 to December 2006 and January 2007 to April 2016. We choose December 2006 as a splitting point for the following two reasons: first, 2007 is the onset of the financial crisis and hence market frictions have increased in the period thereafter. This may have an effect on the cross-section of CFER values as Figure 1.1b has indicated. Second, 2007 coincides with the period where the academic research, which demonstrates that the option-implied measures extracted from individual equity options predict the cross-section of future stock returns, has appeared.[23] McLean and Pontiff (2016) find that the publication of academic research on asset pricing anomalies eliminates the predictability of variables which manifest asset pricing anomalies. Panels A and B of Table 1.13 report the results. The spread portfolio's average return and α_{FFC} decrease by 68 bps and 119 bps, respectively, from the earlier to the more recent sub-sample. However, the average return and alpha of the spread portfolio are still statistically and economically significant. These results show that the predictive power of CFER is not solely driven by financial crisis period.

[Table 1.13 about here.]

1.6.2 Robustness tests: Fama-MacBeth regressions

We complement the robustness tests which use portfolio sorts with FM regressions, where we regress stock returns on stocks' characteristics including the estimated

[23]For instance, Cremers and Weinbaum (2010) and Bali and Hovakimian (2009) working paper versions appeared on the SSRN website in March 2007 and November 2007, respectively.

CFER. These regressions provide additional robustness checks for our results since they employ all firms without imposing portfolio breakpoints and allow for control variables (see Hou et al., 2016). Similar to the estimation in Section 1.5, for each month, we estimate cross-sectional regressions of stock returns on characteristics variables of individual stocks. Then, we calculate the time-series average and the t-statistics of the estimated T cross-sectional intercept and the coefficients on characteristics variables. To ensure that our results are not driven by extreme values, we truncate left-hand side variables at a 1% threshold level.

Table 1.14 reports the results. Model (1) shows that the estimated CFER is positively related to the stock returns. In Model (2), we employ various control variables including market beta, SIZE, log of book-to-market ratio, momentum ($R_{t-12,t-1}$). We also include the previous month return $R_{t-1,t}$, IVOL, asset growth rate and profitability since it is well-known that these variables have predictive power for future stock returns (see e.g., Jegadeesh and Titman (1993) for the short-term reversal, Ang et al. (2006) for IVOL, and Hou et al. (2015) for asset growth and profitability). The coefficient of the estimated CFER is still positive and statistically significant even after controlling for these variables.[24] In Model (3), we further add three liquidity related variables, Amihud's (2002) illiquidity measure, the relative bid-ask spread and the turnover rate. The estimated coefficient of CFER is virtually unchanged from Model (2). In Model (4), we further add IVS. The size of the CFER coefficient and its t-statistic decrease, yet it remains significant at a 1% significance level. On the other hand, the IVS coefficient is not statistically significant. This corroborates our findings that IVS is an approximation (monotonic transformation) of MF-CFER and MF-CFER is a superior predictor of stock returns.

In columns (5) to (10), we report results from conducting FM regressions on two separate stock universes. First, as we have discussed above (Table 1.12, Panel C), NASDAQ stocks are smaller and more volatile than NYSE and Amex stocks. Hence,

[24] In our FM regression results, traditional return predictors such as log book-to-market ratio and momentum are insignificant. To explore this further, we conduct FM regressions by excluding CFER from the set of control variables and using all common stocks including non-optionable stocks. In the case where our stock dataset commences in 1972, these traditional variables have significant coefficients, whereas if we use the data starting from 1996, they become insignificant. This suggests that well-known effects such as value and momentum effects are weaker in the recent period covered by the OM database. Therefore, the insignificant coefficients we obtain for some traditional variables should not be attributed to the narrower universe of optionable stocks neither to the inclusion of CFER in the regressions.

the FM regression results may be driven by the NASDAQ stocks (see Hou et al., 2016). To examine this possibility, we repeat the FM regression by splitting our sample into NYSE/Amex stocks and NASDAQ stocks. Columns (5) and (6) report respective results. The coefficients of CFER are still highly significant regardless of whether we use only NYSE/Amex stocks or NASDAQ stocks. Next, as we have seen in Table 1.7, CFER and various firm and stock characteristics exhibit (inverse) U-shaped relations. Therefore, it might be the case that this non-linear structure affects the FM regression results. To address this issue, we split our initial sample based on the sign of CFER; we split our sample into two parts where the splitting points is a zero CFER value. Hence, the two parts roughly correspond to the left and right part of the U-shaped relations so that each subsample has a monotonic relation between CFER and the firm and stock characteristics. This would be closer to the structure of the FM regressions. Columns (7) and (8) demonstrate that the coefficient of CFER is larger for negative CFER samples, but the coefficient of CFER is significant for both subsamples. Finally, we split our sample into January 1996 to December 2006 and January 2007 to April 2016 as before and we re-apply the FM regressions. We can see from the last two columns that the estimated coefficient on CFER becomes slightly smaller in the latter period, but they are statistically significant in both sub-periods.

[Table 1.14 about here.]

1.7 Conclusions and implications

We derive a novel formula to estimate the *contribution of frictions to expected returns* (CFER) within a formal asset pricing setting. CFER is the wedge between the expected excess return and the covariance risk premium, and captures the effect of market frictions to expected stock returns. We show that properly scaled deviations from put-call parity reliably estimate the underlying stock's CFER as follows. First, we show that deviations from put-call parity contain information about the effect of frictions on both the underlying and the synthetic stock. Then, we empirically estimate the effect of frictions on the synthetic stock and document that it is negligible compared to the observed cross-sectional deviations from put-call parity. Our CFER formula provides a forward-looking measure of the effect of frictions to stock prices and it requires no estimation of parameters.

We estimate CFER for each optionable U.S. common stock. Our empirical findings contribute to two strands of the literature. First, we contribute to the literature on the informational content of option prices. We find that CFER has a strong predictive power, which is robust to the recent data snooping concerns and does not stem from omitted risk factors. Moreover, we show that CFER has superior properties than other measures of deviations from put-call parity and outperforms the implied volatility spread as a predictor of stock returns. We document that the size of market frictions is a more pertinent explanation to the return predictability than option trading activity.

Second, we contribute to the literature on the relation between market frictions and stock returns. The large magnitude and variation of the estimated CFER and its associated strong predictive power suggest that market frictions have a non-negligible effect on even large optionable stocks. We reconcile this finding with Hou et al. (2018), who document that almost all previously documented friction-related anomalies vanish once micro-cap stocks are not included in the analysis. This is possible thanks to the ability of CFER to identify outperforming and underperforming stocks due to its theoretical foundation as a risk-adjusted signed measure of expected return (alpha) which captures the effect of market frictions; on the contrary, common measures of market frictions do not possess this property. Moreover, we theoretically show that the upper bound of the alpha of CFER-sorted spread portfolios is at least twice the round-trip transaction costs, thus accommodating the magnitude of alpha of the CFER sorted spread portfolio.

1.A Proofs

1.A.1 Proof of Theorem 1.2.1

First, note that dividing both sides of equation (1.9) by S_t yields

$$1 = \mathbb{E}_t^{\mathbb{P}}[m^*_{t,t+1} R_{t,t+1}] + \frac{1}{S_t}\left(M_t^S - \frac{\partial TC_t}{\partial \theta_t^S} + \frac{\beta}{u'(c_t)} \mathbb{E}_t^{\mathbb{P}}\left[\frac{\partial V_{t+1}}{\partial \theta_t^S} \right] \right). \tag{1.A.1}$$

Application of the property of covariances, $Cov_t(X,Y) = \mathbb{E}_t[XY] - \mathbb{E}_t[X]\mathbb{E}_t[Y]$, to equation (1.A.1) yields

$$\mathbb{E}_t^{\mathbb{P}}[m_{t,t+1}^* R_{t,t+1}] = Cov_t^{\mathbb{P}}(m_{t,t+1}^*, R_{t,t+1}) + \frac{\mathbb{E}_t^{\mathbb{P}}[R_{t,t+1}]}{R_{t,t+1}^0} \tag{1.A.2}$$

because $\mathbb{E}_t^{\mathbb{P}}[m_{t,t+1}^*] = 1/R_{t,t+1}^0$. Substituting (1.A.2) to (1.A.1) and rearranging terms yields equation (1.11), where $CFER_{t,t+1}$ is defined in equation (1.12). □

1.A.2 Proof of Proposition 1.2.1

Deviations from put-call parity $\widetilde{S}_t(K) - S_t$ can be rewritten as the difference between the observed call price $C_t(K)$ and the hypothetical call price $\widetilde{C}_t(K)$, which is calculated as if put-call parity would hold, that is,

$$\widetilde{C}_t(K) = P_t(K) + S_t - \frac{K + D_{t+1}}{R_{t,t+1}^0}. \tag{1.A.3}$$

Let $BS_{call}(IV)$ ($BS_{put}(IV)$) be the Black-Scholes call (put) option function viewed as a function of the implied volatility parameter. Then, by the definition of the BS-IV, $IV_t^c(K)$ and $IV_t^p(K)$ satisfy $C_t(K) = BS_{call}(IV_t^c(K))$ and $P_t(K) = BS_{put}(IV_t^p(K))$, respectively. Moreover, it follows that

$$\widetilde{C}_t(K) = BS_{put}(IV_t^p(K)) + S_t - \frac{K + D_{t+1}}{R_{t,t+1}^0} = BS_{call}(IV_t^p(K)), \tag{1.A.4}$$

because the pair of the Black-Scholes European call and put option prices with the same volatility satisfies the put-call parity. This shows that $C_t(K) - \widetilde{C}_t(K) = BS_{call}(IV_t^c(K)) - BS_{call}(IV_t^p(K))$. Therefore, a first-order Taylor series approximation of $BS_{call}(IV^c(K)) - BS_{call}(IV^c(K))$ around the mid volatility point $(IV_t^c(K) + IV_t^p(K))/2$ yields

$$C_t(K) - \widetilde{C}_t(K) = BS_{call}(IV_t^c(K)) - BS_{call}(IV_t^p(K)) \approx \mathcal{V}_t(K)(IV_t^c(K) - IV_t^p(K)), \tag{1.A.5}$$

where $\mathcal{V}_t(K)$ is the Black-Scholes *vega*, $\partial BS_{call}(\sigma)/\partial\sigma$, evaluated at the mid IV point $(IV_t^c(K) + IV_t^p(K))/2$. By substituting this approximation into equation (1.15), we obtain equation (1.28). This derivation shows that the approximation error in (1.28) originates from the higher-order terms of the Taylor series approximation of

$BS_{call}(IV_t^c(K)) - BS_{call}(IV_t^p(K)).$ □

1.A.3 Proof of Proposition 1.2.2

Substituting the definition of S_t^U and S_t^L into equation (1.30) yields

$$DOTS_t = \frac{1}{S_t}\left(C_t^{mid}(K) - P_t^{mid}(K) - S_t + \frac{1}{2}\left(1 + \frac{1}{R^0_{t,t+1}}\right)K + \frac{D_{t+1}}{2R^0_{t,t+1}}\right), \quad (1.A.6)$$

where C_t^{mid} and P_t^{mid} are the mid price of American options. By the definition of η_t^c and η_t^p, $C_t := C_t^{mid} - \eta_t^c$ and $P_t := P_t^{mid} - \eta_t^p$ are the European option prices. Then, by substituting the definition of the synthetic stock price (equation (1.13)), we obtain

$$DOTS_t = \frac{\widetilde{S}_t(K) - S_t}{S_t} + \frac{1}{S_t}\left[\left(\eta_t^c - \frac{D_{t+1}}{2R^0_{t,t+1}}\right) - \left(\eta_t^p - \frac{1}{2}\left(1 - \frac{1}{R^0_{t,t+1}}\right)\right)\right]. \quad (1.A.7)$$

The first term in the right hand side of equation (1.A.7) is $CFER^{MF}_{t,t+1}(K)/R^0_{t,t+1}$. This completes the proof of equation (1.29). □

1.A.4 Proof of Proposition 1.2.3

First, to prove the upper bound of equation (1.34), consider the case where the agent buys (i.e., $\Delta\theta_t^S > 0$) and longs (i.e., $\theta_t^S > 0$) the stock at time t. Once dividing by S_t, the first-order condition for the stock given the margin constraint function (equation (1.7)) and the transaction cost function (equation (1.5)) yields

$$1 = \mathbb{E}_t^{\mathbb{P}}[m^*_{t,t+1}R_{t,t+1}] - \frac{\lambda_t^{MC}}{u'(c_t)}\mu_t^S - \rho + \frac{\beta}{u'(c_t)}\mathbb{E}_t^{\mathbb{P}}\left[\frac{1}{S_t}\frac{\partial V_{t+1}}{\partial \theta_t^S}\right]. \quad (1.A.8)$$

The last term in equation (1.A.8) equals

$$\frac{\beta}{u'(c_t)}\mathbb{E}_t^{\mathbb{P}}\left[\frac{1}{S_t}\frac{\partial V_{t+1}}{\partial \theta_t^S}\right] = \rho\mathbb{E}_t^{\mathbb{P}}\left[\beta\frac{u'(c_{t+1})}{u'(c_t)}\frac{S_{t+1}}{S_t}sgn(\Delta\theta^S_{t+1})\right]. \quad (1.A.9)$$

due to the envelop theorem $\partial V_{t+1}/\partial\theta_t^S = u'(c_{t+1})\rho S_{t+1} sgn(\Delta\theta^S_{t+1})$.

We can further approximate equation (1.A.9) as

$$\begin{aligned}\rho\mathbb{E}_t^{\mathbb{P}}\left[\beta\frac{u'(c_{t+1})}{u'(c_t)}\frac{S_{t+1}}{S_t}sgn(\Delta\theta^S_{t+1})\right] &\approx \rho\mathbb{E}_t^{\mathbb{P}}\left[\beta\frac{u'(c_{t+1})}{u'(c_t)}R_{t,t+1}sgn(\Delta\theta^S_{t+1})\right] \\ &\approx \rho\mathbb{E}_t^{\mathbb{P}}\left[m^*_{t,t+1}R_{t,t+1}sgn(\Delta\theta^S_{t+1})\right].\end{aligned} \quad (1.A.10)$$

At the first approximation, we replace the ex-dividend return S_{t+1}/S_t with the cum-dividend return $R_{t,t+1} = (S_{t+1} + D_{t+1})/S_t$. Its approximation error is negligible because the product of the dividend-to-price ratio D_{t+1}/S_t (the difference between the ex- and cum-dividend return) and the proportional transaction costs ρ is small; given typical values of these variables, the product is less than one basis point. At the second approximation, we employ an approximation $u'(c_{t+1}) \approx \partial V_{t+1}/\partial W_{t+1}$. This relation holds exactly when the margin constraints are not binding. Even when the margin constraints are binding, one can show that the approximation error is negligible given the empirical estimate of the shadow price of the margin constraint.[25]

Substituting equation (1.A.10) into (1.A.8) yields

$$1 + \rho \approx \mathbb{E}_t^{\mathbb{P}} \left[m^*_{t,t+1} R_{t,t+1} (1 + \rho \cdot sgn(\Delta\theta^S_{t+1})) \right] - \frac{\lambda_t^{MC}}{u'(c_t)} \mu_t^S. \tag{1.A.11}$$

Since $m^*_{t,t+1} R_{t,t+1}$ is non-negative and $1 - \rho \leq 1 + \rho \cdot sgn(\Delta\theta^S_{t+1})$, we obtain

$$1 + \rho \gtrsim (1 - \rho) \mathbb{E}_t^{\mathbb{P}} [m^*_{t,t+1} R_{t,t+1}] - \frac{\lambda_t^{MC}}{u'(c_t)} \mu_t^S. \tag{1.A.12}$$

Dividing both sides of equation (1.A.12) by $1 - \rho$ and using $\mathbb{E}_t^{\mathbb{P}}[m^*_{t,t+1} R_{t,t+1}] = 1 + R^0_{t,t+1} CFER_{t,t+1}$, which is implied by equation (1.9), yield

$$CFER_{t,t+1} \lesssim \frac{2\rho}{(1-\rho) R^0_{t,t+1}} + \frac{\lambda_t^{MC} \mu_t^S}{(1-\rho) R^0_{t,t+1} u'(c_t)}. \tag{1.A.13}$$

Since $(1 - \rho) R^0_{t,t+1}$ in the denominators is close to one, further approximation proves the right inequality in equation (1.34).

To prove the lower bound case of equation (1.34), consider a situation where the agent short-sells the stock. In this case, the first-order condition for the stock yields

$$1 - (\rho + \varsigma) = \mathbb{E}_t^{\mathbb{P}} [m^*_{t,t+1} R_{t,t+1}] + \frac{\lambda_t^{MC}}{u'(c_t)} \mu_t^S + \frac{\beta}{u'(c_t)} \mathbb{E}_t^{\mathbb{P}} \left[\frac{1}{S_t} \frac{\partial V_{t+1}}{\partial \theta_t^S} \right]. \tag{1.A.14}$$

We follow the same approximation argument for the last term in equation (1.A.14) to obtain the following equation in analogy to equation (1.A.11):

$$1 - (\rho + \varsigma) \approx \mathbb{E}_t^{\mathbb{P}} \left[m^*_{t,t+1} R_{t,t+1} (1 + \rho \cdot sgn(\Delta\theta^S_{t+1})) \right] + \frac{\lambda_t^{MC}}{u'(c_t)} \mu_t^S. \tag{1.A.15}$$

[25]The envelop theorem yields $\partial V_{t+1}/\partial W_{t+1} = u'(c_{t+1}) + \lambda^{MC}_{t+1} = u'(c_{t+1})(1 + \lambda^{MC}_{t+1}/u'(c_{t+1}))$. The "approximation coefficient," $1 + \lambda^{MC}_{t+1}/u'(c_{t+1})$, is close to one given the estimate of the shadow price of the margin constraints, which is at most about 0.8% per month (see the discussion in Section 1.3.3).

This time, we use the inequality $1 + \rho \cdot sgn(\Delta\theta^S_{t+1}) \leq 1 + \rho$ and obtain

$$1 - (\rho + \varsigma) \lesssim (1 + \rho)\mathbb{E}^{\mathbb{P}}_t[m^*_{t,t+1}R_{t,t+1}] + \frac{\lambda^{MC}_t}{u'(c_t)}\mu^S_t. \tag{1.A.16}$$

Then, some more algebra proves the left inequality in equation (1.34).

Note that the approximate equality holds at (1.A.12) for the upper bound ((1.A.16) for the lower bound) when $sgn(\Delta\theta^S_{t+1}) = -1$ ($sgn(\Delta\theta^S_{t+1}) = 1$) almost surely. This means that CFER attains the maximum (minimum) possible value when the agent expects to sell (buy) the stock in the next period for sure. Such a situation may arise when the stock market-maker is forced to increase (decrease) her stock inventory drastically due to large selling (buying) pressure at time t and hence she wishes to decrease (increase) the stock inventory level in the next period for sure. In sum, the extreme CFER value occurs when the agent expects to unwind her time t trading in the next period for sure. In this case, she demands CFER to deviate at least by the round-trip transaction costs for compensating the contemporaneous and expected future transaction costs. Our discussion here is a generalization of He and Modest's (1995) result, where they implicitly assume that the agent always unwinds her position taken at time t in the next period.

Finally, the latter statement of Proposition (equation (1.35)) follows from the inequality range, equation (1.34). □

1.B CFER estimation: Effect of frictions on options

Theorem 1.2.2 shows that CFER equals the sum of the model-free $CFER^{MF}_{t,t+1}(K)$ term and the effect of frictions on the synthetic stock $U_{t,t+1}(K) + T_{t,t+1}(K)$, where $U_{t,t+1}(K)$ represents the effect of constraints on allocations and $T_{t,t+1}(K)$ represents the effect of transaction costs for trading options. In this Section, we provide necessary material and discussion which complement our discussion in Sections 1.2.3 and 1.3.3. Specifically, in Appendix 1.B.1, we explain how we assess the unobservable $U_{t,t+1}(K)$, which is proportional to $M^c_t(K) - M^p_t(K)$ (equation (1.16)). In Appendix 1.B.2, we explain how we treat $T_{t,t+1}(K)$ in our estimation framework.

1.B.1 Estimation of $U_{t,t+1}$: Effect of constraints on options

1.B.1.1 Embedded leverage model

Frazzini and Pedersen (2012) document that option returns are lower (compared to the covariance risk premium term) by 1.25% per month per unit of embedded leverage. In line with their empirical findings, we model the CFER term arising due to the embedded leverage effect as $k\Omega_t(K)$, where $\Omega_t(K)$ is option's embedded leverage and k is the sensitivity of option returns to the embedded leverage. Call (put) option's embedded leverage is defined as $\Omega_t^c(K) = |\Delta_t^c(K)S_t/C_t(K)|$ ($\Omega_t^p(K) = |\Delta_t^p(K)S_t/P_t(K)|$), where $\Delta_t^c(K)$ ($\Delta_t^p(K)$) is the call (put) option's delta. Then, option returns are expressed as

$$\begin{aligned} \mathbb{E}_t^{\mathbb{P}}[R_{t,t+1}^c(K)] - R_{t,t+1}^0 &= -R_{t,t+1}^0 Cov_t^{\mathbb{P}}(m_{t,t+1}^*, R_{t,t+1}^c(K)) + k\Omega_t^c(K), \\ \mathbb{E}_t^{\mathbb{P}}[R_{t,t+1}^p(K)] - R_{t,t+1}^0 &= -R_{t,t+1}^0 Cov_t^{\mathbb{P}}(m_{t,t+1}^*, R_{t,t+1}^p(K)) + k\Omega_t^p(K), \end{aligned} \tag{1.B.1}$$

where $R_t^c(K) = (S_{t+1} - K)^+/C_t(K)$ and $R_t^p(K) = (K - S_{t+1})^+/P_t(K)$ are the return of call and put options, respectively.[26] Given that two equations in (1.B.1) are equivalent to $M_t^c(K) = kC_t\Omega_t^c(K)/R_{t,t+1}^0$ and $M_t^p(K) = kP_t\Omega_t^p(K)/R_{t,t+1}^0$, $U_{t,t+1}(K)$ satisfies the following equation under this embedded leverage model (the superscript EL stands for "embedded leverage"):

$$U_{t,t+1}^{EL}(K) = -\frac{R_{t,t+1}^0}{S_t}\left(\frac{kC_t(K)\Omega_t^c(K)}{R_{t,t+1}^0} - \frac{kP_t(K)\Omega_t^p(K)}{R_{t,t+1}^0}\right) = k(|\Delta_p(K)| - \Delta_c(K)). \tag{1.B.2}$$

1.B.1.2 Margin constraints model

First, we prove equations (1.26) and (1.27). Given equation (1.22), it suffices to show

$$-\frac{1}{S_t}[\mu_t^c(K)C_t(K)sgn(\theta_t^c(K)) - \mu_t^p(K)P_t(K)sgn(\theta_t^p(K))] = E_t(K), \tag{1.B.3}$$

where the sign function $sgn(x)$ returns 1 (-1) if x is positive (negative). We can further calculate the left-hand side of equation (1.B.3) for each one of four possible combinations of the signs of $\theta_t^c(K)$ and $\theta_t^p(K)$ as follows.

When $\theta_t^c(K) > 0$ and $\theta_t^p(K) > 0$, the margin rule is $\mu_t^c(K) = \mu_t^p(K) = 1$ and

[26] For simplicity, we abstract from option transaction costs based on our discussion in Section 1.2.5.2.

the left-hand side of equation (1.B.3) boils down to $(C_t - P_t)/S_t$. When $\theta_t^c(K) < 0$ *and* $\theta_t^p(K) < 0$, the margin rule are given by $\mu_t^c(K)C_t(K) = 0.2S_t - (K - S_t)^+$ and $\mu_t^p(K)P_t(K) = 0.2S_t - (S_t - K)^+$ under our assumption that $8/9 \leq K/S_t \leq 1.1$. Therefore, the left-hand side of equation (1.B.3) simplifies to

$$\frac{1}{S_t}\left[-(0.2S_t - (K - S_t)^+) + (0.2S_t - (S_t - K)^+)\right] = -\frac{S_t - K}{S_t}. \tag{1.B.4}$$

When $\theta_t^c(K) > 0$ and $\theta_t^p(K) < 0$, the margin rule is $\mu_t^c(K) = 1$ and $\mu_t^p(K)P_t(K) = 0.2S_t - (S_t - K)^+$ and the left-hand side of equation (1.B.3) is calculated as

$$\frac{1}{S_t}\left[C_t + (0.2S_t - (S_t - K)^+)\right] = 0.2 + [C_t - (S_t - K)^+]/S_t. \tag{1.B.5}$$

Finally, when $\theta_t^c(K) < 0$ and $\theta_t^p(K) > 0$, the margin rule becomes $\mu_t^c(K)C_t(K) = 0.2S_t - (K - S_t)^+$ and $\mu_t^p(K) = 1$ and the left-hand side of equation (1.B.3) becomes

$$-\frac{1}{S_t}\left[P_t + (0.2S_t - (K - S_t)^+)\right] = -\left(0.2 + [P_t - (K - S_t)^+]/S_t\right). \tag{1.B.6}$$

These complete the proof of equation (1.27). □

Next, we prove that the spread between the uncollateralized and collateralized bond rates coincides with $R_{t,t+1}^0 \lambda_t^{MC}/u'(c_t)$ if we extend our model to include both the collateralized and uncollateralized risk-free bonds. Let $R_{t,t+1}^u$ be the return of the uncollateralized bond and θ_t^u be the agent's position on the uncollateralized bond. The equations for the consumption (1.4) and the margin constraint (1.7) change to

$$c_t = W_t - \theta_t^0 - \theta_t^u - \theta_t^S S_t - \sum_{K \in \mathcal{K}_t}\left[\theta_t^c(K)C_t(K) + \theta_t^p(K)P_t(K)\right] - TC_t(\Delta\boldsymbol{\theta}_t),$$

$$\begin{aligned} g_t^{MC}(\boldsymbol{\theta}_t) &= W_t - \theta_t^u - |\theta_t^S|\mu_t^S S_t - \\ &= \sum_{K \in \mathcal{K}_t}\left[|\theta_t^c(K)|\mu_t^c(K)C_t(K) + |\theta_t^p(K)|\mu_t^p(K)P_t(K)\right] \geq 0, \end{aligned}$$

respectively. For simplicity, we assume that no transaction costs are needed to trade the risk-free bonds. Then, the first-order conditions of the collateralized bond (θ_t^0) is unchanged and given by $1 = \mathbb{E}_t^{\mathbb{P}}[m_{t,t+1}^* R_{t,t+1}^0]$, whereas the first order condition of the uncollateralized bond (θ_t^u) is $1 = \mathbb{E}_t^{\mathbb{P}}[m_{t,t+1}^* R_{t,t+1}^u] - \lambda_t^{MC}/u'(c_t)$. From these two first order conditions, we obtain $R_{t,t+1}^u - R_{t,t+1}^0 = R_{t,t+1}^0 \lambda_t^{MC}/u'(c_t)$. □

1.B.1.3 Why is $U_{t,t+1}$ negligible although market frictions affect option returns?

Even though Frazzini and Pedersen (2012) and Hitzemann et al. (2017) document that the embedded leverage effect and the margin constraints have a non-negligible effect on option returns, respectively, our analysis suggests that these two types of market frictions have a negligible effect on $\overline{U}_{t,t+1}$. This is possible due to the following two reasons.

First, the findings in the previous literature on option returns suggest that the ratios $M_t^c(K)/C_t(K)$ and $M_t^p(K)/P_t(K)$ are not negligible. To see this point for the case of call options (put options can be treated similarly), the transformation of the first-order condition of the call option, equation (1.10), yields

$$\mathbb{E}_t^{\mathbb{P}}[R_{t,t+1}^c(K)]-R_{t,t+1}^0 = -R_{t,t+1}^0 Cov_t^{\mathbb{P}}(m_{t,t+1}^*, R_{t,t+1}^c(K))-R_{t,t+1}^0\frac{M_t^c(K)}{C_t(K)}+R_{t,t+1}^0\frac{\frac{\partial TC_t}{\partial \theta_t^c(K)}}{C_t(K)}, \tag{1.B.7}$$

where the second term in the right-hand side denotes the effect of constraints on call option returns. On the other hand, $U_{t,t+1}(K)$ becomes small if $M_t^c(K)/S_t$ and $M_t^p(K)/S_t$ are small (equation (1.16)). These ratios are much smaller than the ratios $M_t^c(K)/C_t(K)$ and $M_t^p(K)/P_t(K)$ because the denominators of the former (the stock price) is much larger than those of the latter (option prices).[27] Therefore, it is possible that the effect of market frictions on option returns is not negligible, yet $U_{t,t+1}(K)$ is negligible.

Second, $U_{t,t+1}(K)$ is proportional to the difference between $M_t^c(K)$ and $M_t^p(K)$ and hence they would mostly offset each other in the case where they have the same sign and they are of similar size. Their signs are always the same for the embedded leverage effect model because $M_t^c(K) = k\Delta_t^c(K)$ and $M_t^p(K) = k|\Delta_t^p(K)|$ always have the same sign (see equation (1.B.2)). For the margin constraints model, $M_t^c(K)$ and $M_t^p(K)$ have the same sign when the agent's allocations to call and put options have the same sign. We find that call option and put options' "expensiveness" are strongly correlated, and hence the signs of $\theta_t^c(K)$ and $\theta_t^p(K)$ determined based on equation (1.36) are the same for approximately 90% of call and put option pairs used

[27] For example, $C_t(K)/S_t$ is less than 0.05 for a one-month at-the-money option price under the Black and Scholes (1973) model when a typical value of 40% for equity's volatility is used.

to estimate AVE-CM CFER, regardless of the three choices of the threshold value s. Moreover, the degree of the offsetting effect is stronger for near at-the-money (ATM) options because the absolute value of the delta of call and put options are similar, and the amount of margins required for call and put options are similar when the sign of allocations to them (θ_t^c and θ_t^p) are the same. These results suggest that the offsetting effect between $M_t^c(K)$ and $M_t^p(K)$ reduces the size of $U_{t,t+1}(K)$ for most of the cases, especially for near ATM options which we use in the estimation of CFER.

1.B.2 Transaction costs for trading options and option mid-prices

1.B.2.1 A theoretical approximation

For brevity, we describe the model for call prices. Put prices are described similarly. We assume that the agent in our model is an option market-maker. In line with the theoretical model considered in Muravyev (2016), we assume that the bid and ask option prices $C_t^{bid}(K)$ and $C_t^{ask}(K)$, respectively, are the end-of-period pre-trade price set by the market-maker conditional on getting a fixed moderate size trade x. These are given by

$$\begin{aligned} C_t^{bid}(K) &= \mathbb{E}_{t_b}^{\mathbb{P}}[m_{t,t+1}^*(\boldsymbol{\theta}_t + x)(S_{t+1} - K)^+] + M_t^c(K) - \varphi x + \vartheta_t^b, \\ C_t^{ask}(K) &= \mathbb{E}_{t_a}^{\mathbb{P}}[m_{t,t+1}^*(\boldsymbol{\theta}_t - x)(S_{t+1} - K)^+] + M_t^c(K) + \varphi x + \vartheta_t^a, \end{aligned} \tag{1.B.8}$$

where $\mathbb{E}_{t_b}^{\mathbb{P}}$ ($\mathbb{E}_{t_a}^{\mathbb{P}}$) denotes the conditional expectation given the information up to time t *and* getting a sell (buy) order and $\varphi \geq 0$ is the unit fixed costs. The terms ϑ_t^b and ϑ_t^a represent microstructure noise caused by tick size and other frictions in the bid and ask price, respectively, and assumed to be zero mean random variables which are uncorrelated with other variables.

Let $C_t^{exTC}(K) = \mathbb{E}_t^{\mathbb{P}}[m_{t,t+1}^*(\boldsymbol{\theta}_t)(S_{t+1} - K)^+] + M_t^c(K)$ be the ex-transaction costs call price. Then,

$$\begin{aligned} C_t^{bid}(K) &= C_t^{exTC}(K) + I_t^{bid}(K) + J_t^{bid}(K) - \varphi x + \vartheta_t^b, \\ C_t^{ask}(K) &= C_t^{exTC}(K) + I_t^{ask}(K) + J_t^{ask}(K) + \varphi x + \vartheta_t^a, \end{aligned} \tag{1.B.9}$$

and $I_t^{bid}(K)$ and $I_t^{ask}(K)$ denote the adverse selection risk part,

$$I_t^{bid}(K) = \mathbb{E}_{t_b}^{\mathbb{P}}[m_{t,t+1}^*(\boldsymbol{\theta}_t + x)(S_{t+1} - K)^+] - \mathbb{E}_t^{\mathbb{P}}[m_{t,t+1}^*(\boldsymbol{\theta}_t + x)(S_{t+1} - K)^+],$$
$$I_t^{ask}(K) = \mathbb{E}_{t_a}^{\mathbb{P}}[m_{t,t+1}^*(\boldsymbol{\theta}_t - x)(S_{t+1} - K)^+] - \mathbb{E}_t^{\mathbb{P}}[m_{t,t+1}^*(\boldsymbol{\theta}_t - x)(S_{t+1} - K)^+], \tag{1.B.10}$$

while $J_t^{bid}(K)$ and $J_t^{ask}(K)$ denote the inventory risk part,

$$\begin{aligned} J_t^{bid}(K) &= \mathbb{E}_t^{\mathbb{P}}[(m_{t,t+1}^*(\boldsymbol{\theta}_t + x)(S_{t+1} - K)^+ - m_{t,t+1}^*(\boldsymbol{\theta}_t)(S_{t+1} - K)^+], \\ J_t^{ask}(K) &= \mathbb{E}_t^{\mathbb{P}}[(m_{t,t+1}^*(\boldsymbol{\theta}_t - x)(S_{t+1} - K)^+ - m_{t,t+1}^*(\boldsymbol{\theta}_t)(S_{t+1} - K)^+]. \end{aligned} \tag{1.B.11}$$

Then, the mid-call price $C_t^{mid}(K) = (C_t^{ask}(K) + C_t^{bid}(K))/2$ equals

$$C_t^{mid}(K) = C_t^{exTC}(K) + \frac{I_t^{bid}(K) + I_t^{ask}(K)}{2} + \frac{J_t^{bid}(K) + J_t^{ask}(K)}{2} + \frac{\vartheta_t^b + \vartheta_t^a}{2}. \tag{1.B.12}$$

Under the standard assumption in the market microstructure literature that the bid and ask prices are linear in the impact of the order flow (see Muravyev, 2016 and references therein), it follows that $I_t^{bid}(K) \approx -I_t^{ask}(K)$. Moreover, for a moderate size of expected trade x, the first-order Taylor series approximation of the IMRS parts $m_{t,t+1}^*(\boldsymbol{\theta}_t + x)$ and $m_{t,t+1}^*(\boldsymbol{\theta}_t - x)$ around $\boldsymbol{\theta}_t$ shows that the inventory risk part is also symmetric in the first-order approximation level, that is, $J_t^{bid}(K) \approx -J_t^{ask}(K)$. Note also that the fixed costs part φ does not appear in equation (1.B.12) in line with Muravyev (2016); this shows that the fixed costs part does not affect the mid-option price albeit it is documented to be the biggest part of the bid-ask spread. This leads to the following approximation:

$$C_t^{mid}(K) \approx C_t^{exTC}(K) + \frac{\vartheta_t^b + \vartheta_t^a}{2}, \tag{1.B.13}$$

that is, apart from random microstructure noise caused by tick size and other frictions, the mid-option price provides a good approximation of the ex-transaction costs option price.

1.B.2.2 Evaluation of measurement errors: Simulation results

To reinforce that the mid-option price is approximately equal to the ex-transaction costs option price, we assess the size of the market microstructure error terms in equation (1.B.13) by employing a simulation setting when the AVE-CM CFER needs to be calculated. We assume that ϑ_t^a and ϑ_t^b follow a zero-mean *i.i.d.* uniform distribution over $[-d/2, d/2]$, where $d > 0$ is the option tick size. This setup is in line with Bliss and Panigirtzoglou (2002) and Dennis and Mayhew (2009), who assume that observed option prices contain zero-mean measurement errors arising from various frictions such as illiquidity and discrete option price quotes; if the true option price is rounded to the nearest discrete quote, the maximal size of the measurement error is $d/2$. The market microstructure term in the estimated AVE-CM CFER is the average of $4N$ *i.i.d.* uniform random variables scaled by $R^0_{t,t+1}/S_t$ when N pairs of call and put options are used; each call and put option pair involves four measurement error variables (errors in the bid and ask prices of call and put option, respectively). We calculate the probability distribution of this error term numerically.

In our dataset, the one-month risk-free rate is on average $R^0_{t,t+1} = 1.002$ (i.e., 20 bps per month), the median stock price is around $S_t = \$30$. The median number of pairs of call and put options used to estimate AVE-CM CFER is $N = 3$. We follow Dennis and Mayhew (2009) and set d to \$0.1, which is the biggest possible tick size of option quotes according to the CBOE contract specifications. Then, we find that the magnitude of the market microstructure noise term in AVE-CM CFER is less than 4.6 bps with probability greater than 90%.[28] Given that the AVE-CM CFER ranges from -1.24% to +0.89% per month in a 5th to 95th percentile range (see Table 1.1), we conclude that the simulated bias of the microstructure noise stemming from the last term in equation (1.B.13) is negligible.

Furthermore, we examine an alternative, more conservative setup where the information risk and the inventory risk components are not necessarily symmetric. In this situation, the maximum size of the measurement error in the mid-price (i.e., the last three terms in equation (1.B.12) altogether) is the half of the bid-ask spread. To assess the size of the measurement error in AVE-CM CFER under this situation, we assume that the measurement error in each option follows a zero-mean *i.i.d.* uniform distribution over $[-d/2, d/2]$, where d is now set to the half of the bid-ask spread. The measurement error term in the estimated AVE-CM CFER is the average of $2N$ *i.i.d.* uniform random variables (errors in the mid call and put prices, respectively) scaled

[28] For simplicity, we assume that the equally-weighted average is taken across strikes and two maturities.

by $R^0_{t,t+1}/S_t$. We set $d = \$0.25$, which is the average bid-ask spread of the options we use for the calculation of AVE-CM CFER. We find that the measurement error term in AVE-CM CFER is less than 16 bps with probability greater than 90%, which is still negligible compared to the magnitude of AVE-CM CFER.

1.C Description of variables

Relative bid-ask spread (BAS): We calculate the daily relative bid-ask spread as $BAS^i_d = (S^{ask,i}_d - S^{bid,i}_d)/(0.5(S^{ask,i} + S^{bid,i}))$. Then, we average the daily bid-ask spread over the past one year. We require there are at least 200 non-missing observations. Data are obtained from the CRSP database.

Amihud's illiquidity measure: We calculate daily Amihud's (2002) illiquidity measure as the ratio of the absolute daily return to the dollar trading volume, $Illiq^i_d = |R^i_d|/(S^i_d Vol^i_d)$, where R^i_d and Vol^i_d are the daily return and the trading volume of i-th stock on day d. Then, we average daily illiquidity measure over the past one year. We require there are at least 200 non-missing observations. The stock returns, stock prices, and trading volumes are obtained from the CRSP database. The trading volume of the NASDAQ equities is adjusted by following Gao and Ritter (2010).

SIZE: Size is the natural logarithm of the market equity. The market equity is calculated as the product of the number of outstanding share with the price of the stock at the end of each month. Data are obtained from the CRSP database.

Idiosyncratic volatility (IVOL): In each month, we regress the daily excess returns over the past 12 months on the Fama and French (1993) three factors to obtain the residual time-series ε^i_d. Then, we calculate the idiosyncratic volatility (IVOL) as

$$IVOL^i_t = \sqrt{\frac{1}{N(d)-1}\sum_{d\in D}(\varepsilon^i_d)^2},$$

where D is the set of non-missing days in the past 12 months. We require there are at least 200 non-missing observations. Stock return data are obtained from the CRSP database and the Fama and French (1993) three factors data are obtained from Kenneth French's website.

Beta: In each month, we regress daily stock excess returns over past 12 months on the daily excess market return to obtain the beta. We require there are at least 200 non-missing observations. Stock return data are obtained from the CRSP database. We use the excess market return provided at Kenneth French's website.

Relative short interest (RSI): The relative short interest (RSI) is calculated as the ratio of the number of short interest to the number of outstanding share. The short interest data is obtained from the Compustat North America, Supplemental Short Interest File via the WRDS. Until the end of 2006, the Compustat records the short interest at the middle of any given month (typically 15th day of each month). Since 2007, the short interest file contains the short interest at the middle of months and the end of months. We use the end-of-month short interest data since 2007 because we sort stocks in portfolios at the end-of-each month in our analysis. The number of outstanding share is obtained from the CRSP database.

Estimated shorting fee (ESF): We follow Boehme et al. (2006) to calculate the estimated shorting fee as

$$\begin{aligned} ESF = {} & 0.07834 + 0.05438 \cdot VRSI - 0.00664 \cdot VRSI^2 + 0.000382 \cdot VRSI^3 \\ & - 0.5908 \cdot Option + 0.2587 \cdot Option \cdot VRSI \\ & - 0.02713 \cdot Option \cdot VRSI^2 + 0.0007583 \cdot Option \cdot VRSI^3, \end{aligned}$$

where $VRSI$ is the *vicile* ranking of the RSI, that is, $VRSI$ takes the value 1 if the firm's RSI is below 5th percentile, 2 if the RSI is between 5th and 10th percentile and so on. *Option* is a dummy variable that takes 1 if option trading volume in the month is non-zero and takes 0 otherwise. Option trading volume data is obtained from the OM database.

Book-to-Market equity (B/M): We follow Davis et al. (2000) to measure book equity as stockholders' book equity, plus balance sheet deferred taxes and investment tax credit (Compustat annual item TXDITC) if available, minus the book value of preferred stock. Stockholders' equity is the value reported by Compustat (item SEQ), if it is available. If not, we measure stockholders' equity as the book value of common equity (item CEQ) plus the par value of preferred stock (item PSTK), or the book value of assets (item AT) minus total liabilities (item LT). Depending on availability, we use redemption (item PSTKRV), liquidating (item PSTKL), or par value (item PSTK) for the book value of preferred stock. From June of each year t to May of $t+1$, the book-to-market equity (B/M) is calculated as the ratio of the book equity for the fiscal year ending in calendar year $t-1$ to the market equity at the end of December of year $t-1$. We treat non-positive B/M data as missing.

Profitability: We follow Fama and French (2015) to measure profitability as revenues (Compustat annual item REVT) minus cost of goods sold (item COGS) if available, minus selling, general, and administrative expenses (item XSGA) if available, minus interest expense (item XINT) if available all divided by (non-lagged) book equity. From June of year t to May of $t+1$, we assign profitability for the fiscal year ending in calendar year $t-1$.

Investment: We follow Fama and French (2015) to measure investment as the change in total assets (Compustat annual item AT) from the fiscal year ending in year $t-1$ to the fiscal year ending in t, divided by $t-1$ total assets. From June of year t to May of $t+1$, we assign investment for the fiscal year ending in calendar year $t-1$.

Turnover rate: We calculate daily turnover rate as the ratio of trading volume to the number of outstanding share. Then, we average daily turnover rate over the past one year. We require there are at least 200 non-missing observations. Trading volume and the number of outstanding share are obtained from the CRSP database. The trading volume of the NASDAQ equities is adjusted by following Gao and Ritter (2010).

Implied-volatility spread (IVS): We follow Bali and Hovakimian (2009) to construct IVS. Specifically, we keep IV data for options which have (i) positive bid price, (ii) positive open interest, (iii) bid-ask spread is smaller than

50% of the mid price. Then, we average all available IVS extracted from options with maturities between 30 days and 91 days and with the absolute value of the log moneyness $|\log(K/S)|$ smaller than 0.1.

DOTS: We follow Goncalves-Pinto et al. (2019) to keep pairs of call and put options with the same maturity and strike if (i) their day-to-maturity is between 8-days and 31-days, (ii) their IV does not exceed 250%, (iii) their bid prices are strictly positive and (iv) their open interest is greater than zero.

On each end of month t, DOTS of i-th stock at j-th strike price is calculated as follows:

$$DOTS^i_{t,j} = \frac{\frac{S^{i,U}_j + S^{i,L}_j}{2} - S^i_t}{S^i_t},$$

where $S^{i,U}_j = C^{i,ask}_t(K_j) - P^{i,bid}_t(K_j) + K_j + PVD^i_t$ and $S^{i,L}_j = C^{i,bid}_t(K_j) - P^{i,ask}_t(K_j) + PVK^i_{t,j}$. PVD^i_t and $PVK^i_{t,j}$ are the present value of dividend payments and the strike price K_j. Then, DOTS of i-th stock in month t is calculated as

$$DOTS^i_t = 100 \cdot \sum_{j=1}^{J} \frac{CPBAS^i_t(K_k)^{-1}}{\sum_{k=1}^{J} CPBAS^i_t(K_j)^{-1}} DOTS^i_{t,j},$$

where $CPBAS^i_t(K) = C^{i,ask}_t(K) - C^{i,bid}_t(K) + P^{i,ask}_t(K) - P^{i,bid}_t(K)$ is the sum of call and put option bid-ask spreads for i-th stock at strike K and J is the number of option pairs. Option and dividend data are obtained from the OM database.

1.D The risk-free bond market with market frictions

In the main body, we assume that market frictions have no effect on the risk-free bond to keep the exposition simple. In this Appendix, we provide the extended model where we relax this assumption. Then, we show that this modification has a negligible effect on our model-free CFER measure. This justifies our approach in the main model to employ a simplifying assumption regarding the effect of frictions on the risk-free bond market.

1.D.1 The generalized definition of CFER

For simplicity, we assume that trading the risk-free bond does not incur transaction costs. This is a reasonable approximation because transaction costs for trading the risk-free bond is much smaller than those for trading individual stocks and options. Therefore, in what follows, the effect of market frictions means the effect originated from the constraints on the allocation (equation (1.6)). In the case where market frictions affect the risk-free bond, the first-order condition for the bond price is analogous to equation (1.9) and it is given by

$$1 = \mathbb{E}_t^{\mathbb{P}}[m_{t,t+1}^* R_{t,t+1}^0] + M_{t,t+1}^0, \tag{1.D.1}$$

where $M_{t,t+1}^0 = \left[\sum_{l=1}^{L} \lambda_t^l \partial g_t^l(\boldsymbol{\theta}_t)/\partial\theta_t^0\right]/u'(c_t)$. By defining $\widetilde{R}_{t,t+1}^0 = 1/\mathbb{E}_t^{\mathbb{P}}[m_{t,t+1}^*]$, equation (1.D.1) can be transformed into

$$\widetilde{R}_{t,t+1}^0 - R_{t,t+1}^0 = \widetilde{R}_{t,t+1}^0 M_{t,t+1}^0. \tag{1.D.2}$$

The difference $\widetilde{R}_{t,t+1}^0 - R_{t,t+1}^0$ can be interpreted as the effect of frictions on the risk-free *rate*. Consistent with this interpretation, equation (1.D.2) shows that $\widetilde{R}_{t,t+1}^0 = R_{t,t+1}^0$ holds if and only if the effect of frictions on the risk-free bond is zero (i.e., $M_{t,t+1}^0 = 0$).

By using the new expression for $\mathbb{E}_t^{\mathbb{P}}[m_{t,t+1}^*]$ to transform the Euler equation (1.9), the asset pricing equation is generalized to

$$\begin{aligned}\mathbb{E}_t^{\mathbb{P}}[R_{t,t+1}] - R_{t,t+1}^0 = &-\frac{Cov_t^{\mathbb{P}}(m_{t,t+1}^*, R_{t,t+1})}{\mathbb{E}_t^{\mathbb{P}}[m_{t,t+1}^*]} + (\widetilde{R}_{t,t+1}^0 - R_{t,t+1}^0) \\ &- \frac{\widetilde{R}_{t,t+1}^0}{S_t}\left(M_{t,t+1}^S - \frac{\partial TC_t}{\partial\theta_t^S} + \frac{\beta}{u'(c_t)}\mathbb{E}_t^{\mathbb{P}}\left[\frac{\partial V_{t+1}}{\partial\theta_t^S}\right]\right).\end{aligned} \tag{1.D.3}$$

Since CFER is the part of the expected return which is not explained by the covariance risk premium, the definition of CFER is generalized to

$$CFER_{t,t+1} = -\widetilde{R}_{t,t+1}^0 \frac{N_t^S}{S_t} + (\widetilde{R}_{t,t+1}^0 - R_{t,t+1}^0), \quad where \tag{1.D.4}$$

$$N_t^S = M_t^S - \frac{\partial TC_t}{\partial\theta_t^S} + \beta\mathbb{E}_t^{\mathbb{P}}\left[\frac{\partial V_{t+1}}{\partial\theta_t^S}\right]. \tag{1.D.5}$$

Note that equation (1.D.4) nests the definition of CFER presented in the main body of this Chapter; in the case where the risk-free rate is not affected by frictions, the equality $\widetilde{R}^0_{t,t+1} = R^0_{t,t+1}$ holds and equation (1.D.4) boils down to equation (1.12). Moreover, this generalization of the definition of CFER does not affect the cross-sectional variation in CFER because $\widetilde{R}^0_{t,t+1} - R^0_{t,t+1}$ is common across all stocks; the variation in CFER is determined again by the value of the first term in equation (1.D.4) just as it was the case with the definition of CFER in the main body. As a result, the evidence on the cross-sectional predictability of CFER presented in Section 1.4 is not affected by assuming that there is no effect of market frictions on the risk-free bond market.

1.D.2 Scaled deviations from put-call parity

For the ease of exposition, in what follows we fix K and suppress the strike argument K in option-related terms. Under the generalized framework, the synthetic stock price $\widetilde{S}_t = C_t - P_t + (K + D_{t+1})/R^0_{t,t+1}$ becomes

$$\begin{aligned}\widetilde{S}_t = \mathbb{E}^{\mathbb{P}}_t[m^*_{t,t+1}(S_{t+1} + D_{t+1})] + (M^c_t - M^p_t) - \left(\frac{\partial TC_t}{\partial \theta^c_t} - \frac{\partial TC_t}{\partial \theta^p_t}\right) \\ + \left(\frac{1}{R^0_{t,t+1}} - \frac{1}{\widetilde{R}^0_{t,t+1}}\right)(K + D_{t+1}).\end{aligned} \tag{1.D.6}$$

The last term in the right-hand side of equation (1.D.6) reflects the fact that the effect of frictions on the risk-free bond transmits to the synthetic stock price because the synthetic stock position involves the investment in the risk-free bond by the amount of $K + D_{t+1}$. Then, taking the difference between S_t and $\widetilde{S}_t$ yields

$$S_t - \widetilde{S}_t = N^S_t - (M^c_t - M^p_t) + \left(\frac{\partial TC_t}{\partial \theta^c_t} - \frac{\partial TC_t}{\partial \theta^p_t}\right) - \frac{\widetilde{R}^0_{t,t+1} - R^0_{t,t+1}}{R^0_{t,t+1}\widetilde{R}^0_{t,t+1}}(K + D_{t+1}). \tag{1.D.7}$$

Scaling the both sides of equation (1.D.7) by $-R^0_{t,t+1}/S_t$ yields

$$\begin{aligned}CFER^{MF}_{t,t+1} &= \frac{R^0_{t,t+1}}{S_t}(\widetilde{S}_t - S_t) \\ &= -R^0_{t,t+1}\frac{N^S_t}{S_t} - U_{t,t+1} - T_{t,t+1} + \left(\widetilde{R}^0_{t,t+1} - R^0_{t,t+1}\right)\frac{K + D_{t+1}}{\widetilde{R}^0_{t,t+1}S_t},\end{aligned} \tag{1.D.8}$$

where $U_{t,t+1}(K)$ and $T_{t,t+1}(K)$ is the same as in Theorem 1.2.2. Then, subtracting equation (1.D.8) from (1.D.4) and rearranging terms yields

$$CFER_{t,t+1} = CFER^{MF}_{t,t+1} + U_{t,t+1} + T_{t,t+1} - (\widetilde{R}^0_{t,t+1} - R^0_{t,t+1}) \left[\frac{K + D_{t+1}}{\widetilde{R}^0_{t,t+1} S_t} - 1 + \frac{N^S_{t,t+1}}{S_t} \right]. \tag{1.D.9}$$

Equation (1.D.9) shows that $CFER_{t,t+1}$ now contains an additional unobservable term

$$\underbrace{-(\widetilde{R}^0_{t,t+1} - R^0_{t,t+1}) \left[\frac{K + D_{t+1}}{\widetilde{R}^0_{t,t+1} S_t} - 1 \right]}_{A_1} \underbrace{-(\widetilde{R}^0_{t,t+1} - R^0_{t,t+1}) \frac{N^S_{t,t+1}}{S_t}}_{A_2} \tag{1.D.10}$$

in addition to $U_{t,t+1}$ and $T_{t,t+1}$ terms (see equation (1.20)).

1.D.3 The evaluation of the additional term

In what follows, we demonstrate that the additional terms in equation (1.D.10) due to a non-zero effect of frictions to the risk-free bond market is negligible. Therefore, the model-free CFER (i.e., scaled deviations from put-call parity) still proxies the true CFER accurately even when we allow the risk-free bond to be affected by frictions.

We begin by discussing a plausible value for $\widetilde{R}^0_{t,t+1} - R^0_{t,t+1}$, which measures the size of the effect of market frictions on the risk-free rate. The effect of frictions on the risk-free rate has been studied extensively from the perspective of the *risk-free rate puzzle*; the empirically observed risk-free is too low given standard theoretical models with plausible values for the preference parameters. A strand of studies consider a model with market frictions, especially the borrowing constraints, which makes the observed risk-free rate (i.e., $R^0_{t,t+1}$) lower than $\widetilde{R}^0_{t,t+1}$. The consensus in this literature is that $\widetilde{R}^0_{t,t+1} - R^0_{t,t+1} > 0$.

However, there is no consensus on its empirical magnitude. Kogan et al. (2007) consider a borrowing constraints model and report a calibrated simulation result that the short-term risk-free rate is lowered possibly by 1.5% per year. Constantinides et al. (2002) calibrate their borrowing constrained model and report that the borrowing constraints lower the long-term bond rate by about 4% to 6% per year. Even though this value is much higher than that in Kogan et al. (2007), it may be due to the difference in the maturity of bond under consideration. Heaton and Lucas (1996) examine whether borrowing constraints and transaction costs can solve the equity risk premium puzzle and the risk-free rate puzzle simultaneously. They report that unrealistically large transaction costs are necessary to decrease the risk-free rate to solve the risk-free rate puzzle, and in such a case, the model's risk-free rate decreases by about 4% per year. Given these results in the previous literature, we set $\widetilde{R}^0_{t,t+1} - R^0_{t,t+1}$ to 0.5% per month (6% per year) in the subsequent discussion.

Now, we evaluate A_1 and A_2 in equation (1.D.10) separately. First, we transform A_1 as

$$A_1 = \frac{\widetilde{R}^0_{t,t+1} - R^0_{t,t+1}}{\widetilde{R}^0_{t,t+1}} \left[\left(\frac{K}{S_t} - 1 + \frac{D_{t+1}}{S_t} \right) - (\widetilde{R}^0_{t,t+1} - 1) \right]. \tag{1.D.11}$$

Then, taking the absolute value and using $\widetilde{R}^0_{t,t+1} \geq R^0_{t,t+1} \geq 1$ yields

$$|A_1| \leq |\widetilde{R}^0_{t,t+1} - R^0_{t,t+1}| \times \left(\left| \frac{K}{S_t} - 1 \right| + \frac{D_{t+1}}{S_t} + |\widetilde{R}^0_{t,t+1} - 1| \right). \tag{1.D.12}$$

To further evaluate this inequality, recall that we use only near at-the-money options ($|K/S_t - 1| \leq 0.1$). Given a plausible yet conservative dividend yield of 4% per year, D_{t+1}/S_t is about 0.01 (= 4%/4) because U.S. firms typically pay dividends quarterly. The value of $\widetilde{R}^0_{t,t+1} - 1 = (\widetilde{R}^0_{t,t+1} - R^0_{t,t+1}) + (R^0_{t,t+1} - 1)$ is at most 1% per month (i.e., $\widetilde{R}^0_{t,t+1} \leq 1.01$) under our numerical assumption on $\widetilde{R}^0_{t,t+1} - R^0_{t,t+1}$ because the observed risk-free rate $R^0_{t,t+1} - 1$ is at most 0.5% per month during our sample period. Therefore, we obtain

$$|A_1| \leq 0.5\% \times (0.1 + 0.01 + 0.01) \approx 6 \text{ bps}. \tag{1.D.13}$$

Note that this is an upper bound and usually $|A_1|$ is much smaller because the moneyness of options used is much closer to one (i.e., $|K/S_t - 1| \approx 0$). Therefore, we can conclude that the A_1 term is negligible compared to the variation in the estimated $CFER^{MF}_{t,t+1}$.

Next, we evaluate A_2. To this end, by ignoring the negligible $U_{t,t+1}$, $T_{t,t+1}$ and A_1 terms in equation (1.D.9), we obtain the following approximation relation:

$$CFER^{MF}_{t,t+1} \approx CFER_{t,t+1} + A_2 = -R^0_{t,t+1} \frac{N^S_{t,t+1}}{S_t} + (\widetilde{R}^0_{t,t+1} - R^0_{t,t+1}). \tag{1.D.14}$$

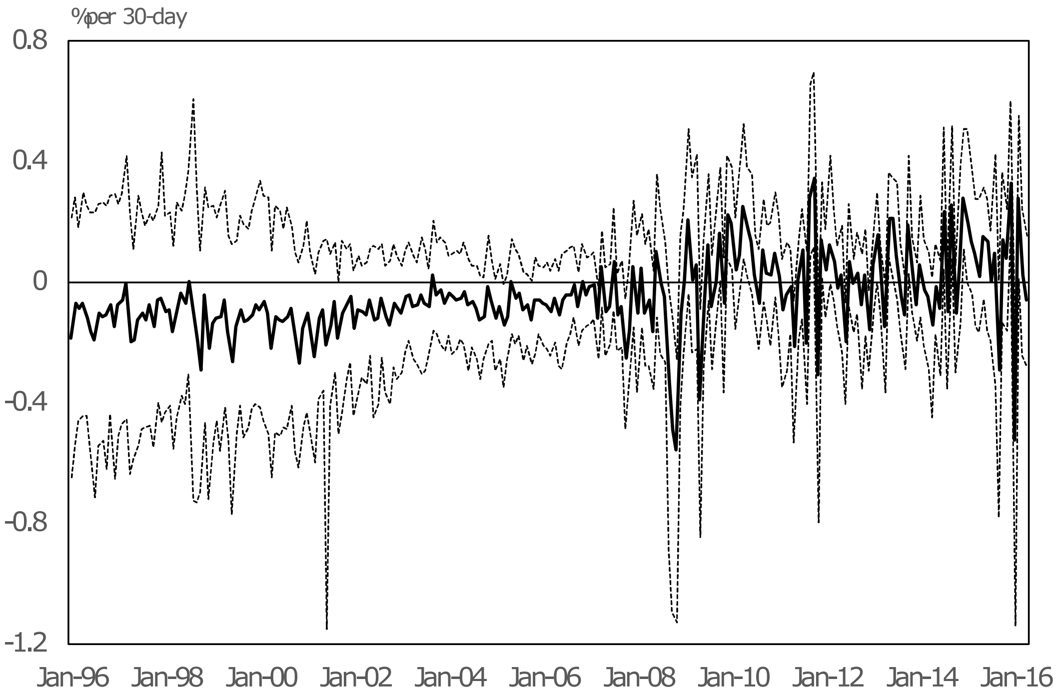
With some more algebra, we obtain the following approximation relation:

$$A_2 = (\widetilde{R}^0_{t,t+1} - R^0_{t,t+1})\frac{N^S_{t,t+1}}{S_t} \approx -\frac{(\widetilde{R}^0_{t,t+1} - R^0_{t,t+1})}{R^0_{t,t+1}} CFER^{MF}_{t,t+1} + \frac{(\widetilde{R}^0_{t,t+1} - R^0_{t,t+1})^2}{R^0_{t,t+1}}. \tag{1.D.15}$$

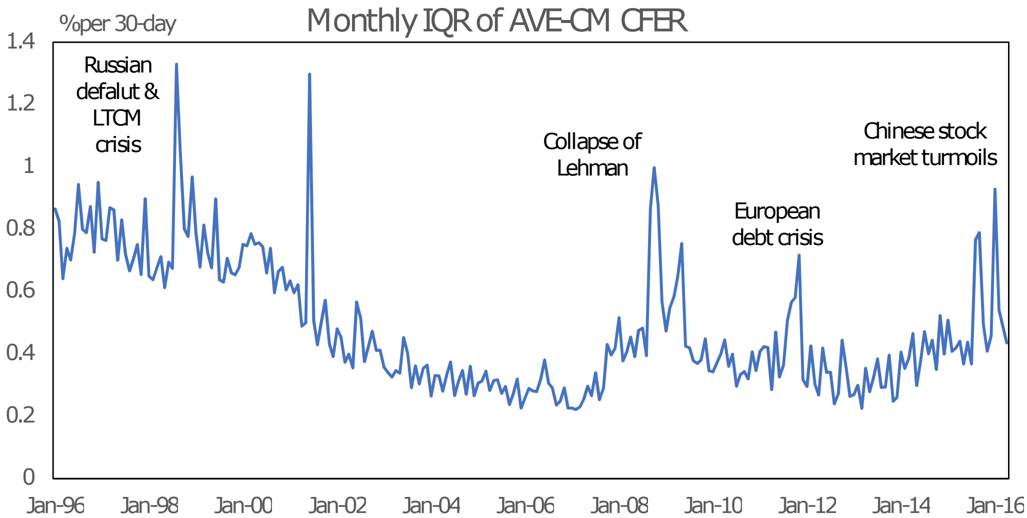
Under the assumption of $\widetilde{R}^0_{t,t+1} - R^0_{t,t+1} = 0.5\%$, the value of the second term in the right-hand side of equation (1.D.15) is negligible (0.25 bps). The first term in the right-hand side of equation (1.D.15) is proportional to $CFER^{MF}_{t,t+1}$ by the factor of $(\widetilde{R}^0_{t,t+1} - R^0_{t,t+1})/R^0_{t,t+1} \approx 0.5\%$. Therefore, this term results in a negligible relative error; when $|CFER^{MF}_{t,t+1}|$ is smaller than 3%, it results in at most only 1.5 bps of absolute error.[29] All in all, we conclude that A_2 is negligible.

To sum up, the additional term in the $CFER^{MF}_{t,t+1}$ formula caused by the effect of frictions on the risk-free rate, equation (1.D.10), is negligible. Therefore, the model-free CFER (i.e., the scaled deviations from put-call parity) is still a good proxy of the true CFER even when market frictions affect the risk-free bond market.

[29]The 1st percentile to 99th percentile range of AVE-CM CFER is from -2.9% to +2.1%. Therefore, $|CFER^{MF}_{t,t+1}| \leq 3\%$ holds for almost all samples except a small number of outliers.



(a) Time Series of the monthly median, 25th and 75th percentile of AVE-CM CFER



(b) Time Series of the monthly IQR of AVE-CM CFER

Figure 1.1. Time Series of the monthly IQR of AVE-CM CFER

Figure 1.1a illustrates the time-series of the monthly median (solid line), the 25th and 75th percentile points (two dotted lines) of AVE-CM CFER. Figure 1.1b illustrates the time-series of the monthly IQR (difference between the 75th and 25th percentile points) of AVE-CM CFER. At the end of each month, we calculate the median, 25th and 75th percentile points, and IQR of the individual stocks' AVE-CM CFER values. The unit of the y-axis is % per 30-day. The estimation period spans January 1996 to April 2016 (244 months).

Table 1.1. Estimated CFER: Summary statistics

Entries in Panel A report the summary statistics of the estimated model-free CFER at the end of each month for the four different ways of estimating CFER. These are denoted by a combination of the method of choosing strikes (AVE or ATM) and the method of choosing maturities (CM or CLS) of options. In AVE methods, (1) and (3), we average CFER across available strikes, whereas in ATM methods, (2) and (4), we choose the strike closest to the forward price. In CM methods, (1) and (2), we interpolate CFER across the estimated CFER of traded maturities to obtain a 30-day constant maturity CFER, while in CLS methods, (3) and (4), we choose the traded maturity closest to 30 days. The row for N reports the total number of month-stock CFER observations, the row for IQR reports the interquartile range (75th minus 25th percentile values), and the last row, % of $CFER < 0$, reports the proportion of observations with negatively estimated CFER. The estimation period spans January 1996 to April 2016 (244 months). The unit of statistics (except skewness, kurtosis, and % of $CFER < 0$) is % per 30-day. Entries in Panel B report the pairwise Pearson correlation coefficients between the four estimated CFER measures.

Panel A: Summary statistics of CFER				
	(1) AVE-CM	(2) ATM-CM	(3) AVE-CLS	(4) ATM-CLS
N	333,234	333,234	347,073	347,073
mean	-0.09	-0.09	-0.10	-0.10
standard deviation	0.88	0.89	1.09	1.10
skewness	-1.95	-1.88	-1.59	-1.53
kurtosis	69.32	68.92	69.97	69.20
minimum	-27.67	-27.67	-35.60	-35.60
5th percentile	-1.24	-1.25	-1.54	-1.55
Median	-0.04	-0.04	-0.04	-0.04
95th percentile	0.89	0.89	1.14	1.15
maximum	24.96	24.96	32.72	32.72
IQR	0.47	0.46	0.60	0.60
% of CFER< 0	55.3%	55.1%	54.9%	54.6%
Panel B: Correlation between different measures of CFER				
	(1) AVE-CM	(2) ATM-CM	(3) AVE-CLS	(4) ATM-CLS
(1) AVE-CM	1			
(2) ATM-CM	0.986	1		
(3) AVE-CLS	0.989	0.974	1	
(4) ATM-CLS	0.973	0.989	0.984	1

Table 1.2. AVE-CM CFER-sorted decile portfolios: Cross-sectional predictability

Entries in Panel A report the average CFER, average post-ranking return and results for the risk-adjusted returns (α) of the AVE-CM CFER-sorted value-weighted decile portfolios and the spread portfolio, with respect to the CAPM and Carhart (1997) four-factor model. On the last trading day of each month t, stocks are sorted in ascending order based on AVE-CM CFER and then value-weighted decile portfolios are formed. We then calculate the return of these portfolios and the spread portfolio in the succeeding month-$(t+1)$. Entries in Panel B report the average CFER, average post-ranking return and alphas of the AVE-CM CFER-sorted equally-weighted decile portfolios and the spread portfolio. The estimation period spans January 1996 to April 2016 (244 months) for both Panels. t-statistics are adjusted for heteroscedasticity and autocorrelation and reported in parentheses. The unit of the average returns and alphas (average CFER) is % per month (30-day). N is the average number of stocks in each decile portfolio.

	AVE-CM CFER-sorted decile portfolios										Spread
	1 (Lowest)	2	3	4	5	6	7	8	9	10 (Highest)	10-1
Panel A: Value-weighted portfolios											
Ave. CFER	-1.31	-0.53	-0.31	-0.18	-0.09	-0.01	0.07	0.18	0.36	0.93	2.24
Ave. return	-0.18	0.23	0.54	0.49	0.79	0.89	0.99	0.92	1.15	1.46	1.64
	(-0.36)	(0.60)	(1.58)	(1.64)	(2.57)	(2.82)	(3.20)	(2.73)	(3.19)	(3.38)	(5.77)
α_{CAPM}	-1.15	-0.58	-0.23	-0.25	0.04	0.13	0.24	0.13	0.33	0.55	1.70
	(-5.69)	(-3.41)	(-1.96)	(-2.47)	(0.42)	(1.29)	(2.30)	(1.16)	(2.30)	(2.77)	(5.91)
α_{FFC}	-1.11	-0.62	-0.26	-0.23	0.00	0.12	0.24	0.18	0.44	0.75	1.86
	(-6.52)	(-3.74)	(-2.28)	(-2.28)	(0.02)	(1.16)	(2.38)	(1.50)	(2.54)	(3.52)	(6.56)
N	134.9	135.0	134.9	135.1	134.7	135.3	134.9	135.0	134.9	135.0	—
Panel B: Equally-weighted portfolios											
Ave. CFER	-1.58	-0.55	-0.32	-0.19	-0.09	-0.01	0.07	0.18	0.37	1.14	2.73
Ave. return	-0.35	0.49	0.65	0.76	0.86	0.97	0.93	1.02	1.08	1.38	1.73
	(-0.65)	(1.09)	(1.54)	(1.94)	(2.24)	(2.62)	(2.42)	(2.58)	(2.49)	(2.76)	(9.10)
α_{CAPM}	-1.41	-0.46	-0.24	-0.12	0.00	0.11	0.07	0.12	0.14	0.35	1.76
	(-5.62)	(-2.63)	(-1.59)	(-0.86)	(0.01)	(0.97)	(0.63)	(1.00)	(0.77)	(1.53)	(9.60)
α_{FFC}	-1.31	-0.45	-0.27	-0.10	-0.04	0.07	0.04	0.13	0.17	0.50	1.81
	(-9.61)	(-3.79)	(-2.55)	(-0.99)	(-0.47)	(0.79)	(0.40)	(1.48)	(1.41)	(2.61)	(9.42)
N	134.9	135.0	134.9	135.1	134.7	135.3	134.9	135.0	134.9	135.0	—

Table 1.3. AVE-CM Full CFER-sorted decile portfolios: Cross-sectional predictability

Entries report the average return and the five risk-adjusted returns (α's) with respect to the CAPM, Fama and French (1993) three-factor model, Carhart (1997) four-factor model, Fama and French (2015) five-factor model, and Stambaugh and Yuan (2017) mispricing-factor model, of the spread portfolio of the CFER-sorted value-weighted decile portfolios. The first two rows report the results based on the model-free AVE-CM CFER, which ignores U_t. Panel A shows the results based on the fully-estimated CFER, where U_t is estimated based on the embedded leverage model. Panel B shows the results based on the fully-estimated CFER, where U_t is estimated based on the margin constraints model. For the margin constraints model, we consider nine alternative fully-estimated AVE-CM CFER. Each one of the nine alternative CFERs corresponds to the value of $R^0_{t,+1}\lambda^{mc}_t$ and the value of the reference point of the option expensiveness s, which determines the value of $E_t(K)$. The analysis spans January 1996 to April 2016 (244 months). t-statistics are adjusted for heteroscedasticity and autocorrelation and reported in the parentheses. The t-statistics of the difference between the baseline result and each one of the fully-estimated CFER-sorted results are reported in square brackets. The unit of all variables is % per 30 days.

		Ave. Ret	α_{CAPM}	α_{FF3}	α_{FFC}	α_{FF5}	α_{SY}
Panel A: Embedded leverage model							
Baseline result		1.64	1.70	1.78	1.86	1.58	1.70
		(5.77)	(5.91)	(6.36)	(6.56)	(5.63)	(5.21)
Fully-estimated		1.54	1.61	1.66	1.75	1.41	1.56
		(6.22)	(6.37)	(6.30)	(6.71)	(5.98)	(5.63)
		[-0.60]	[-0.53]	[-0.74]	[-0.66]	[-0.95]	[-0.76]
Panel B: Margin constraints model							
$R^0_{t,t+1}\psi =$	$s = 0.00$	1.58	1.64	1.75	1.83	1.49	1.63
10% per year		(5.28)	(5.30)	(5.95)	(6.16)	(5.45)	(4.84)
		[-0.56]	[-0.50]	[-0.27]	[-0.23]	[-0.72]	[-0.42]
	$s = 0.01$	1.61	1.66	1.77	1.85	1.50	1.65
		(5.42)	(5.36)	(5.99)	(6.25)	(5.56)	(4.94)
		[-0.25]	[-0.31]	[-0.11]	[-0.08]	[-0.55]	[-0.26]
	$s = 0.02$	1.73	1.79	1.89	1.97	1.65	1.78
		(5.58)	(5.49)	(6.03)	(6.35)	(5.66)	(5.07)
		[0.82]	[0.73]	[0.84]	[0.68]	[0.45]	[0.42]
$R^0_{t,t+1}\psi =$	$s = 0.00$	1.57	1.63	1.71	1.76	1.49	1.56
5% per year		(5.46)	(5.55)	(6.14)	(6.09)	(5.54)	(5.12)
		[-0.94]	[-0.99]	[-0.91]	[-1.03]	[-0.98]	[-1.12]
	$s = 0.01$	1.63	1.69	1.79	1.85	1.53	1.65
		(5.52)	(5.42)	(5.97)	(6.28)	(5.46)	(4.95)
		[-0.10]	[-0.12]	[0.07]	[-0.04]	[-0.44]	[-0.30]
	$s = 0.02$	1.69	1.76	1.86	1.93	1.62	1.76
		(5.54)	(5.61)	(6.15)	(6.43)	(5.63)	(5.20)
		[0.55]	[0.59]	[0.72]	[0.49]	[0.27]	[0.34]
$R^0_{t,t+1}\psi =$	$s = 0.00$	1.60	1.66	1.74	1.78	1.49	1.56
Time-varying		(5.50)	(5.65)	(6.20)	(6.18)	(5.42)	(5.09)
		[-0.64]	[-0.69]	[-0.70]	[-0.86]	[-1.06]	[-1.14]
	$s = 0.01$	1.60	1.66	1.74	1.78	1.50	1.57
		(5.54)	(5.70)	(6.24)	(6.19)	(5.46)	(5.15)
		[-0.58]	[-0.63]	[-0.65]	[-0.87]	[-0.96]	[-1.09]
	$s = 0.02$	1.64	1.71	1.79	1.82	1.55	1.62
		(5.59)	(5.74)	(6.24)	(6.16)	(5.55)	(5.20)
		[0.10]	[0.10]	[0.10]	[-0.42]	[-0.39]	[-0.64]

Table 1.4. CFER-adjusted excess returns: Alphas of CFER-sorted portfolios

Entries in Panel A report the intercepts α_{CAPM} and α_{FFC} of the regressions of CFER-adjusted excess returns $R_{t,t+1} - R^0_{t,t+1} - CFER_{t,t+1}$ on a set of risk factor(s) of the CAPM and Carhart (1997) four-factor model, respectively (i.e., equation (1.32)). On the last trading day of each month t, stocks are sorted in ascending order based on AVE-CM CFER and then value-weighted decile portfolios are formed. We then calculate the average CFER as well as the return in the succeeding month-$(t+1)$ of these portfolios and the spread portfolio to calculate the CFER-adjusted excess return. Entries in Panel B report α_{CAPM} and α_{FFC}, where we eliminate CFER observations below 1st percentile and above 99th percentile point. Entries in Panel C report α_{CAPM} and α_{FFC}, where we form quintile portfolios instead of the decile portfolios. The estimation period spans January 1996 to April 2016 (244 months) for all Panels. t-statistics are adjusted for heteroscedasticity and autocorrelation and reported in the parentheses. The unit of all variables is % per month.

	AVE-CM CFER-sorted value-weighted decile portfolios										Spread
	1 (Lowest)	2	3	4	5	6	7	8	9	10 (Highest)	10-1
Panel A: Decile sort with all available CFER											
α_{CAPM}	0.17	-0.05	0.08	-0.07	0.14	0.15	0.17	-0.04	-0.02	-0.37	-0.55
	(0.94)	(-0.29)	(0.70)	(-0.66)	(1.41)	(1.42)	(1.65)	(-0.38)	(-0.14)	(-1.95)	(-2.31)
α_{FFC}	0.22	-0.08	0.05	-0.05	0.10	0.13	0.17	0.00	0.09	-0.16	-0.39
	(1.44)	(-0.50)	(0.45)	(-0.48)	(1.01)	(1.31)	(1.68)	(0.04)	(0.49)	(-0.79)	(-1.57)
Panel B: Decile sort where CFER below 1st or above 99th percentile are eliminated											
α_{CAPM}	0.02	0.00	0.11	-0.02	0.11	0.12	0.13	0.06	-0.02	-0.33	-0.36
	(0.13)	(0.02)	(1.03)	(-0.16)	(1.18)	(1.12)	(1.26)	(0.52)	(-0.14)	(-1.78)	(-1.57)
α_{FFC}	0.03	-0.02	0.09	0.00	0.08	0.09	0.13	0.08	0.09	-0.16	-0.19
	(0.18)	(-0.12)	(0.91)	(0.05)	(0.80)	(0.94)	(1.32)	(0.79)	(0.52)	(-0.84)	(-0.88)

Panel C: Quintile sort with all available CFER

	AVE-CM CFER-sorted value-weighted quintile portfolios					Spread
	1 (Lowest)	2	3	4	5 (Highest)	5-1
α_{CAPM}	0.00	0.01	0.14	0.08	-0.17	-0.17
	(-0.02)	(0.15)	(1.70)	(1.00)	(-1.29)	(-0.85)
α_{FFC}	-0.01	0.01	0.11	0.09	-0.02	-0.01
	(-0.08)	(0.17)	(1.46)	(1.14)	(-0.14)	(-0.06)

Table 1.5. Predictive power of $CFER^{MF}_{t,t+1}(K)$, Implied Volatility Spread (IVS), and DOTS

Entries report the average return and five risk-adjusted returns (alphas) with respect to the CAPM, Fama and French (1993) three-factor model, Carhart (1997) four-factor model, Fama and French (2015) five-factor model, and Stambaugh and Yuan (2017) mispricing-factor model of the spread portfolio of AVE-CM $CFER^{MF}_{t,t+1}(K)$, IVS and DOTS-sorted value-weighted decile portfolios. For any given sorting criterion, t-statistics for each average return and alpha are reported within parentheses. t-statistics of the difference between the performance measures of the CFER-sorted and IVS (DOTS)-sorted portfolios are reported within square brackets. For this comparison, the null hypothesis is that each pair of sorting criteria yields equal results and the alternative hypothesis is that the CFER-sorted portfolios outperform. The critical value is 1.64. One asterisk denotes significance at a 5% level of significance. The analysis spans January 1996 to April 2016 (244 months). t-statistics are adjusted for heteroscedasticity and autocorrelation and reported in parentheses. The unit of all variables is % per month.

	Average return	α_{CAPM}	α_{FF3}	α_{FFC}	α_{FF5}	α_{SY}
MF-CFER-sorted	1.64	1.70	1.78	1.86	1.58	1.70
	(5.77)	(5.91)	(6.36)	(6.56)	(5.63)	(5.21)
IVS-sorted	1.17	1.23	1.30	1.38	1.14	1.25
	(4.90)	(4.76)	(4.94)	(5.24)	(4.64)	(4.25)
	[2.03*]	[1.91*]	[1.93*]	[1.87*]	[1.66*]	[1.66*]
DOTS-sorted	1.45	1.42	1.46	1.47	1.31	1.28
	(5.61)	(4.93)	(4.99)	(4.98)	(4.70)	(4.19)
	[1.02]	[1.54]	[1.72*]	[1.99*]	[1.31]	[1.89*]

Table 1.6. Bivariate dependent sort: Controlling for option trading volume

Entries report the result of the bivariate dependent sort, where we first sort stocks based on the aggregate option trading volume (OV), and then within each group of the OV level, we further sort stocks into quintile portfolios by the AVE-CM MF-CFER criterion. Rows correspond to the level of the first sorting variable, OV, and the first to the fifth columns correspond to the level of the second sorting variable,AVE-CM MF-CFER. Sixth to eighth columns report the average returns, Carhart (1997) four-factor alpha, and the average CFER of the spared, respectively, of the spread portfolio (the highest CFER portfolio minus the lowest MF-CFER portfolio). The four right columns report the average value of the option trading volume (OV) and the underlying stocks' characteristics, log market capitalization (SIZE), relative bid-ask spread (BAS) and Amihud's (2002) illiquidity measure of five portfolios within each OV level. The estimation period spans January 1996 to April 2016 (244 months). t-statistics are adjusted for heteroscedasticity and autocorrelation (HAC-adjusted t-stat). The unit of the average returns and alphas (average CFER) is % per month (per 30-days). The unit of BAS is % of the mid stock price and Amihud measure is multiplied by 1000 for the readability.

	Ave. ret. of AVE-CM MF-CFER-sorted portfolios					Spread portfolio (5-1)			Average characteristics			
	1 (lowest)	2	3	4	5 (highest)	Ave. ret.	α_{FFC}	Ave. CFER	OV	SIZE	BAS	Amihud
OV 1	0.29	0.81	0.85	1.17	1.60	1.32	1.31	1.83	9	14.78	0.44	4.40
(smallest)	(0.77)	(2.53)	(2.84)	(3.78)	(4.46)	(6.39)	(5.91)	(17.86)				
2	0.53	0.71	0.94	1.11	1.32	0.79	0.85	1.56	64	15.20	0.42	3.27
	(1.40)	(2.19)	(3.03)	(3.46)	(3.58)	(4.37)	(4.78)	(16.67)				
3	0.53	0.47	0.98	1.26	1.23	0.70	0.80	1.36	245	15.62	0.39	2.24
	(1.31)	(1.40)	(3.29)	(3.96)	(3.21)	(3.00)	(3.32)	(13.29)				
4	0.41	0.42	0.93	0.96	1.40	0.99	1.04	1.16	979	16.21	0.37	1.17
	(0.90)	(1.30)	(3.16)	(3.20)	(3.54)	(3.28)	(3.77)	(11.06)				
OV 5	0.03	0.53	0.83	0.94	0.94	0.91	1.04	0.88	32815	17.87	0.30	0.26
(largest)	(0.07)	(1.70)	(2.56)	(2.98)	(2.42)	(3.26)	(3.27)	(12.65)				

Table 1.7. Characteristics of AVE-CM CFER-sorted value-weighted decile portfolios

Entries report the average value of various characteristics of decile portfolios as well as the difference between the highest CFER decile portfolio and the lowest CFER decile portfolio. On the last trading day of each month t, stocks are sorted in ascending order based on AVE-CM CFER and then value-weighted decile portfolios are formed. We then calculate the value-weighted average value of characteristics. BAS is the relative bid-ask spread, Amihud is Amihud's (2002) illiquidity measure (multiplied by 1000), SIZE is the natural log of the market equity, S_t is the stock price level, IVOL is the idiosyncratic volatility, beta is the regression coefficient of stock returns on the market portfolio return, RSI is the relative short-interest, ESF is the estimated shorting fee, B/M is the book-to-market ratio, and N is the number of average stocks in each portfolio. See Appendix 1.C for the detailed description of each variable. The data period spans January 1996 to April 2016 (244 months). t-statistics are adjusted for heteroscedasticity and autocorrelation and reported in the parentheses.

	AVE-CM CFER-sorted value-weighted decile portfolios										Spread
	1 (Lowest)	2	3	4	5	6	7	8	9	10 (Highest)	10-1
CFER	-1.31	-0.53	-0.31	-0.18	-0.09	-0.01	0.07	0.18	0.36	0.93	2.24
	(-15.88)	(-12.34)	(-10.92)	(-9.52)	(-7.03)	(-1.25)	(7.26)	(13.59)	(15.08)	(16.17)	(15.38)
BAS	0.48	0.39	0.35	0.34	0.33	0.31	0.31	0.34	0.37	0.44	-0.04
	(2.67)	(2.51)	(2.14)	(2.31)	(2.15)	(2.18)	(2.25)	(2.34)	(2.28)	(2.69)	(-2.89)
Amihud	5.60	1.84	0.97	0.55	0.37	0.31	0.36	0.55	1.17	3.82	-1.78
	(6.96)	(6.13)	(5.09)	(5.74)	(6.26)	(5.94)	(7.93)	(7.50)	(7.70)	(6.23)	(-4.76)
SIZE	15.34	16.30	16.81	17.14	17.43	17.47	17.46	17.20	16.66	15.76	0.42
	(221.77)	(192.45)	(320.21)	(386.63)	(372.29)	(292.14)	(282.72)	(255.42)	(268.91)	(189.59)	(3.96)
S_t	36.72	49.00	58.69	62.37	68.08	69.36	70.39	60.62	52.92	39.08	2.35
	(23.40)	(23.50)	(18.87)	(20.09)	(22.06)	(22.03)	(16.44)	(18.72)	(18.38)	(15.68)	(1.33)
IVOL	39.53	31.74	28.24	26.37	24.93	24.69	24.89	26.22	29.07	35.21	-4.32
	(18.06)	(13.67)	(10.23)	(10.25)	(9.40)	(9.46)	(9.82)	(10.76)	(13.13)	(12.95)	(-7.30)
Beta	1.20	1.12	1.05	1.02	1.01	1.01	1.03	1.04	1.07	1.16	-0.04
	(53.61)	(79.84)	(132.19)	(86.01)	(69.39)	(59.49)	(72.21)	(102.76)	(102.44)	(55.50)	(-2.04)
RSI	6.19	3.97	2.99	2.52	2.22	2.09	2.19	2.47	3.14	4.28	-1.90
	(19.54)	(30.61)	(46.87)	(39.76)	(35.74)	(22.78)	(19.30)	(21.57)	(32.74)	(29.84)	(-8.41)
ESF	0.57	0.43	0.35	0.30	0.27	0.25	0.26	0.29	0.36	0.47	-0.11
	(13.46)	(8.30)	(8.82)	(8.37)	(6.80)	(8.83)	(9.00)	(8.96)	(9.20)	(9.08)	(-6.52)
B/M	0.53	0.47	0.44	0.43	0.41	0.40	0.40	0.42	0.44	0.48	-0.05
	(13.32)	(19.51)	(22.86)	(19.74)	(13.87)	(13.75)	(13.49)	(13.55)	(16.70)	(16.02)	(-4.30)
N	134.93	135.02	134.89	135.06	134.69	135.25	134.94	135.00	134.91	135.05	—

Table 1.8. Performance of CFER-sorted portfolios: Bivariate dependent sorts controlling for relative bid-ask spread or SIZE

Entries in Panel A report the result of the bivariate dependent sort, where we first sort stocks based on the relative bid-ask spread (BAS), and then within each group of the BAS level, we further sort stocks into quintile portfolios by the AVE-CM CFER criterion. Rows correspond to the level of the first sorting variable, BAS, and the first to the fifth columns correspond to the level of the second sorting variable,AVE-CM CFER. Sixth to the last columns report the average returns, Fama and French (2015) five-factor alpha, and the average CFER, respectively, of the spread portfolio (the highest CFER portfolio minus the lowest CFER portfolio). Entries in Panel B report the result, where we use SIZE (the log of market equity) as the first sorting variable instead of BAS. The estimation period spans January 1996 to April 2016 (244 months). t-statistics are adjusted for heteroscedasticity and autocorrelation (HAC-adjusted t-stat). The unit of the average returns and alphas (average CFER) is % per month (per 30-days).

	Ave. returns of AVE-CM CFER-sorted portfolios					Ave. return	α_{FFC}	Ave. CFER
	1 (Lo)	2	3	4	5 (Hi)	Spread portfolio (5-1)		
Panel A: Relative bid-ask spread-sorted dependent bivariate sort								
BAS 1	0.11	0.63	0.65	0.76	0.84	0.74	0.86	0.75
(narrowest)	(0.25)	(1.65)	(1.54)	(1.75)	(2.19)	(3.47)	(3.29)	(13.71)
BAS 2	0.31	0.48	0.81	0.70	1.09	0.78	0.69	1.05
	(0.72)	(1.24)	(2.37)	(1.78)	(2.59)	(3.31)	(2.96)	(10.63)
BAS 3	0.41	0.59	0.98	0.93	1.30	0.89	0.93	1.34
	(1.01)	(1.52)	(2.56)	(2.26)	(3.03)	(3.26)	(3.22)	(12.30)
BAS 4	0.14	0.70	0.68	0.85	1.48	1.33	1.48	1.70
	(0.30)	(1.56)	(1.69)	(2.06)	(3.22)	(4.60)	(4.62)	(14.90)
BAS 5	-0.42	0.28	0.85	0.78	1.53	1.95	1.94	2.64
(widest)	(-0.72)	(0.57)	(1.79)	(1.64)	(3.29)	(5.54)	(5.90)	(14.59)
Panel B: Size-sorted dependent bivariate sort								
SIZE 1	-0.62	0.48	0.72	0.96	1.22	1.84	1.79	3.18
(smallest)	(-1.00)	(0.85)	(1.31)	(1.63)	(2.06)	(6.54)	(6.39)	(16.56)
SIZE 2	0.16	0.86	0.85	0.94	1.28	1.12	1.12	1.95
	(0.30)	(1.70)	(1.84)	(2.02)	(2.53)	(4.94)	(4.94)	(14.13)
SIZE 3	0.30	0.78	0.86	0.96	1.26	0.96	1.04	1.46
	(0.66)	(1.83)	(2.06)	(2.33)	(3.07)	(4.82)	(4.86)	(15.77)
SIZE 4	0.70	0.76	0.99	1.20	1.21	0.51	0.57	1.01
	(1.72)	(1.98)	(2.77)	(3.35)	(3.27)	(3.01)	(3.14)	(12.43)
SIZE 5	0.33	0.69	0.78	0.88	0.94	0.61	0.67	0.65
(largest)	(1.03)	(2.52)	(2.48)	(2.96)	(2.89)	(3.50)	(3.65)	(12.64)

Table 1.9. CFER and firms' and stocks' characteristics: Fama-MacBeth regressions

Entries in Panel A report the results from Fama and MacBeth (1973) regressions of AVE-CM CFER on SIZE (log of market equity), relative bid-ask spread (BAS), idiosyncratic volatility (IVOL), Amihud's (2002) illiquidity measure, and relative short interest (RSI), where we use positive CFER subsamples. Entries in Panel B report the results, where we use negative CFER subsamples. Even though the intercept is included in the regressions, we do not report them due to space limitations. The time-series averages of the estimated coefficients of the cross-sectional regressions are reported. t-statistics are adjusted for heteroscedasticity and autocorrelation and reported in the parentheses. The time-series averages of adjusted R^2 and the number of observations N employed in the cross-sectional regressions are reported in the last two rows of each Panel. The data period spans January 1996 to April 2016 (244 months).

	(1)	(2)	(3)	(4)	(5)	(6)
Panel A: Positive CFER subsample						
SIZE	-0.11					-0.06
	(-14.46)					(-14.92)
BAS		1.22				0.55
		(4.52)				(6.24)
IVOL			0.78			0.19
			(15.13)			(3.65)
Amihud				17.94		1.38
				(6.98)		(0.67)
RSI					0.85	-0.44
					(9.96)	(-5.04)
adj. R^2	14.3%	10.2%	9.0%	10.3%	0.7%	16.7%
N	579.6	554.5	565.2	565.7	495.6	451.0
Panel B: Negative CFER subsample						
SIZE	0.15					0.05
	(15.49)					(9.72)
BAS		-1.57				-0.92
		(-4.08)				(-4.48)
IVOL			-1.10			-0.32
			(-12.79)			(-7.11)
Amihud				-20.68		-6.13
				(-6.97)		(-4.26)
RSI					-1.57	-0.31
					(-14.17)	(-4.63)
adj. R^2	14.8%	12.6%	12.5%	10.5%	2.8%	19.0%
N	718.5	672.7	694.9	695.2	595.0	528.5

Table 1.10. Robustness tests: (i) Comparison of methods to estimate CFER, (ii) Removing extreme CFER values

Entries in Panel A report the average return and Carhart (1997) four-factor model alpha of the spread portfolio of CFER-sorted value-weighted decile portfolios, where each column uses one of four estimation methods of CFER. The first row denotes the method of choosing strikes (AVE: taking average across available strikes, ATM: choosing the strike closest to the forward price) and the second row denotes the method of choosing maturities (CM: interpolating traded maturities to construct 30-day constant maturity CFER, CLS: choosing the traded maturity closest to 30 days). Entries in Panel B report the average return and Carhart (1997) four-factor model alpha of the spread portfolio of AVE-CM CFER-sorted value-weighted portfolios. The first column shows the result, where we truncate AVE-CM CFER values at a 1% level, that is, we remove CFER samples below 1st percentile point or above 99th percentile point. The second column reports the result of the modified spread, where we long the second highest CFER portfolio (portfolio 9) and short the second lowest CFER portfolio (portfolio 2). The third column reports the quintile portfolio sort results, and the last column reports the modified spread of the quintile portfolios, where we long the second highest CFER portfolio (portfolio 4) and short the second lowest CFER portfolio (portfolio 2). The estimation period spans January 1996 to April 2016 (244 months). t-statistics are adjusted for heteroscedasticity and autocorrelation (HAC-adjusted t-stat). The unit of the mean returns and alphas are % per month.

Panel A: Comparison between four estimation methods of CFER				
Strike	AVE		ATM	
Maturity	CM	CLS	CM	CLS
Value-weighted decile spread portfolio				
Average return	1.64	1.56	1.49	1.38
	(5.77)	(5.44)	(5.33)	(4.89)
α_{FFC}	1.86	1.77	1.60	1.51
	(6.56)	(6.20)	(5.55)	(5.33)
Equally-weighted decile spread portfolio				
Average return	1.73	1.56	1.67	1.54
	(9.10)	(8.92)	(9.15)	(8.95)
α_{FFC}	1.81	1.65	1.75	1.62
	(9.42)	(9.18)	(9.12)	(9.21)

Panel B: Mitigating the effect of extreme CFER values				
	Truncated decile sort (VW)		Quintile sort (VW)	
	Spread (10-1)	Spread (9-2)	Spread (5-1)	Spread (4-2)
Average return	1.43	0.92	1.11	0.43
	(6.08)	(4.19)	(5.59)	(3.37)
α_{FFC}	1.65	0.93	1.29	0.43
	(7.05)	(3.77)	(5.65)	(3.27)

Table 1.11. Bivariate dependent sort on CFER: Controlling for previous month return

Entries report the result of the bivariate dependent sort, where we first sort stocks based on the previous month return, $R_{t-1,t}$, and then within each group of the bid-ask spread level, we further sort stocks into quintile portfolios by the AVE-CM CFER criterion. Rows correspond to the level of the first sorting variable, the previous month return $R_{t-1,t}$, and the first to the fifth columns correspond to the level of the second sorting variable, AVE-CM CFER. The sixth to last columns report the average return, α_{FFC}, and the average CFER of the CFER-sorted spread portfolios, respectively. All returns are value-weighted returns. The estimation period spans January 1996 to April 2016 (244 months). t-statistics are adjusted for heteroscedasticity and autocorrelation (HAC-adjusted t-stat). The unit of the mean returns and alphas are % per month.

	Ave. returns of AVE-CM CFER-sorted portfolios					Ave. return	α_{FFC}	Ave. CFER
	1 (Lo)	2	3	4	5 (Hi)	Spread portfolio (5-1)		
$R_{t,-1,t}$ 1	-0.51	0.34	1.15	0.44	1.28	1.79	2.02	1.72
(lowest)	(-0.78)	(0.65)	(2.45)	(0.86)	(1.99)	(4.87)	(5.02)	(13.33)
$R_{t,-1,t}$ 2	0.50	0.76	1.04	1.04	1.30	0.80	0.90	1.24
	(1.10)	(2.23)	(2.79)	(2.98)	(3.51)	(2.83)	(2.83)	(13.69)
$R_{t,-1,t}$ 3	0.31	0.35	0.78	1.16	1.44	1.14	1.22	1.12
	(0.82)	(1.06)	(2.44)	(3.71)	(3.96)	(4.39)	(4.40)	(14.42)
$R_{t,-1,t}$ 4	0.58	0.51	0.81	0.76	0.97	0.39	0.52	1.11
	(1.53)	(1.68)	(2.59)	(2.33)	(2.75)	(1.44)	(1.83)	(15.30)
$R_{t,-1,t}$ 5	0.00	0.41	0.58	0.93	0.85	0.85	0.96	1.52
(highest)	(-0.00)	(0.93)	(1.51)	(2.21)	(2.04)	(2.88)	(3.15)	(18.81)

Table 1.12. Robustness tests: Non-synchronicity, Low stock price level, and NYSE breakpoint

Entries in Panel A report the average return and Carhart (1997) four-factor model alpha of the spread portfolio of AVE-CM CFER-sorted value-weighted decile portfolios, where the returns are calculated as the open-to-close return. The open-to-close return is the return from the open price on the first trading date after the portfolio formation in month-t to the close price of the end of month-$t+1$. Entries in Panel B report the average return and α_{FFC} of the AVE-CM CFER-sorted value-weighted decile portfolios, where we discard stocks whose price level is below \$10. Entries in Panel C report the the average return and α_{FFC} of the AVE-CM CFER-sorted value-weighted decile portfolios, where we calculate decile portfolios' breakpoints based on NYSE stocks only. The estimation period spans January 1996 to April 2016 (244 months). t-statistics are adjusted for heteroscedasticity and autocorrelation (HAC-adjusted t-stat). The unit of the mean returns and alphas are % per month.

	AVE-CM CFER-sorted value-weighted decile portfolios										Spread
	1 (Lowest)	2	3	4	5	6	7	8	9	10 (Highest)	10-1
Panel A: Open-to-close return (non-synchronicity)											
Ave. return	-0.21	0.22	0.52	0.48	0.77	0.87	0.97	0.90	1.13	1.39	1.60
	(-0.43)	(0.57)	(1.52)	(1.59)	(2.51)	(2.77)	(3.16)	(2.66)	(3.12)	(3.21)	(5.58)
α_{FFC}	-1.14	-0.63	-0.28	-0.25	-0.01	0.10	0.23	0.16	0.42	0.69	1.83
	(-6.73)	(-3.80)	(-2.42)	(-2.41)	(-0.14)	(0.99)	(2.27)	(1.30)	(2.40)	(3.20)	(6.40)
Panel B: Eliminating stocks whose price is below \$10											
Ave. return	-0.09	0.28	0.49	0.52	0.79	0.91	0.89	0.98	1.05	1.26	1.35
	(-0.18)	(0.79)	(1.49)	(1.75)	(2.62)	(2.82)	(2.99)	(2.95)	(3.00)	(3.23)	(5.79)
α_{FFC}	-1.03	-0.51	-0.29	-0.20	0.01	0.10	0.17	0.23	0.35	0.55	1.57
	(-5.07)	(-3.27)	(-2.92)	(-2.03)	(0.07)	(0.99)	(1.68)	(2.15)	(2.18)	(3.15)	(5.96)
N	122.0	122.2	122.1	122.2	121.8	122.3	122.1	122.2	122.1	122.1	—
Panel C: NYSE breakpoints											
Ave. return	-0.02	0.48	0.51	0.60	0.80	0.90	0.88	1.11	0.93	1.33	1.35
	(-0.04)	(1.38)	(1.51)	(2.01)	(2.58)	(2.79)	(2.96)	(3.48)	(2.66)	(3.26)	(5.33)
α_{FFC}	-0.95	-0.34	-0.27	-0.13	0.00	0.10	0.16	0.36	0.18	0.63	1.58
	(-4.92)	(-2.56)	(-2.28)	(-1.19)	(-0.04)	(0.99)	(1.45)	(3.26)	(1.45)	(3.25)	(5.85)
N	182.0	136.6	124.0	118.7	115.3	115.3	117.3	122.1	134.0	181.1	—

Table 1.13. Robustness test: Sub-sample analysis

Entries in Panels A and B report the average return and Carhart (1997) four-factor model alpha of the spread portfolio of AVE-CM CFER-sorted value-weighted decile portfolios over January 1996 to December 2006 and January 2007 to April 2016, respectively. t-statistics adjusted for heteroscedasticity and autocorrelation and reported in the parentheses. The unit of all variables is % per month.

	AVE-CM CFER-sorted value-weighted decile portfolios										Spread
	1 (Lowest)	2	3	4	5	6	7	8	9	10 (Highest)	10-1
Panel A: Sub-sample, January 1996–December 2006											
Average return	-0.19	0.30	0.57	0.49	0.89	0.99	1.13	0.82	1.42	1.76	1.95
	(-0.31)	(0.60)	(1.33)	(1.19)	(2.25)	(2.41)	(2.77)	(1.90)	(2.79)	(3.04)	(5.02)
α_{FFC}	-1.20	-0.71	-0.36	-0.33	-0.02	0.16	0.33	-0.02	0.74	1.20	2.41
	(-5.48)	(-2.96)	(-1.87)	(-2.18)	(-0.12)	(1.07)	(1.99)	(-0.12)	(2.54)	(3.84)	(6.38)
Panel B: Sub-sample, January 2007–April 2016											
Average return	-0.17	0.15	0.51	0.50	0.66	0.77	0.81	1.03	0.84	1.10	1.27
	(-0.21)	(0.24)	(0.94)	(1.07)	(1.38)	(1.57)	(1.72)	(1.97)	(1.59)	(1.69)	(3.24)
α_{FFC}	-0.87	-0.49	-0.16	-0.10	-0.01	0.09	0.14	0.34	0.18	0.35	1.22
	(-3.36)	(-2.03)	(-0.99)	(-0.78)	(-0.05)	(0.97)	(1.16)	(2.44)	(1.07)	(1.56)	(3.12)

Table 1.14. Predictive power of CFER: Fama-MacBeth regressions

Entries report the results from Fama and MacBeth (1973) regressions of stock returns on AVE-CM CFER, market beta, SIZE (log of market equity), log of book-to-market (logBM), Momentum ($R_{t-12,t-1}$), previous month return $R_{t-1,t}$, idiosyncratic volatility (IVOL), profitability (operational profit to book equity), investment (asset growth rate), Amihud's (2002) illiquidity measure, relative bid-ask spread (BAS), turnover rate (TO), and implied volatility spread (IVS). See Appendix 1.C for detailed definition of these variables. The time-series averages of the estimated coefficients of the cross-sectional regressions are reported. t-statistics are adjusted for heteroscedasticity and autocorrelation and reported in the parentheses. The time-series averages of adjusted R^2 and the observation number N of cross-sectional regressions are reported in the last two rows. Columns (1) to (4) report the results using all samples from January 1996 to April 2016. Columns (5) and (6) report the results using only NYSE/Amex and NASDAQ stocks, respectively. Columns(7) and (8) report the results using only the observations with non-negative and negative AVE-CM CFER, respectively. Columns (9) and (10) report the results using only the observation over 1996-2006 and 2007-2016, respectively.

	All sample				NYSE Amex	NAS DAQ	CFER ≥ 0	CFER < 0	1996–2006	2007–2016
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
CFER	0.62	0.39	0.40	0.30	0.53	0.31	0.46	0.46	0.41	0.38
	(7.78)	(5.60)	(5.60)	(2.93)	(4.12)	(3.23)	(2.78)	(4.01)	(4.14)	(3.70)
Beta		-0.03	-0.02	0.03	-0.02	-0.06	0.19	-0.01	0.19	-0.26
		(-0.12)	(-0.07)	(0.11)	(-0.08)	(-0.20)	(0.63)	(-0.03)	(0.48)	(-0.64)
SIZE		-0.11	-0.12	-0.13	-0.14	0.01	-0.06	-0.11	-0.12	-0.12
		(-1.92)	(-1.95)	(-1.88)	(-2.22)	(0.12)	(-0.88)	(-1.78)	(-1.23)	(-1.71)
logBM		0.11	0.11	0.13	0.11	0.15	0.14	0.08	0.32	-0.13
		(1.12)	(1.14)	(1.26)	(1.28)	(1.36)	(1.24)	(0.85)	(2.29)	(-1.19)
Mom		-0.02	-0.01	-0.06	0.09	-0.11	-0.05	-0.04	0.24	-0.31
		(-0.05)	(-0.03)	(-0.18)	(0.22)	(-0.38)	(-0.14)	(-0.10)	(0.77)	(-0.47)
$R_{t-1,t}$		-1.12	-1.26	-1.05	-0.53	-1.78	-1.55	-0.88	-2.50	0.21
		(-1.58)	(-1.81)	(-1.52)	(-0.64)	(-2.47)	(-1.91)	(-1.18)	(-2.91)	(0.20)
IVOL		-0.01	0.00	0.00	-0.02	0.01	0.00	0.00	0.00	-0.01
		(-1.47)	(-0.51)	(-0.65)	(-2.07)	(0.87)	(-0.27)	(-0.75)	(0.16)	(-1.10)
Profit		0.27	0.28	0.32	0.22	0.27	0.37	0.42	0.53	0.00
		(2.07)	(1.92)	(1.98)	(1.73)	(0.88)	(2.00)	(2.35)	(2.12)	(-0.00)
Invest		-0.33	-0.33	-0.32	-0.46	-0.25	-0.39	-0.28	-0.34	-0.32
		(-3.76)	(-3.64)	(-3.46)	(-2.98)	(-2.48)	(-2.64)	(-2.60)	(-3.24)	(-2.01)
Amihud			-14.62	-0.10	-2.51	0.92	-34.98	-17.24	-2.77	-28.59
			(-2.43)	(-0.53)	(-0.09)	(0.12)	(-2.06)	(-1.08)	(-0.45)	(-2.68)
BAS			0.44	-17.22	0.47	-0.66	0.52	0.70	0.10	0.84
			(0.62)	(-2.28)	(0.53)	(-0.66)	(0.44)	(0.71)	(0.16)	(0.62)
TO			-0.12	0.65	0.18	-0.26	-0.36	-0.08	-0.11	-0.14
			(-0.72)	(0.82)	(0.86)	(-0.94)	(-1.82)	(-0.48)	(-0.37)	(-1.27)
IVS				3.01						
				(1.71)						
Int.	0.88	2.84	2.92	2.97	3.32	1.09	2.17	2.72	3.05	2.76
	(2.17)	(3.03)	(2.83)	(2.69)	(3.00)	(0.52)	(1.85)	(2.56)	(1.84)	(2.29)
Adj R^2	0.2	9.0	9.2	9.8	10.7	7.6	10.1	9.9	10.4	7.8
N	1323	1009	940	782	549	402	430	514	776	1134

Chapter 2

Martingale Restriction, Risk-neutral Skewness, and Future Stock Returns

2.1 Introduction

A voluminous literature has documented that market frictions such as short-sale constraints, margin and transaction costs affect asset returns.[1] While this literature mainly examines the effect of frictions on asset returns under the physical probability measure, relatively little research has been done on the implications of the presence of market frictions under the risk-neutral probability measure and for the calculation of option-implied variables.

In this Chapter, we fill this gap by relating the *martingale restriction* (MR) to a formal theoretical setting which models market frictions. MR is a property of asset prices in the absence of market frictions first examined by Longstaff (1995); if the market is frictionless and arbitrage-free, asset prices discounted by the risk-free rate should be martingales under the risk-neutral measure (the first fundamental theorem

[1] Early studies by He and Modest (1995) and Luttmer (1996) argue that the equity risk premium puzzle can be explained by considering the effect of market frictions on the expected asset returns. Brunnermeier and Pedersen (2009), Gârleanu and Pedersen (2011), Chabakauri (2013), and Frazzini and Pedersen (2014) among others show that margins and leverage constraints affect asset returns. A number of studies document that short-sale constraints affect stock returns (e.g., Chen et al., 2002; Ofek et al., 2004; Asquith et al., 2005; Drechsler and Drechsler, 2014). Moreover, Hou et al. (2018) examine more than 100 market friction-related stock anomaly variables.

of asset pricing; Harrison and Kreps, 1979; Harrison and Pliska, 1981). The violation of MR suggests the presence of market frictions and it would call for a revisit of the voluminous theoretical and empirical literature which is based on the martingale assumption on asset prices.

Our study allows us to make three contributions. First, we revisit Longstaff (1995) and propose a novel and robust way to test MR. Moreover, we document that the MR testing way proposed by Longstaff (1995) may result in spurious judgment on whether MR is violated. Second, we explore the implications of the violation of MR for the estimation of risk-neutral moments (RNMs). RNMs are the building block in the vast literature which examines the informational content of option prices to address a number of questions in finance such as asset pricing, stock return predictability and asset allocation. The Bakshi et al. (2003) (BKM) formulae are the current de facto standard estimation method of RNMs. However, the BKM formulae are based on the assumption that the underlying asset satisfies MR; its violation may lead to biases in the estimated RNMs. We derive generalized formulae to estimate RNMs compatible with the possible violation of MR. Third, we document that the RNS estimated by the original BKM formula (O-RNS) predicts future stock returns, whereas the estimated RNS based on our generalized formula which accounts for the violation of MR (G-RNS) does not. More importantly, we find that the predictive power of O-RNS stems from its bias component caused by the violation of MR. These findings shed light on the ongoing debate on the mechanism behind the ability of O-RNS to predict future stock returns as we will discuss in the literature review Section below.

MR predicts that the risk-neutral expected asset return equals the risk-free rate. Our MR testing approach examines whether there is a wedge between the risk-neutral expected asset return and the risk-free rate. A non-zero wedge signifies that MR is violated due to the presence of market frictions. We show that this wedge term equals the CFER (contribution of frictions to expected returns) term of Hiraki and Skiadopoulos (2019) (HS).[2] By considering a general setting in which an agent faces transaction costs and constraints on her portfolio allocation choice caused by market frictions, HS show that the expected excess return of the asset under the physical probability measure equals the sum of CFER and the covariance between the in-

[2]Chapter One of this thesis is based on Hiraki and Skiadopoulos (2019).

tertemporal marginal rate of substitution (IMRS) and the asset return. We define the "risk-neutral" probability measure in the HS setting by using the (scaled) IMRS as a Radon-Nikodým derivative. Then, we show that the expected excess return under this risk-neutral measure equals CFER, that is, the CFER term is the wedge between the risk-neutral expected return and the risk-free rate. Therefore, we can test MR by testing whether CFER is zero or not. Our CFER-based approach provides a robust and easily implementable way to test MR; HS show that CFER can be reliably estimated in a model-free manner from a properly scaled deviations from put-call parity. On the other hand, we document that the *implied stock price approach* to test MR proposed by Longstaff (1995) may reject MR spuriously even if MR holds. In particular, we document that empirical results found in the implied stock price literature is not evidence for the violation of MR, but a mechanical reflection of a well-known stylized fact that the implied volatility (IV) of out-of-the money (OTM) equity put options are greater than those of at-the-money (ATM) options (negative IV skew).

Next, we show that the original BKM formulae do not correctly estimate RNMs if MR is violated (i.e., CFER is non-zero). This is because the derivation of the original BKM formulae relies on the implicit assumption that the risk-neutral expected asset return equals the risk-free rate, which does not hold under the violation of MR. The bias in the estimation of RNMs arises because a wrong risk-neutral expected return is employed to calculate the *central* moments (i.e., moments around the mean) of returns. We remedy this drawback by providing the generalized BKM formulae, in which we set the risk-neutral expected asset return to the sum of the risk-free rate and the CFER of the underlying asset.

In addition, we investigate the behavior of the Black and Scholes (1973) (BS) IV under the violation of MR because this is related to the estimation of RNS in two ways. First, IV is often used to estimate RNMs as an input to interpolate discretely observed option prices.[3] We document, by means of simulation, that using the standard BS-IV as an input to interpolate option prices results in incorrectly estimated RNS in the case where MR is violated. Second, the slope of IV curves (e.g., difference between OTM option's IV and ATM option's IV) is frequently interpreted as a proxy of RNS

[3]Since the theoretical formulae of BKM requires a continuum of option prices with respect to the strike price, typically discretely observed option prices are interpolated and extrapolated to obtain a continuum of option prices. To this end, BS-IVs of options rather than their option prices are typically interpolated and extrapolated.

(e.g., Bali et al., 2018). However, we show that in the presence of market frictions, a non-zero IV slope measure does not necessarily mean that the underlying distribution is not log-normal, in contrast to the common perception. To remedy these issues, we propose the *robust* IV, which accounts for the possible violation of MR in the calculation of IV. Our simulation result shows that using the robust IV as an input to interpolate option prices reduces the estimation bias in the RNS. Moreover, the robust IV curve is a more appropriate ingredient to calculate the slope of the IV curve because a zero slope of the robust IV curve means that the underlying distribution is log-normal.

We test whether the S&P 500 index and U.S. individual equities violate MR based on our new testing approach and find that both the index and individual stocks frequently violate MR. We find that the 30-day risk-neutral expected return of the S&P 500 index deviates from the risk-free rate on average by 1% to 2% per year. The degree of the violation of MR is generally larger for individual stocks. We find that the time-series average deviation of the 30-day risk-neutral expected return of individual stocks from the risk-free rate is greater than 1% per year for almost all stocks and greater than 3% for more than half of stocks. This finding validates our motivation to theoretically and empirically investigate the effect of MR on the estimated RNMs.

Next, we compare the estimated O-RNMs and G-RNMs. We find that the violation of MR has a substantial effect on the estimated RNS but not on the estimated model-free implied volatility and risk-neutral kurtosis. Importantly, we find that the difference between O-RNS and G-RNS is economically significant; we find that O-RNS predicts the cross-section of future returns in line with the literature (e.g., Stilger et al., 2017), whereas G-RNS does not. Furthermore, we empirically document that the predictive power of O-RNS stems from its estimation bias component caused by the violation of MR. The bias component is highly correlated with CFER in line with the fact that CFER measures the degree of the violation of MR. In a nutshell, O-RNS predicts future returns because its bias component contains predictive power inherited from the strong predictability of CFER documented in HS.

The remaining part of this Chapter is organized as follows. In Section 2.2, we review the related literature and discuss our contributions. In Section 2.3, we outlay the theoretical framework for testing MR. We also discuss the relation between our approach and the existing implied stock price approach proposed by Longstaff (1995). In Section 2.4, we propose the generalized BKM formulae for the estimation of RNMs and the robust IV, both of which account for the possible violation of MR. Section 2.5 explains data sources and empirical procedures for the estimation of CFER and RNMs as well as their summary statistics. Section 2.6 investigates the return predictive power of O-RNS and G-RNS. Section 2.7 concludes this Chapter.

2.2 Related literature and our contributions

We contribute to three strands of literature. The first has to do with studies which document that O-RNS predicts stock returns. While it has been well-documented that O-RNS has return predictive power, the sign of its predictive relation and the mechanism of the predictability are still open questions. Conrad et al. (2013) document that RNS *negatively* predicts future returns, consistent with theoretical models in which risk-averse agents accept lower returns in exchange for positive skewness (e.g., Harvey and Siddique, 2000; Mitton and Vorkink, 2007). On the other hand, a number of more recent studies document that RNS *positively* predict future returns, yet there is no unanimous consensus regarding the mechanism which generates positive return predictive power.

There are two proposed explanations for the positive predictive relation. The first one is a *limits-of-arbitrage* explanation: O-RNS signals the future abnormal stock performance due to the presence of market frictions. For example, Rehman and Vilkov (2012) find that stocks with more negative O-RNS underperform stocks with less negative O-RNS. They argue that this is because stocks with more negative O-RNS are relatively overpriced and this overpricing is not corrected fast due to limits-of-arbitrage. Stilger et al. (2017) document that it is the interplay between the stock overpricing and short-sale constraints that explains the result in Rehman and Vilkov (2012). Gkionis et al. (2018), in contrast to the previous literature which focuses on the negative informational content of O-RNS for future stock returns, document that stocks with relatively high O-RNS signal future stock outperformance. They argue that these stocks are underpriced and they have a greater downside risk which may discourage arbitrage activities in the underlying stock compared to stocks with more negative O-RNS values. Chordia et al. (2019) document that *"the positive RNS-return relation is more pronounced for stocks with higher limits-to-arbitrage and illiquidity"* (p.28), although they argue that the limits-of-arbitrage story cannot fully explain the predictive power of O-RNS.

The second explanation relies on the idea that the option market attracts informed traders (e.g., Easley et al., 1998); informed traders resort to the option market to utilize their private information on the future stock performance and consequently private information is incorporated into option prices and in option-implied measures (e.g., RNS) prior to the stock price. For example, pessimistic informed traders may utilize their private information through the option market by buying OTM put options and selling OTM call options. According to the demand-based option pricing theory (e.g., Bollen and Whaley, 2004; Gârleanu et al., 2009), this trading pressure would increase (decrease) put (call) option prices, resulting in a lower RNS. Chordia et al. (2019) argue that this informed option trading channel drives the predictive power of O-RNS based on their finding that O-RNS observed prior to scheduled and non-scheduled corporate news strongly predicts post-news stock returns. Bali et al. (2018) decompose O-RNS into a systematic part and an unsystematic part and find that the unsystematic component is robustly related to expected stock returns. They argue that the unsystematic RNS component represents firm-specific information which informed investors can utilize.

We shed light on the ongoing debate on the O-RNS predictive power in three important ways. First, our findings support the limits-of-arbitrage explanation as the mechanism behind the predictive power of O-RNS. We theoretically and empirically show that the presence of market frictions (i.e., the violation of MR) results in estimation biases in O-RNS. Furthermore, we document that it is the bias component of O-RNS caused by the violation of MR that predicts future stock returns; we find that the bias component is highly correlated with CFER in line with the fact that CFER measures the degree of the violation of MR (i.e., the deviation of the risk-neutral expected return from the risk-free rate). Since CFER is the effect of frictions on the expected stock returns, O-RNS signals the effect of frictions on the expected stock returns via its correlation with CFER.

Second, our CFER-based story can explain both the positive (higher O-RNS predicting future outperformance) and negative (lower O-RNS predicting future underperformance) informational contents simultaneously. This is in contrast to the existing literature which finds it challenging to explain the positive and negative informational contents in a single type of market frictions. For example, the explanation based on the short-sale constraints employed by Stilger et al. (2017) can explain only the negative informational contents of O-RNS, whereas Gkionis et al.'s (2018) explanation based on the downside risk can explain only the positive informational contents of O-RNS.

Third, we find that the predictive power of O-RNS is more pertaining to the market frictions than the informed option trading explanation. We compare the predictive power of O-RNS in the two subsamples: stocks with non-zero option trading volume and those without any option trading. Then, we find that the predictive power of O-RNS does not differ qualitatively between these two subsamples. This finding is at odds with the informed option trading story, which predicts that the predictive power of option-implied variables is more pronounced when option trading volume is larger. On the other hand, this finding supports the limits-of-arbitrage explanation for the predictive power of RNS, because stocks with zero option trading tend to be smaller stocks which face a considerable degree of market frictions. This finding is largely in line with Goncalves-Pinto et al. (2019), who show that their option-based return predictor, DOTS, does not exhibit the different level of predictability among stocks with zero and non-zero option trading volume.

In addition to the literature on the predictive power of RNS, this Chapter contributes to the two strands of literature. First, we contribute to the literature on testing MR. Longstaff (1995) tests MR for the S&P 100 index and finds that it violates MR. Follow-up studies examine MR for other major equity indices including the S&P 500 (Strong and Xu, 1999), the German DAX index (Neumann and Schlag, 1996; Huang et al., 2016), and the South Korean KOSPI 200 index (Guo et al., 2013). All these studies test MR by examining the difference between the market stock price and the implied stock price extracted from an option pricing model (typically the Black and Scholes, 1973 model). We contribute to this literature in two ways. First, we demonstrate that the implied stock price approach may result in spurious rejection of MR even if MR holds. For example, we show that the BS-model-based implied stock price approach does not constitute a valid MR test in the case where the true risk-neutral underlying distribution is not log-normal. Second, our CFER-based MR test approach is free from the drawback of the implied stock price approach. This is because the estimation of CFER is model-free; the estimation of CFER and hence our CFER-based MR test approach do not depend on a specific option pricing model.

We also contribute to the literature which assesses biases in empirically estimated option-based measures. Bliss and Panigirtzoglou (2002) and Hentchel (2003) examine the effect of measurement errors in option prices on the estimation of risk-neutral distributions and IVs, respectively. Dennis and Mayhew (2009) investigate the effect of measurement errors in option prices and discretely observed strikes on the estimation of RNMs based on the BKM formulae. Ammann and Feser (2019) extend Dennis and Mayhew (2009) and investigate robust interpolation methodologies to calculate RNMs in the presence of measurement errors and limited number of option price observations. These studies assume that the underlying satisfies MR and focus on the effect of practical issues such as discretely traded strikes and measurement errors to the calculation of option-based quantities. Our study differs from these studies in that the bias in RNMs caused by the violation of MR occurs even under an ideal situation where a continuum of option prices are observable without any measurement errors.

Our analysis in this Chapter focuses on a specific issue of the skewness of stocks, namely, the return predictive power of RNS estimated based on the BKM formulae. However, it is worth noting here that the literature on the skewness of asset returns has much broader coverage. For instance, there are theoretical and empirical studies which show that both the systematic skewness and idiosyncratic skewness of securities measured under the physical measure are relevant to their returns. Kraus and Litzenberger (1976) propose a three-moment CAPM in which the systematic skewness affects expected returns and Harvey and Siddique (2000) find that the systemic skewness risk is priced. Mitton and Vorkink (2007) find that unsystematic skewness affect equilibrium asset prices. Barberis and Huang (2008) find a negative relation between individual stock skewness and expected return, while Boyer et al. (2010) find that the expected idiosyncratic skewness predicts cross-section of stock returns negatively. This strand of studies is closely related to the literature on lottery-like stocks, which argues that investors have preference on stocks with positively skewed payoff distribution and hence these stocks command lower returns. Kumar (2009), Bali et al. (2011), and Filippou et al. (2018) among others provide empirical support for this story.

Another important research topic of the skewness is the skew risk premium. A seminal work by Kozhan et al. (2013) proposes an option-based skew swap that pays the difference between the implied skew and realized skew. They find that the skew risk premium of the S&P 500 index exists and positive, that is, on average the implied skew is lower than the realized skew. Ruf (2012) studies the skew risk premium of commodity futures options and find that the skew risk premium increases when the degree of limits-of-arbitrage increases. Harris and Qiao (2018) find that individual stocks' skew risk premium is positive and the skew risk premium is negatively related to future stock returns. Lin et al. (2019) develop an equilibrium asset and option pricing model, which yields an analytical expression for the relation between the market equity premium and the skew risk premium.

2.3 Theoretical framework

2.3.1 The martingale restriction: Definition and implications

We consider a financial market where three types of assets trade: the risk-free bond, a risky asset (the stock) and options written on the risky asset. We assume that the instantaneous risk-free rate is constant and given by r_f. The gross risk-free bond return from time t to T is denoted by $R^f_{t,T}$ where $R^f_{t,T} = e^{r_f(T-t)}$. We denote the stock price at time t by S_t. The stock pays dividends and $\widetilde{D}_{t,T}$ denotes the time T value of the aggregate dividends paid over the period $(t, T]$, that is,

$$\widetilde{D}_{t,T} = \sum_{t<j\leq T} R^f_{j,T} D_j. \tag{2.1}$$

We define the cum-dividend gross stock return from t to T by $R_{t,T} = (S_T + \widetilde{D}_{t,T})/S_t$. The time-$t$ price of a call (put) option with strike price K and maturity T is denoted by $C_t(K,T)$ ($P_t(K,T)$) and $\tau = T - t$ denotes the length of return horizon and options' time-to-maturity. We call the discounted value of the asset with the dividends reinvested (if applicable) the *discounted (cum-dividend) price process*. For example, the discounted cum-dividend stock price process is defined as $\{(S_t + \widetilde{D}_{0,t})/R^f_{0,t}\}_{t\geq 0}$. For the discounted price process of non-dividend paying assets such as options, we simply call it, for example, the discounted call price process $\{C_t(K,T)/R^f_{0,t}\}_{t\geq 0}$.

The first fundamental theorem of asset pricing (FFTAP, Harrison and Kreps, 1979; Harrison and Pliska, 1981) states that in a frictionless market, the no-arbitrage condition is equivalent to the existence of a *risk-neutral probability measure* $\mathbb{Q}$, under which the discounted (cum-dividend) price process of any asset is a martingale. In other words, FFTAP yields the following *martingale restriction* on the dynamics of the traded asset prices as a testable implication.

Definition 2.3.1 (Martingale restriction)**.** *The martingale restriction (MR) is said to be satisfied if the discounted (cum-dividend) price process of any traded asset is a martingale under a risk-neutral probability measure.*

MR yields a number of important implications regarding the price and the expected return of any traded asset. First, under MR, the martingale property of the discounted cum-dividend stock price yields

$$\frac{S_t + \widetilde{D}_{0,t}}{R^f_{0,t}} = \mathbb{E}^{\mathbb{Q}}_t\left[\frac{S_T + \widetilde{D}_{0,T}}{R^f_{0,T}}\right] \quad \Leftrightarrow \quad S_t = \frac{1}{R^f_{t,T}}\mathbb{E}^{\mathbb{Q}}_t[S_T + \widetilde{D}_{t,T}], \tag{2.2}$$

where $\mathbb{E}^{\mathbb{Q}}_t[.]$ denotes the expectation under the $\mathbb{Q}$-measure conditional on the information set at time t. Equation (2.2) is the so-called risk-neutral valuation formula, that is, the current stock price equals the discounted value of the expected future cum-dividend stock price under $\mathbb{Q}$. Second, equation (2.2) implies that the expected stock return $R_{t,T}$ satisfies

$$\mathbb{E}^{\mathbb{Q}}_t[R_{t,T}] = R^f_{t,T}, \tag{2.3}$$

that is, the $\mathbb{Q}$-expected stock return equals the risk-free rate.

Third, in the asset pricing literature, FFTAP is stated under the physical measure $\mathbb{P}$ (see e.g., Aït-Sahalia and Lo, 1998; Cochrane, 2005). For each risk-neutral measure, a corresponding stochastic discount factor (SDF) is defined as $m_{t,T} = e^{-r_f\tau}(d\mathbb{Q}/d\mathbb{P})$, where $d\mathbb{Q}/d\mathbb{P}$ is the Radon-Nikodým derivative for the change of measure from $\mathbb{P}$ to $\mathbb{Q}$. Then, by definition, $\mathbb{Q}$ and $m_{t,T}$ are related through the following change of measure equation,

$$\mathbb{E}_t^{\mathbb{P}}[m_{t,T}(S_T + \widetilde{D}_{t,T})] = \frac{1}{R_{t,T}^f}\mathbb{E}_t^{\mathbb{Q}}[S_T + \widetilde{D}_{t,T}]. \tag{2.4}$$

Equations (2.2) and (2.4) yield

$$1 = \mathbb{E}_t^{\mathbb{P}}[m_{t,T}R_{t,T}]. \tag{2.5}$$

Moreover, equation (2.5) yields the following asset pricing equation:

$$\mathbb{E}_t^{\mathbb{P}}[R_{t,T}] - R_{t,T}^f = -R_{t,T}^f Cov_t^{\mathbb{P}}(m_{t,T}, R_{t,T}), \tag{2.6}$$

that is, the expected excess return is determined by the covariance between the SDF and the asset's return.

Testing MR is of paramount importance because the above argument shows that the important asset pricing relations including equations (2.2) to (2.6) hold if and only if MR holds. Specifically, we will document that the violation of equation (2.3) results in biases in the estimated RNMs based on BKM formulae.

2.3.2 Testing MR via CFER

Our approach to test MR relies on the idea that the violation of MR implies the presence of market frictions. Equivalently stated, a violation of equation (2.3) implies the presence of a wedge between the risk-neutral expected stock return $\mathbb{E}_t^{\mathbb{Q}}[R_{t,T}]$ and the risk-free rate $R_{t,T}^f$, which captures the effect of market frictions. We quantify this wedge term by using the Hiraki and Skiadopoulos (2019) (HS) general asset pricing setting which accounts for market frictions.

The HS model is an agent-based model, where a marginal agent chooses her consumption c_t and asset allocations $\boldsymbol{\theta_t}$ (a vector of allocations to each asset) over the stock, the risk-free bond and one-period options to maximize her expected lifetime utility $\sum_{j\geq t}\mathbb{E}_t^{\mathbb{P}}[u(c_j)]$.[4] This maximization problem is subject to the following dy-

[4]The time-separable utility form is employed for the sake of simplicity. The HS setting works under more general utility functions including the recursive utility.

namics of the agent's wealth W_t,

$$c_t = W_t - \theta_t^0 - \theta_t^S S_t - \sum_K [\theta_t^c(K)C_t(K) + \theta_t^p(K)P_t(K)] - TC_t(\boldsymbol{\theta}_t - \boldsymbol{\theta}_{t-1}),$$

$$W_{t+1} = \theta_t^0 R_{t,t+1}^f + \theta_t^S(S_{t+1} + D_{t+1}) + \sum_K \left[\theta_t^c(K)(S_{t+1} - K)^+ + \theta_t^p(K)(K - S_{t+1})^+\right],$$

where TC_t is the transaction costs function and $(x)^+ = \max(x, 0)$. The asset allocation is also subject to L constraints,

$$g_t^l(\boldsymbol{\theta}_t) \geq 0, \quad l = 1\ldots, L. \tag{2.7}$$

With these constraints, the Bellman equation of the optimization problem is given by

$$V_t(W_t, \theta_{t-1}^S) = \max_{c_t, \boldsymbol{\theta}_t} \left\{u(c_t) + \beta \mathbb{E}_t^{\mathbb{P}}[V_{t+1}(W_{t+1}, \theta_t^S)]\right\}.$$

The allocation on the stocks at time $t-1$ is treated as a state variable because it affects the transaction costs and hence the agent's decision making. On the other hand, since we assume that options and the risk-free bond are one-period asset, their past allocations do not enter as state variables. The first-order condition of the agent regarding the allocation on the stock yields

$$S_t = \mathbb{E}_t^{\mathbb{P}}[m_{t,t+1}^*(S_{t+1} + D_{t+1})] + \sum_{l=1}^{L} \frac{\lambda_t^l}{u'(c_t)} \frac{\partial g_t^l}{\partial \theta_t^S} - \frac{\partial TC_t}{\partial \theta_t^S} + \frac{\beta}{u'(c_t)} \mathbb{E}_t^{\mathbb{P}}\left[\frac{\partial V_{t+1}}{\partial \theta_t^S}\right], \tag{2.8}$$

where $m_{t,t+1}^* = \beta(\partial V_{t+1}/\partial W_{t+1})/u'(c_t)$ is the agent's intertemporal marginal rate of substitution (IMRS) between time t and $t+1$, and λ_t^l is the Lagrange multiplier of the constraint, equation (2.7). By transforming equation (2.8), HS derive the following asset pricing equation,

$$\mathbb{E}_t^{\mathbb{P}}[R_{t,t+1}] - R_{t,t+1}^f = -R_{t,t+1}^f Cov_t^{\mathbb{P}}(m_{t,t+1}^*, R_{t,t+1}) + CFER_{t,t+1}, \tag{2.9}$$

where

$$CFER_{t,t+1} = -\frac{R_{t,t+1}^f}{S_t}\left(\sum_{l=1}^{L} \frac{\lambda_t^l}{u'(c_t)} \frac{\partial g_t^l}{\partial \theta_t^S} - \frac{\partial TC_t}{\partial \theta_t^S} + \frac{\beta}{u'(c_t)} \mathbb{E}_t^{\mathbb{P}}\left[\frac{\partial V_{t+1}}{\partial \theta_t^S}\right]\right), \tag{2.10}$$

is the *contribution of frictions to the expected return* of the stock, which represents the wedge between the expected stock return and its covariance risk premium term caused by constraints on the portfolio allocations (e.g., margin constraints) and transaction costs.

Under this framework, we define the new probability measure $\mathbb{Q}^*$ by using the scaled IMRS $R^f_{t,T} m^*_{t,T}$ as a Radon-Nikodým derivative via the change of measure formula (equation (2.4) with $m_{t,T}$ being replaced with $m^*_{t,T}$). In other words, we define the $\mathbb{Q}^*$-measure by utilizing the agent-based equilibrium model structure instead of resorting to FFTAP to claim the existence of such a measure; the latter route is not available because the market is not frictionless. Under this definition of the $\mathbb{Q}^*$-measure, we obtain the following result.

Lemma 2.3.1. *Under the HS model, the following relation holds.*

$$\mathbb{E}_t^{\mathbb{Q}^*}[R_{t,t+1}] = R^f_{t,t+1} + CFER_{t,t+1}. \tag{2.11}$$

Proof. See Appendix 2.A.1. □

Equation (2.11) implies that we can still call the $\mathbb{Q}^*$-measure the "risk-neutral measure." This is because equation (2.9) collapses to (2.11) if the agent has a risk-neutral preference in the sense that the IMRS $m^*_{t,t+1}$ is constant.

Lemma 2.3.1 allows us to draw three remarks. First, the comparison of equations (2.3) and (2.11) shows that the asset violates MR if and only if CFER is non-zero. Second, Lemma 2.3.1 shows that the $\mathbb{Q}^*$-expected return of the risky asset is not observable and it may vary across different assets given that CFER is not observable and it is asset specific. This invalidates a fundamental principle which prevails in the case of no frictions: the expected return of a risky asset under the risk-neutral measure equals the observable risk-free rate and hence it is not an unknown parameter. Third, from an empirical point of view, equation (2.11) shows that one can test whether MR holds by testing whether CFER is zero. This test can be implemented because HS show that CFER can be reliably estimated as a scaled deviation from put-call parity.

2.3.3 Relation to the implied stock price approach

Longstaff (1995) studies first whether a stock price satisfies MR empirically by testing whether equation (2.2) holds. To this end, he assumes that the unobservable right-hand side of equation (2.2) can be estimated by the *implied stock price* S^*_t extracted from the observed market option prices, that is, he assumes that

$$S_t^* = e^{-r_f \tau} \mathbb{E}_t^{\mathbb{Q}^*}[S_T + \widetilde{D}_{t,T}]. \tag{2.12}$$

Under this assumption, Longstaff (1995) and subsequent studies typically test MR by investigating whether the relative difference between the implied stock price and the observed stock price

$$\Delta_t = \frac{S_t^* - S_t}{S_t} \tag{2.13}$$

is different from zero.

To extract the implied stock price, we proceed as follows. First, an option pricing model $h(\xi, S_t, K, \tau, r_f, \widetilde{D}_{t,T})$, where ξ denotes the set of parameter(s), is selected. For example, a typical choice in the literature is the Black and Scholes (1973) (BS) model. In this case, ξ consists of one parameter, the volatility. Then, S_t^* is calculated by minimizing the sum of squared errors between the observed option prices $O_t(K,T)$ and model-based option prices $h(\xi, S_t^*, K, \tau, r_f, \widetilde{D}_{t,T})$ by treating the current stock price argument as a parameter to be estimated in addition to the model parameters in ξ:

$$\min_{\xi, S_t^*} \sum \left[O_t(K,T) - h(\xi, S_t^*, K, \tau, r_f, \widetilde{D}_{t,T}) \right]^2. \tag{2.14}$$

At the first glance, the implied stock price and CFER are two distinct variables because the former is about the *price* of the stock whereas the latter is about the *return* of the stock. However, dividing both sides of equation (2.12) by S_t and using equation (2.11) yield

$$\Delta_t = \frac{CFER_{t,T}}{R_{t,T}^f}. \tag{2.15}$$

Therefore, as long as equation (2.12) holds, we can view the implied stock price approach as an alternative way to estimate CFER. In this case, the implied stock price approach and the CFER approach would be equivalent ways to test MR because $\Delta_t = 0$ is equivalent to $CFER_{t,T} = 0$. The correspondence between the implied stock price and the expected stock return (or CFER) is not surprising given the one-to-one correspondence between the discounted present value of assets and the discount rate; the expected future stock price discounted by the risk-free rate (i.e., the implied stock price in equation (2.12)) coincides with the observed market price if and only if the expected stock return (i.e., the discount rate for the stock) equals the risk-free rate.

However, in the case where equation (2.12) does not hold, the two approaches are not equivalent; adding and subtracting $e^{-r_f\tau}\mathbb{E}_t^{\mathbb{Q}^*}[S_T + \widetilde{D}_{t,T}]$ to the numerator of equation (2.13) yields

$$\Delta_t = \frac{S_t^* - e^{-r_f\tau}\mathbb{E}_t^{\mathbb{Q}^*}[S_T + \widetilde{D}_{t,T}]}{S_t} + \frac{CFER_{t,T}}{R_{t,T}^f}, \tag{2.16}$$

where the first term in the right-hand side does not vanish unless equation (2.12) holds. Most importantly, in this case, the implied stock price approach does not constitute a valid test of MR because Δ_t will differ from zero even if CFER is zero (i.e., even if MR holds). Interestingly, the literature has largely overlooked investigating the validity of the crucial assumption of the implied stock price approach, equation (2.12), and hence the validity of the implied stock price approach as a way of testing MR.

We explore the extent to which the implied stock price approach constitutes a valid test of MR. We consider the most frequently considered situation in the literature, where the option price model is set to the BS model. The following Proposition showcases a situation where the implied stock price approach based on the BS option pricing model constitutes a valid test of MR.

Proposition 2.3.1. *Assume that options satisfy MR but the underlying stock may violate MR. Assume further that (i) the true $\mathbb{Q}^*$-distribution of S_T is log-normal, and (ii) there are at least two observed option prices input to the minimization problem described in equation (2.14). Then,*

$$\Delta_t^{BS} := \frac{S_{t,BS}^* - S_t}{S_t} = \frac{CFER_{t,T}}{R_{t,T}^f}, \tag{2.17}$$

where $S_{t,BS}^*$ is given as the solution of equation (2.14) in the case where the option pricing model used to extract the implied stock price is the BS model.

Proof. See Appendix 2.A.2. □

Proposition 2.3.1 shows that the implied stock price approach using the BS model constitutes a valid test of MR (i.e., MR holds if and only if $\Delta_t^{BS} = 0$) *if* the true $\mathbb{Q}^*$-distribution follows a log-normal distribution. On the other hand, we show, by means of simulation where we assume the stock price satisfies MR, that the BS model-based implied stock price approach spuriously rejects MR if the true $\mathbb{Q}^*$-distribution of the future stock price is not log-normal.

The simulation setup is as follows. We choose $S_t = 100$, $\tau = 1/6$ and $r_f = 4\%$. For dividends, we set $\widetilde{D}_{t,T} = 0.5$, which corresponds to an about 3% annualized dividend yield. Call and put options trade at strikes $K = 80, 85, \ldots, 120$ (i.e., nine strikes). We assume that both the underlying stock and options satisfy MR under the true $\mathbb{Q}^*$ measure. In this case, option prices are given by the risk-neutral valuation formula

$$C_t(K,T) = e^{-r_f \tau}\mathbb{E}_t^{\mathbb{Q}^*}[(S_T - K)^+], \quad \text{and} \quad P_t(K,T) = e^{-r_f \tau}\mathbb{E}_t^{\mathbb{Q}^*}[(K - S_T)^+], \tag{2.18}$$

where $(x)^+ = \max(x, 0)$. We assume that the $\mathbb{Q}^*$-distribution of the stock price may not be log-normal and hence the BS implied volatility (BS-IV) curve may be *skewed*, that is, the graph of the BS-IV as a function of the option moneyness (the strike-to-stock price ratio) may not be flat. We model the BS-IV curve as a linear function of moneyness $IV(K/S_t) = \sigma_{ATM} + k(K/S_t - 1)$. We set the at-the-money (ATM) volatility, σ_{ATM} to 20%. The coefficient k controls the slope of the IV skew. We examine five realistic values for the slope coefficient k: $k = \pm 1/2$, $k = \pm 1/4$ and $k = 0$. For example, $k = -1/2$ corresponds to the situation where the BS-IV at $K = 80$ (i.e., moneyness = 0.8) is 30% and that at $K = 120$ (i.e., moneyness = 1.2) is 10%. Then, we convert the given BS-IV at each traded strike to the respective call and put option prices by using the BS option pricing functions with deterministic dividend payments, that is,

$$C_t(K,T) = BS_{call}(S_t, K, \tau, r_f, \widetilde{D}_{t,T}, IV) = (S_t - e^{-r_f \tau}\widetilde{D}_{t,T})\Phi(d_1) - e^{-r_f \tau} K \Phi(d_2), \tag{2.19}$$

where Φ is the cumulative density function of the standard normal distribution and

$$d_1 = \frac{\log\left(\frac{S_t - e^{-r_f \tau}\widetilde{D}_{t,T}}{K}\right) + (r_f + \frac{1}{2}IV^2)\tau}{IV\sqrt{\tau}}, \quad \text{and} \quad d_2 = d_1 - IV\sqrt{\tau}. \tag{2.20}$$

The put option price is given by

$$P_t(K,T) = BS_{put}(S_t, K, \tau, r_f, \widetilde{D}_{t,T}, IV) = e^{-r_f \tau} K \Phi(-d_2) - (S_t - e^{-r_f \tau}\widetilde{D}_{t,T})\Phi(-d_1). \tag{2.21}$$

This construction ensures that the underlying satisfies MR, too. To see this,

$C_t(K) - P_t(K) = S_t - e^{-r_f\tau}\widetilde{D}_{t,T} - e^{-r_f\tau}K$ follows from the definition of BS-IV via the BS functions (equations (2.19) and (2.21)), while $C_t(K) - P_t(K) = e^{-r_f\tau}\mathbb{E}_t^{\mathbb{Q}^*}[S_T - K]$ follows from equation (2.18). These two equations yield $S_t = e^{-r_f\tau}\mathbb{E}_t^{\mathbb{Q}^*}[S_T + \widetilde{D}_{t,T}]$.[5] Note that converting the *same* BS-IV to obtain the call and put option prices is crucial to ensure that the underlying satisfies MR; in Section 2.4.4, we will show that the violation of MR is equivalent to the non-zero implied volatility spread (i.e., the difference between the call BS-IV and put BS-IV at the same strike). Given these generated option prices, the two parameters, the implied stock price S_t^* and the BS volatility parameter σ^*, are obtained by solving equation (2.14).

Table 2.1 reports the simulation result. In line with the literature on testing MR, we estimate S_t^* and σ^* using a set of call options and a set of put options, separately. We also report the estimate of (annualized) CFER as $\widehat{CFER} = R_{t,T}^f\widehat{\Delta}_t/\tau$ under the implied stock price approach (equation (2.15)). Here, we view the implied stock price approach as an alternative way to estimate the underlying stock CFER by imposing its key assumption, equation (2.12). We can see that the two parameters, S_t^* and σ^*, are correctly estimated under $k = 0$, that is, the implied stock price approach identifies correctly that MR holds ad hence it constitutes a valid test of MR in this case. This is in line with Proposition 2.3.1 because $k = 0$ corresponds to a situation where the true $\mathbb{Q}^*$-distribution follows a log-normal distribution. On the contrary, when k is negative (positive), S_t^* is estimated to be lower (higher) than S_t regardless of using either call options or put options. Therefore, even though we assume that the underlying satisfies MR, the implied stock price approach spuriously rejects MR when the slope of the IV curve is non-zero.

[Table 2.1 about here.]

We explain the mechanism which yields $S_t^* > S_t$ in the case where $k < 0$ (i.e., negative IV skew case). The implied stock price S_t^* is chosen so that the BS-IVs of the *model-based* option prices $h(\xi, S_t^*, K, \tau, r_f, \widetilde{D}_{t,T})$ fit the negatively skewed BS-IVs of the *observed* option prices as close as possible. For a set of call options, for this to occur, in-the-money (ITM) calls should be more "expensive" (in terms of the BS-IV) compared to out-of-the-money (OTM) calls. To achieve this, the implied stock

[5]Note that we make no assumption that the BS model is the true model, but we simply use it as a translation mechanism to map IVs to option prices. In fact, the designed simulation admits that the BS model is not the true model unless $k = 0$.

price approach sets S_t^* higher than S_t. This is equivalent to setting a positive $\widehat{CFER}$ (equation (2.15)) and hence setting the estimate of the expected stock return under $\mathbb{Q}^*$ greater than the risk-free rate (equation (2.11)). Therefore, choosing $S_t^* > S_t$ makes call options more expensive by setting the expected stock return higher. In addition, this effect is stronger for ITM call options than OTM call options. This is because one can benefit from a higher expected stock price only in the case where call options expire in-the-money, which is more likely for ITM options than OTM options. Figure 2.1a shows the effect of choosing a higher value for S_t^*; compared to the "initial" BS model-based IV curve where the parameters are set to $S_t^* = S_t$ and $\sigma^* = \sigma_{ATM}$ (the gray thin line), the BS model-based IV with a higher S_t^* (the blue dotted line) locates closer to the true simulated IV curve (the black thick line). Finally, σ^* is estimated lower than the average IV level (i.e., σ_{ATM}) to offset increases in call option prices due to a higher S_t^* which would prevent fitting the level of the given IV curve. Figure 2.1b shows the effect of choosing a lower value for σ^*; the BS model-based IV curve with lower σ^* (the red dashed line) locates closer to the true simulated IV line compared to the BS model-based IV curve with $\sigma^* = \sigma_{ATM}$ (the blue dotted line). For the case of put options, a negative IV skew means that OTM puts are more "expensive" (in terms of the IV) compared to OTM puts. To create such a pattern, again a higher S_t^* is chosen. This is because choosing $S_t^* > S_t$ makes put options cheaper and this effect is stronger for ITM put options than OTM ones (Figure 2.1c) for an analogous reason to the call case. Finally, σ^* is estimated higher than σ_{ATM} to offset decreases in put option prices due to a higher S_t^* which again would prevent fitting the level of the IV curve (Figure 2.1d).

[Figure 2.1 about here.]

Our simulation results are in line with the empirical results reported in the literature; Longstaff (1995) and the follow-up studies report that the estimated stock price is greater than the current stock price (i.e., $S_t^* > S_t$) in more than 90% of the cases in their samples. Although the previous literature views this as evidence for the violation of MR, our simulation results suggest that the findings in the previous literature is a mechanical reflection of the empirical stylized fact that index options exhibit negative IV skews.

A remark is in order at this point. Our simulation exercise shows that the implied stock approach rejects MR spuriously in the case where the BS model is chosen to calculate the implied stock price and an IV skew exists. In other words, even if CFER equals zero, Δ_t^{BS} may not equal zero, which implies that the key assumption of the implied stock price approach, equation (2.12), may not supported by the data. In addition, the spurious rejection of MR under the implied stock price approach may also be a result of the selected (possibly mis-specified) option pricing model. This manifests that the implied stock price approach is subject to the joint hypothesis problem, that is, testing a property of the true risk-neutral distribution (i.e., MR) by assuming a specific model for the risk-neutral distribution.

On the contrary, testing MR under the CFER-approach circumvents the joint hypothesis problem as well as the assumption shown by equation (2.12); our estimated CFER equals scaled deviations from put-call parity and hence its calculation does not require an option pricing model nor equation (2.12). Note, however, that there is the following caveat about our CFER-based approach for testing MR. Deviations from put-call parity means that either the underlying stocks or synthetic stocks (i.e., options) or both violate MR due to the effect of market frictions. Therefore, our CFER-based MR testing approach would result in a spurious judgment when the underlying stock does not violate MR but deviations from put-call parity occurs due to the violation of MR by options. Empirical findings in HS suggest that such a situation is not likely to occur. However, it should be noted that their conclusion relies on specific models of the effect of frictions on option prices (Frazzini and Pedersen (2012) and Hitzemann et al. (2017)), and hence their judgment is model dependent.

2.4 Estimation of risk-neutral moments: Theory

In this Section, we develop the theory for the estimation of RNMs under the possible violation of MR. In Section 2.4.1, we provide the *generalized* Bakshi et al. (2003) (BKM) formulae for the estimation of RNMs in the case where the underlying asset violates MR.

Then, in Section 2.4.2, we investigate the behavior of BS-IV under the violation of MR. This is of importance for the purposes of studying the effect of violations of MR to RNMs for two reasons. First, IV is an essential ingredient to obtain a continuum of option prices from discretely observed option prices and hence to estimate RNMs; our simulation result in Section 2.4.3 shows that the usage of the original BKM formulae and the standard IV results in biases in the estimated RNS in the case where MR is violated. Second, the slope of IV curves (e.g., the difference between OTM option's IV and ATM option's IV) is frequently interpreted as a proxy of RNS (e.g., Bali et al., 2018). However, in Section 2.4.4, we show that a non-zero IV slope measure does not necessarily mean the underlying distribution is not log-normal in the case where MR is violated. This is in contrast to the common perception that the presence of IV skew indicates departure from log-normality. To remedy these issues, we propose the *robust* IV, which accounts for the possible violation of MR in the calculation of IV.

2.4.1 The generalized Bakshi et al. (2003) formula

To derive the original BKM formulae, both the underlying and options are assumed to satisfy MR. We relax the assumption that the underlying satisfies MR, whereas we keep the MR assumption for the option prices. Admittedly, the assumption that option prices satisfy MR is strong because it implies that option prices are not affected by market frictions, in contrast to the previous literature (e.g., Frazzini and Pedersen, 2012; Hitzemann et al., 2017) which documents that market frictions affect option prices. Note that we do not need this assumption for the estimation of CFER and hence for our CFER-based test for the violation of MR by the underlying. Indeed, HS do not assume that options satisfy MR, but they empirical find that the effect of frictions on options seems to have negligible impact on the estimation of CFER. On the other hand, we need this assumption to derive the generalized BKM formulae. Otherwise, relaxing the MR assumption for options would not allow the estimation of RNMs in a model-free manner since model-free ways to estimate the effect of frictions on options are yet to be proposed. Our formulae also generalize BKM in that we allow the underlying asset to pay dividends.

Let $r_{t,T} = \log((S_T + \widetilde{D}_{t,T})/S_t)$ be the *log cum-dividend return* of the underlying stock. The (generalized) BKM formulae estimate the central moments of the distribution of $r_{t,T}$. The model-free implied volatility (MFIV), risk-neutral skewness (RNS), and risk-neutral kurtosis (RNK) of the $\mathbb{Q}^*$-distribution of $r_{t,T}$ are denoted by $MFIV_{t,T}$, $RNS_{t,T}$ and $RNK_{t,T}$, respectively. We consider the derivative which delivers a payoff

$(r_{t,T})^n$ with a positive integer n at time T. We call this derivative the "power-n log return contract" and we denote its present value as $M(n)_{t,T} = e^{-r_f\tau}\mathbb{E}_t^{\mathbb{Q}^*}[(r_{t,T})^n]$.[6]

Proposition 2.4.1 (Generalized BKM formulae)**.** *Assume that the underlying may violate MR, yet option prices satisfy MR. Then, the MFIV, RNS and RNK are given by*

$$MFIV_{t,T} = \sqrt{\frac{e^{r_f\tau}M(2)_{t,T} - \widetilde{\mu}_{t,T}^2}{\tau}}, \tag{2.22}$$

$$RNS_{t,T} = \frac{e^{r_f\tau}M(3)_{t,T} - 3e^{r_f\tau}\widetilde{\mu}_{t,T}M(2)_{t,T} + 2\widetilde{\mu}_{t,T}^3}{[e^{r_f\tau}M(2)_{t,T} - \widetilde{\mu}_{t,T}^2]^{3/2}}, \tag{2.23}$$

$$RNK_{t,T} = \frac{e^{r_f\tau}M(4)_{t,T} - 4e^{r_f\tau}\widetilde{\mu}_{t,T}M(3)_{t,T} + 6e^{r_f\tau}\widetilde{\mu}_{t,T}^2M(2)_{t,T} - 3\widetilde{\mu}_{t,T}^4}{[e^{r_f\tau}M(2)_{t,T} - \widetilde{\mu}_{t,T}^2]^2}, \tag{2.24}$$

where

$$\widetilde{\mu}_{t,T} = CFER_{t,T} + e^{r_f\tau} - 1 - \frac{e^{r_f\tau}}{2}M(2)_{t,T} - \frac{e^{r_f\tau}}{6}M(3)_{t,T} - \frac{e^{r_f\tau}}{24}M(4)_{t,T} \tag{2.25}$$

is the forth-order Taylor series approximation of the risk-neutral expected log stock return (i.e., $\widetilde{\mu}_{t,T} \approx \mathbb{E}_t^{\mathbb{Q}^*}[r_{t,T}]$). The value of the power-n log return contracts $M(n)_{t,T}$ is given by

$$M(n)_{t,T} = \int_{S_t - \widetilde{D}_{t,T}}^{\infty} \eta(K; S_t, n)C_t(K,T)dK + \int_0^{S_t - \widetilde{D}_{t,T}} \eta(K; S_t, n)P_t(K,T)dK, \tag{2.26}$$

where

$$\eta(K; S_t, n) = \frac{n}{(K + \widetilde{D}_{t,T})^2}\left[(n-1)\log\left(\frac{K + \widetilde{D}_{t,T}}{S_t}\right)^{n-2} - \log\left(\frac{K + \widetilde{D}_{t,T}}{S_t}\right)^{n-1}\right]. \tag{2.27}$$

Proof. See Appendix 2.A.3. □

Proposition 2.4.1 provides the formulae to calculate RNMs in a setting where MR may be violated. These formulae modify the BKM formulae in two ways. First, the risk-neutral mean of the log stock return $\widetilde{\mu}_{t,T}$ is modified (equation (2.25)). The expression for the risk-neutral mean log stock return in BKM's original study (equation (39) of BKM) does not contain the CFER term, reflecting their assumption that the

[6] BKM call the power-2, -3, and -4 log return contracts the volatility, cubic and quartic contracts. They also denote $M(n)_{t,T}$ $(n = 2, 3, 4)$ by $V_{t,T}$, $W_{t,T}$, and $X_{t,T}$, respectively.

underlying satisfies MR. Therefore, unless CFER is zero, the original BKM formulae mis-estimate the RNMs of the log stock return because they use a mis-measured mean of the log stock return to calculate central moments.

The second modification is that our formulae consider the cum-dividends return whereas BKM assume that the underlying pays no dividends. The inclusion of dividends modifies the new formulae in two ways compared to BKM. First, the boundary between the call and put integration regions in equation (2.26) changes to $S_t - \widetilde{D}_{t,T}$ from S_t in the original BKM formulae. Second, the "weighting" function, equation (2.27), also changes. These two changes occur because the payoff function changes from the power of the ex-dividend log return $(\log(S_T/S_t))^n$ to that of the cum-dividend log return $(\log((S_T + \widetilde{D}_{t,T})/S_t))^n$.

Regarding our modification on the dividend paying stocks, Proposition 2.4.1 provides RNMs of the cum-dividend return. The next Proposition provides formulae for the RNMs of the *ex-dividend return*, $r^{ex}_{t,T} = \log(S_T/S_t)$.

Proposition 2.4.2 (Generalized BKM formulae for the ex-dividend return)**.** *To estimate the RNMs of the ex-dividend return, the value of power-n contracts is given by*

$$M(n)_{t,T} = \int_{S_t}^{\infty} \eta^{ex}(K; S_t, n) C_t(K,T) dK + \int_0^{S_t} \eta^{ex}(K; S_t, n) P_t(K,T) dK, \quad (2.28)$$

where

$$\eta^{ex}(K; S_t, n) = \frac{n}{K^2}\left[(n-1)\log\left(\frac{K}{S_t}\right)^{n-2} - \log\left(\frac{K}{S_t}\right)^{n-1}\right]. \quad (2.29)$$

Moreover, the expected log stock return, equation (2.25), should be modified to the following equation:

$$\widetilde{\mu}_{t,T} = CFER_{t,T} + e^{r_f\tau} - \frac{\widetilde{D}_{t,T}}{S_t} - 1 - \frac{e^{r_f\tau}}{2}M(2)_{t,T} - \frac{e^{r_f\tau}}{6}M(3)_{t,T} - \frac{e^{r_f\tau}}{24}M(4)_{t,T}. \quad (2.30)$$

The functional forms of equations (2.22) to (2.24) do not change.

Proof. See Appendix 2.A.4. □

Note that equations (2.28) and (2.29) are the same as the original BKM formulae. This is because both our formulae in Proposition 2.4.2 and BKM's original formulae are about the payoff function $\log(S_T/S_t)^n$. However, in contrast to BKM, we assume that the underlying asset pays dividends and this modifies the mean of the log stock return, equation (2.30), which now contains an additional term, $-\widetilde{D}_{t,T}/S_t$. This reflects the fact that the ex-dividend expected return is lower than the cum-dividend expected return by $\widetilde{D}_{t,T}/S_t$.

Finally, we briefly discuss the possible bias in the estimation of RNMs due to the assumption that option prices obey the MR. To this end, let us assume that options violate MR and option prices are expressed as $C_t(K) = \mathbb{E}_t^{\mathbb{Q}^*}[(S_T - K)^+]/R_{t,T}^f + M_t^c(K)$ and $P_t(K) = \mathbb{E}_t^{\mathbb{Q}^*}[(K - S_T)^+]/R_{t,T}^f + M_t^p(K)$, where $M_t^c(K)$ and $M_t^p(K)$ denote the effect of frictions on call and put option, respectively. By revisiting the proof of Proposition 2.4.1 with these notations, it follows that $C_t(K,T)$ and $P_t(K,T)$ in equation (2.A.7) should be replaced with $C_t(K,T) - M_t^c(K)$ and $P_t(K,T) - M_t^p(K)$, respectively. This change leads to the modification of equation (2.26) to

$$\begin{aligned} M(n)_{t,T} = &\int_{S_t - \widetilde{D}_{t,T}}^{\infty} \eta(K; S_t, n) C_t(K,T) dK + \int_0^{S_t - \widetilde{D}_{t,T}} \eta(K; S_t, n) P_t(K,T) dK \\ &- \int_{S_t - \widetilde{D}_{t,T}}^{\infty} \eta(K; S_t, n) M_t^c(K) dK - \int_0^{S_t - \widetilde{D}_{t,T}} \eta(K; S_t, n) M_t^p(K) dK. \end{aligned} \tag{2.31}$$

Equation (2.31) shows that, when options violate MR, the power-n log return contract price equals the sum of the two integral terms in the right-hand side of equation (2.26) (the two integrals in the first line) and the integrals of the effect of frictions on option prices across different strikes (the two integrals in the second line). This means that our formula in Proposition 2.4.1 results in the estimation bias in the power-n log return contract price, which equals the sum of the last two terms in equation (2.31).

Unfortunately, this bias component cannot be estimated in a model-free manner because there are no model-free ways to estimate the effect of frictions on options, $M_t^c(K)$ and $M_t^p(K)$. Nevertheless, to get a feeling of the magnitude of this bias, let us assume that the effect of frictions on option prices, $M_t^c(K)$ and $M_t^p(K)$, is proportional to their option prices with the constant coefficient a, that is, $M_t^c(K) = aC_t(K)$ and $M_t^p(K) = aP_t(K)$ for any strike K. In this case, the bias component equals a times the right-hand side of equation (2.26). This means that our estimation formula, equation (2.26), has a relative error of a. Therefore, the bias in the power-n log return contract is not large as long as the relative size of the friction-related term of options to their option prices is not large. Moreover, this bias component can become even smaller if the sign of effect of frictions $M_t^c(K)$ and $M_t^p(K)$ varies across strikes (for example due to measurement errors) as they offset each other to some extent when the integral across strikes are taken. In any case, examining the bias component under more realistic settings is beyond the scope of this thesis and hence it is best left for future research.

2.4.2 IV and the violation of MR: Log-normal case

Next, we investigate the behavior of BS-IV under the violation of MR. To provide an analytically tractable discussion, we consider a situation where the underlying stock may violate MR, yet it follows a log-normal distribution. In particular, we assume that the future stock price S_T follows a log-normal distribution under the $\mathbb{Q}^*$-measure, that is,

$$S_T \sim \text{Lognormal}\left(\log(S_t - e^{-r_S\tau}\widetilde{D}_{t,T}) + r_S\tau - \frac{\sigma^2\tau}{2}, \sigma^2\tau\right), \tag{2.32}$$

where r_S is the continuously compounded risk-neutral expected return of the stock, $e^{r_S\tau} = \mathbb{E}_t^{\mathbb{Q}^*}[R_{t,T}]$, σ is the constant volatility (i.e., annualized standard deviation) parameter, and $\widetilde{D}_{t,T}$ is the aggregate dividend payment between t and T assumed to be known at time t.[7]

In the presence of market frictions (i.e., non-zero CFER), MR does not hold and r_S will differ from r_f since

$$R_{t,T}^f + CFER_{t,T} = \mathbb{E}_t^{\mathbb{Q}^*}[R_{t,T}] = e^{r_S\tau} \neq e^{r_f\tau} = R_{t,T}^f, \tag{2.33}$$

where the first equality comes from equation (2.11).

Next, we calculate the price of options under this setting. To this end, we assume that option prices satisfy MR and thus option prices are still given by equation (2.18).

[7] To show that r_S is the continuously compounded risk-neutral expected return, we calculate the mean of the log-normal distribution, equation (2.32). This yields

$$\mathbb{E}_t^{\mathbb{Q}^*}[S_T] = \exp\left(\log(S_t - e^{-r_S\tau}\widetilde{D}_{t,T}) + r_S\tau\right) = e^{r_S\tau}(S_t - e^{-r_S\tau}\widetilde{D}_{t,T}) = e^{r_S\tau}S_t - \widetilde{D}_{t,T}.$$

Some more algebra shows that $\mathbb{E}_t^{\mathbb{Q}^*}[R_{t,T}] = e^{r_S\tau}$ under the deterministic dividend $\widetilde{D}_{t,T}$ assumption.

Then, the call option price is given by

$$C_t(K,T) = e^{(r_S - r_f)\tau} \times e^{-r_S\tau}\mathbb{E}_t^{\mathbb{Q}^*}[(S_T - K)^+] = e^{(r_S - r_f)\tau} BS_{call}(S_t, K, \tau, r_S, \widetilde{D}_{t,T}, \sigma), \tag{2.34}$$

where $BS_{call}(S, K, \tau, r, D, \sigma)$ is the BS function with deterministic dividend payment, equation (2.19). The relation $e^{-r_S\tau}\mathbb{E}_t^{\mathbb{Q}^*}[(S_T - K)^+] = BS_{call}(S_t, K, \tau, r_S, \widetilde{D}_{t,T}, \sigma)$ holds because the left-hand side is the price of a call in the case where the "risk-free rate" is r_S. Similarly, the put option price is given by

$$P_t(K,T) = e^{(r_S - r_f)\tau} BS_{put}(S_t, K, \tau, r_S, \widetilde{D}_{t,T}, \sigma). \tag{2.35}$$

Two remarks are in order regarding equations (2.34) and (2.35). First, when MR is satisfied (i.e., $r_S = r_f$), the pricing formulae boil down to the standard BS functions with deterministic dividends. Second, under the violation of MR, the option pricing functions are a function of seven arguments ($S_t, K, \tau, \sigma, r_f, \widetilde{D}_{t,T}$ and r_S) rather than a function of six arguments as in the case of the BS model. This is because the risk-free rate used to discount future expected payoff (equations (2.34) and (2.35)) and the continuously compounded expected stock return differ in the case where MR does not hold.

The BS-IV is defined implicitly via the following equations:

$$IV_t^c = BS_{call}^{-1}(C_t; S_t, K, \tau, r_f, \widetilde{D}_{t,T}), \quad \textit{and} \quad IV_t^p = BS_{put}^{-1}(P_t; S_t, K, \tau, r_f, \widetilde{D}_{t,T}), \tag{2.36}$$

where $BS_{call}^{-1}(C; S, K, \tau, r, \widetilde{D})$ ($BS_{put}^{-1}(P; S, K, \tau, r, \widetilde{D})$) is the inverse function of the BS call (put) option function, viewed as a correspondence between the volatility parameter and option price given other parameters. The following Proposition shows the properties of the BS-IV defined through equation (2.36) in the case where option prices (C_t and P_t in the respective right-hand side arguments in equation (2.36)) are given by equations (2.34) and (2.35), respectively.[8]

Proposition 2.4.3. *(a) Let $IV_t^c(K)$ ($IV_t^p(K)$) be the BS-IV of the call (put) option*

[8] Therefore, we define BS-IV via the BS model which prevails in the case of frictionless markets but then we use a setting with frictions in the underlying asset to see how frictions affect BS-IV. Our approach is not internally inconsistent. The fact that we use BS-IV does not mean that market option prices and the underlying asset price are not affected by frictions; the BS formula serves as a translation mechanism between market option prices and IV and it does not imply that the BS formula is the correct option pricing formula.

price given by equation (2.34) (equation (2.35)), that is,

$$C_t(K,T) = e^{(r_S - r_f)\tau} BS_{call}(S_t, K, \tau, r_S, \widetilde{D}_{t,T}, \sigma) = BS_{call}(S_t, K, \tau, r_f, \widetilde{D}_{t,T}, IV_t^c(K)) \tag{2.37}$$

$$P_t(K,T) = e^{(r_S - r_f)\tau} BS_{put}(S_t, K, \tau, r_S, \widetilde{D}_{t,T}, \sigma) = BS_{put}(S_t, K, \tau, r_f, \widetilde{D}_{t,T}, IV_t^p(K)), \tag{2.38}$$

Then, $IV_t^c(K)$ and $IV_t^p(K)$ satisfy the following approximate equations, respectively:

$$IV_t^c(K) \approx \sigma + \frac{CFER_{t,T}}{\sqrt{\tau} R_{t,T}^f} \frac{S_t}{S_t - e^{-r_f\tau}\widetilde{D}_{t,T}} \frac{\Phi(d_1)}{\phi(d_1)}, \tag{2.39}$$

$$IV_t^p(K) \approx \sigma - \frac{CFER_{t,T}}{\sqrt{\tau} R_{t,T}^f} \frac{S_t}{S_t - e^{-r_f\tau}\widetilde{D}_{t,T}} \frac{\Phi(-d_1)}{\phi(-d_1)}, \tag{2.40}$$

where $\phi(x)$ and $\Phi(x)$ are the probability density function and the cumulative density function of the standard normal distribution and

$$d_1 = \frac{\log \frac{S_t - e^{-r_f\tau}\widetilde{D}_{t,T}}{K} + (r_f + \frac{\sigma^2}{2})\tau}{\sigma\sqrt{\tau}}. \tag{2.41}$$

(b) When $CFER_{t,T} = 0$, then $IV_t^c(K) = IV_t^p(K) = \sigma$ holds. On the other hand, when $CFER_{t,T}$ is positive (negative), $IV_t^c(K)$ is greater (smaller) than σ for any strike and decreasing (increasing) in K, whereas $IV_t^p(K)$ is smaller (greater) than σ for any strike and increasing (decreasing) in K.

Proof. See Appendix 2.A.5. □

Proposition 2.4.3 showcases two interesting findings about the properties of the BS-IV. First, in the case where the underlying violates MR, the BS-IV does not correctly estimate the volatility parameter of the log-normal distribution.[9] For instance, when CFER is positive, the call (put) BS-IVs overestimate (underestimate) the true σ. Figure 2.2 manifests this finding by showing the IVs given by equations (2.39) and (2.40) as a function of moneyness with time-to-maturity $\tau = 1/8$ (about 45 days), the

[9]Hentchel (2003) analyzes how the IV as an estimate of the true volatility is biased due to measurement errors in option prices, the underlying prices, the risk-free rate (which equals the expected stock return under the MR assumption) and other parameters. Our result is related to his analysis because the bias in our case can be interpreted as a result of using the wrong value for the expected stock return when inverting the BS function.

risk-free rate $r_f = 3\%$, the annualized CFER $CFER_{t,T}/\tau = 1\%$, the deterministic dividend-to-stock $\widetilde{D}_{t,T}/S_t = 0.5\%$ (i.e., 4% annualized dividend yield) and the log-normal volatility parameter $\sigma = 20\%$.[10] This pattern appears because call (put) options are more expensive (cheaper) compared to the BS model case (i.e., $r_S = r_f$) when $CFER_{t,T} > 0$ because a higher stock drift r_S makes call (put) options more (less) likely to expire in-the-money.

[Figure 2.2 about here.]

Second, the BS-IVs are not constant across strikes in the case where MR is violated despite the fact that the risk-neutral distribution of the underlying is log-normal. For instance, in the case where CFER is positive, Proposition 2.4.3 (b) shows that the call (put) IV curve is decreasing (increasing) in K. Therefore, the IV skew in this case is a result of market frictions and it would spuriously suggest that the underlying distribution is not log-normal. This finding is related to Leland (1985) and Çetin et al. (2006), who also show that market frictions may result in skewed IV curve even if the underlying follows a log-normal distribution.

To remedy these drawbacks of the conventional BS-IV, we propose the *robust IV* (henceforth, we will call the conventional IV the *standard* BS-IV to distinguish it from the robust IV).

Definition 2.4.1 (Robust IV). *Let r_S be the continuously compounded $\mathbb{Q}^*$-expected stock return, which will differ from r_f if market frictions are present. Then, the robust IV, IV_{rob}, is defined implicitly via*

$$\begin{aligned} e^{-(r_S - r_f)\tau} C_t(K,T) &= BS_{call}(S_t, K, \tau, r_S, \widetilde{D}_{t,T}, IV^c_{rob}), \\ e^{-(r_S - r_f)\tau} P_t(K,T) &= BS_{put}(S_t, K, \tau, r_S, \widetilde{D}_{t,T}, IV^p_{rob}). \end{aligned} \tag{2.42}$$

We implicitly define the robust IV by equating the option price *adjusted* by $e^{-(r_S - r_f)\tau}$ with the BS model price which uses r_S rather than r_f as an input. In contrast, the standard BS-IV is obtained by equating the *unadjusted* option price with the BS model price which uses the r_f as an input. In the case where the underlying satisfies MR (i.e., $r_S = r_f$), equation (2.42) boils down to the definition of the

[10] Note that the numerical value of CFER is in line with the evidence we will present in Section 2.5, which shows that the CFER of the S&P 500 may be as extreme as ±1%.

standard BS-IV. This suggests that the robust IV generalizes the standard BS-IV by accounting for the possible violation of MR.

The generalization of the standard IV to the robust IV is of importance because the latter eliminates the two drawbacks of the standard IV discussed above, arising from violations of MR. First, the robust IV correctly estimates the volatility parameter σ when the underlying distribution is log-normal. This can be shown by comparing equations (2.34), (2.35), and (2.42). Second, in the case where the underlying follows a log-normal distribution, the robust IV is constant across strikes and equals to σ even when MR is violated. This suggests that the slope of the robust IV curve is a better proxy of the risk-neutral skewness of the underlying stock because a zero IV slope calculated based on the robust IV would ensure that the asset price is log-normally distributed.

2.4.3 Estimation of RNMs: Simulation results

The violation of MR may affect the estimation of RNMs via two channels. The first channel is via the theoretical formulae used for their computation, and the second channel is via the IVs used in the process of obtaining a continuum of option prices by interpolation given that in practice option prices trade for discrete rather than a continuum of strikes. To quantify how these two channels may affect the estimation of RNMs, we conduct a simulation exercise.

We assume that the stock price follows the log-normal distribution, equation (2.32), and the stock may violate MR (i.e., $r_S \neq r_f$).[11] The current stock price is set to $S_t = 100$, the risk-free rate is $r_f = 3\%$, the time-to-maturity is $\tau = 1/12$. For simplicity, we assume no dividend payment. The volatility parameter is set to $\sigma = 30\%$. We consider three cases of the expected stock return r_S: $r_S = r_f$, $r_S = r_f + 3\%$, and $r_S = r_f - 3\%$. Since the log stock return, $\log(S_T/S_t)$, follows a normal distribution, the true MFIV, RNS, RNK are $\sigma = 30\%$, zero, and three, respectively, regardless of the value of r_S.

We generate a realistic set of option prices as follows. We assume that options trade at $K = 94, 95, \ldots, 108$. These strikes approximately cover the delta range from

[11] This setup is similar in spirit to Dennis and Mayhew (2009), who analyze how microstructural biases affect the estimation of RNS under the log-normality assumption.

0.2 to 0.8, corresponding to the delta range in the OptionMetrics Volatility Surface file, which we use in the subsequent empirical analysis. We calculate these option prices based on equations (2.34) and (2.35). Then, we calculate the standard BS-IV of these options. Unless $r_S = r_f$, the standard BS-IV curve is not flat (Proposition 2.4.3). On the other hand, the robust IV curve is flat at the level of $\sigma = 30\%$.

In line with the standard estimation procedures in the literature (e.g., Stilger et al., 2017), we interpolate the discrete option IVs by the cubic Hermite polynomial in the moneyness-IV metric and extrapolate IVs horizontally beyond the highest and lowest strikes. We follow Stilger et al. (2017) and interpolate and extrapolate call IVs and put IVs, separately. Then, we convert the interpolated IV curves to 1001 option prices over the equally-spaced strikes in the moneyness range $[1/3, 3]$. Finally, we numerically calculate the integrals in the BKM formulae and calculate the RNMs. To perform these procedures, we have two choices regarding the type of IV to be interpolated and two choices regarding the type of the RNM formulae. These yield four estimation specifications, OS-, OR-, GS-, and GR-RNMs. The first letter of the prefix stands for the type of the formulae (the "Original" BKM formulae or the "Generalized" BKM formulae) and the second letter stands for the type of the IVs used for the IV interpolation (the "Standard" IVs or the "Robust" IVs).

Table 2.2 reports the result of the estimated RNMs by the four specifications. We can see that MFIV and RNK are not affected by the violation of MR in our simulated situation. On the other hand, RNS is affected both by the choice of the BKM formula and the employed IVs. Unless *both* the generalized formulae and the robust IV are employed, the estimated RNS differs from the true value of zero. This result shows that both the explicit consideration of CFER in the BKM formulae and the use of the robust IV for the inter- and extrapolation matter for the estimation of RNS.

[Table 2.2 about here.]

Why does the input IV to the interpolation and extrapolation process matter? Roughly speaking, this is because the extrapolation of the standard IV results in biases in the extrapolated option prices, and these biases transmit to the power-n log return contract.

To fix ideas, we consider a positive CFER case. In this case, the call and put standard IV curves are given by the two dotted lines in Figure 2.3. In the extrapolation step, the IV curves are horizontally extended beyond the observed moneyness range (i.e., the two solid lines in Figure 2.3). Since the extrapolated call (put) line is above (below) the true standard IV line, this means that the extrapolation of the standard IV curve results in the overestimation (underestimation) of the extrapolated call (put) option prices in the case CFER is positive.

[Figure 2.3 about here.]

Then, how the overestimation of OTM call options and underestimation of OTM put options transmit to the power-n log return contract prices? We answer this question based on the following Lemma about the sign of the weighting function $\eta(K; S_t, n)$ in equation (2.26).

Lemma 2.4.1. *For odd-degree power-n log return contracts, $\eta(K; S_t, n)$ satisfies the following inequalities.*

$$\eta(K; S_t, n) \begin{cases} < 0 & K/S_t < 1, \\ \geq 0 & 1 \leq K/S_t \leq e^{n-1}, \\ < 0 & K/S_t > e^{n-1}. \end{cases} \tag{2.43}$$

For even-degree power-n log return contracts, $\eta(K; S_t, n)$, satisfies the following inequalities.

$$\eta(K; S_t, n) \begin{cases} \geq 0 & K/S_t \leq e^{n-1}, \\ < 0 & K/S_t > e^{n-1}. \end{cases} \tag{2.44}$$

Proof. See Appendix 2.A.7. □

To simplify our discussion, we ignore the respective last cases in equations (2.43) and (2.44) about the far deep call OTM region $K/S_t > e^{n-1}$. This simplification is innocuous for empirical applications because call option prices in this deep OTM region are typically negligible and do not much affect the calculation of the call integral in equation (2.26).

Equation (2.43) shows that the η function of the *odd*-degree power-n contract is negative in the put OTM region and positive in the call OTM region. Therefore, the undervaluation of OTM put options and the overvaluation of OTM call options *both* result in an upward bias in the power-3 contract price which enters in the calculation of RNS. Therefore, a positive CFER results in an upward bias in the estimated RNS via the bias in OTM option prices arisen during the extrapolation process.

On the other hand, equation (2.44) shows that the η function of the *even*-degree power-n contract is always non-negative. Consequently, the overvaluation in OTM call prices and undervaluation in OTM put prices have the opposite effect on the power-2 and power-4 contract prices, leaving these prices largely unaffected. This explains why MFIV and RNK are largely unaffected by the bias arisen during the extrapolation process.

2.4.4 Relation to the IV slope measure

Xing et al. (2010) (XZZ) find that the difference between the standard IV of an OTM put and that of an ATM call, $XZZ = IV^p_{OTM} - IV^c_{ATM}$ *negatively* predicts the cross-section of future stock returns. Their finding is consistent with the recent literature on RNS which documents a positive relation between RNS and future realized stock returns, given that the XZZ measure is a crude measure of RNS (e.g., a higher value of XZZ implies that the risk-neutral distribution of the asset returns is skewed to the left).

However, the XZZ measure can be decomposed in two components:

$$XZZ = IV^p_{OTM} - IV^c_{ATM} = (IV^p_{OTM} - IV^p_{ATM}) + (IV^p_{ATM} - IV^c_{ATM}) = POMA - IVS, \tag{2.45}$$

where $POMA = IV^p_{OTM} - IV^p_{ATM}$ is the *put out-minus-at* IV skew measure studied by Doran and Krieger (2010) and Fu et al. (2016) and $IVS = IV^c_{ATM} - IV^p_{ATM}$ is the *implied volatility spread* studied by Bali and Hovakimian (2009) and Cremers and Weinbaum (2010) among others.

Regarding the POMA component, Proposition 2.4.3 shows that POMA is not equal to zero even when the underlying price follows a log-normal distribution, in the case where MR is violated.[12] The IVS component also biases the calculation of RNS because it is the difference between the call and put option IVs at the same strike and hence it is not related to RNS. Instead, Proposition 2.4.3 shows that a non-zero IVS

[12]Existing literature reports either insignificant predictive power of POMA (Fu et al., 2016) or even positive relation between POMA and future returns (Doran and Krieger, 2010). This suggests that the POMA component should not be the source of *negative* relation between XZZ and future returns.

arises in the case where MR is violated; given the property of the cumulative standard normal density function $\Phi(x) + \Phi(-x) = 1$, subtracting equation (2.40) from (2.39) yields

$$IV_t^c(K) - IV_t^p(K) \approx \frac{CFER_{t,T}}{R_{t,T}^f} \frac{S_t}{S_t - e^{-r_f \tau} \widetilde{D}_{t,T}} \frac{1}{\sqrt{\tau}\phi(d_1)} = \frac{S_t}{R_{t,T}^f \mathcal{V}_{BS}(K)} CFER_{t,T}, \tag{2.46}$$

where $\mathcal{V}_{BS}(K) = (S_t - e^{-r_f \tau} \widetilde{D}_{t,T})\phi(d_1)\sqrt{\tau}$ is the BS vega (i.e., the sensitivity of option prices to the change in the volatility parameter). Therefore, a non-zero IVS arises if CFER is non-zero, that is, it is an evidence of the violation of MR.[13] These considerations cast doubt on whether the XZZ measure proxies RNS accurately.

However, the XZZ measure will proxy RNS accurately if the IV slope is calculated using the robust IV rather than the standard IV. First, as we have discussed at the end of Section 2.4.2, POMA (i.e., the slope of the put IV curve) calculated based on the robust IV curve is a better proxy of RNS because a zero robust IV slope means that the asset price is distributed log-normally. Second, the following Proposition shows that IVS equals zero and hence it does not affect the XZZ measure once the robust IVs are used.

Proposition 2.4.4. *Assume that call and put option prices satisfy MR, that is, they satisfy the risk-neutral valuation formulae (equation (2.18)). Let IV_{rob}^c and IV_{rob}^p be the robust IV calculated from the call and the put option prices with the same strike price, respectively. Then, $IV_{rob}^c = IV_{rob}^p$ holds regardless of the distribution of the underlying asset.*

Proof. See Appendix 2.A.6. □

In Section 2.6.5, we document that the XZZ slope measure calculated based on the robust IV does not predict future stock returns. We also document that the predictive power of the XZZ measure based on the standard IV stems from its mechanical negative correlation with IVS and CFER.

[13] HS show that equation (2.46) holds regardless of the distribution of the future underlying price.

2.5 Data and estimation results

2.5.1 Dataset

We obtain daily option prices and IVs for January 1996 to December 2017 from the two OptionMetrics Ivy DB (OM) data files, the *Volatility Surface* (VS) file and the *Standardized Options* (SO) file. OM estimate the IV surface of call options and put options separately, and these two files provide the estimated IVs and corresponding European option prices at standardized maturities such as 30, 60, 91, 182, and 365 day-to-maturity. We use 30-day-to-maturity data because subsequently we use the estimated RNS as a sorting variable to construct monthly-rebalanced portfolios. The VS file provides option IVs, option prices, and strike prices at 13 standardized delta-denominated strikes (calls at 0.2, 0.25, . . . , 0.8 delta and puts at -0.2, -0.25, . . . , -0.8 delta). The SO file provides call and put option prices at the forward at-the-money (ATM), $K = F_{t,T} = e^{r_f \tau} S_t - \widetilde{D}_{t,T}$. We also obtain the underlying price and the risk-free rate data from the OM database.

Note that the theoretical formulae to estimate CFER and RNMs rely on European option prices, whereas exchange listed U.S. options are American style. However, the use of the OM VS and SO files is not a problem for our empirical estimation procedures because these files provide the IVs and corresponding European option prices adjusted for the early exercise premium. Our treatment is in line with recent studies using the OM dataset (e.g., Martin and Wagner, 2018).

We obtain stock return data from CRSP. We also calculate the market beta, firm size, book-to-market, profitability, and investment characteristics from CRSP and Compustat databases. See Appendix 2.B.1 for the detailed description of these characteristics variables. Our universe of individual equities is the U.S. common stocks traded at NYSE/Amex/NASDAQ. To obtain this universe, we link the OM database with the CRSP database and keep only the common stocks (CRSP SHRCD: 10 or 11) traded at NYSE/Amex/NASDAQ (CRSP EXCHCD 1, 2, or 3). Appendix 2.B.2 explains how we link the OM, CRSP, and Compustat databases. Finally, we obtain standard risk factors (such as the Fama and French, 1993 three factors) from the respective author's website.

2.5.2 Estimation procedures for the option-implied variables

We estimate CFER for U.S. individual common stocks traded at NYSE/Amex/NASDAQ as well as the S&P 500 index. HS document empirically that the scaled deviations from put-call parity reliably estimate CFER accurately, that is,

$$CFER_{t,T} \approx \frac{R^f_{t,T}}{S_t}(\widetilde{S}_t(K) - S_t), \tag{2.47}$$

where $\widetilde{S}_t(K) = C_t(K,T) - P_t(K,T) + K/R_{f,t,T} + \widetilde{D}_{t,T}$ is the price of a synthetic stock price. To calculate the synthetic stock prices, we use the pairs of put and call prices with 30 day-to-maturity at the forward ATM strike recorded in the OM SO file.We back out the dividend payment $\widetilde{D}_{t,T}$ from the forward price recorded in the OM SO file. Appendix 2.B.3 provides the detailed estimation procedure of CFER.

We estimate the individual stocks' risk-neutral moments (RNMs) via the original BKM formulae adjusted for dividend payments (O-RNMs) and via the generalized BKM formulae (G-RNMs), separately, to examine to what extent the violation of MR in the underlying affects the estimated value of RNMs.

To estimate the O-RNMs, we use equations (2.26) and (2.27) by setting CFER equal to zero in equation (2.25) because O-RNMs are computed assuming that MR holds. We follow Stilger et al. (2017) and Borochin and Zhao (2018) to interpolate the (standard) IVs obtained from the OM VS file by the cubic Hermite polynomial interpolation in the moneyness-IV metric, and extrapolate horizontally beyond the observed option moneyness. We interpolate and extrapolate call IVs and put IVs, separately. Then, we convert the splined IV curves to 1001 option prices at equally-spaced strikes over the moneyness range $[1/3, 3]$ and numerically calculate the integrals. Note that, strictly speaking, the O-RNMs estimation formulae do not exactly coincide with the original formulae in BKM due to the modification regarding dividend payments. In Section 2.6.1, we will document that this modification does not drive our empirical findings. We estimate G-RNMs just as we do for O-RNMs with the difference being that we (i) use the estimated 30-day CFER in equation (2.25), and (ii) interpolate the robust IV instead of the standard IV since our simulation result in Section 2.4.3 shows that using the standard IV may bias RNMs. To obtain the robust IVs, we use option prices in the SO file in conjunction with equation (2.42). We disentangle how these two modifications in the IV input and the employed formulae affect the empirically estimated RNS in Section 2.6.1.

We treat both the O-RNMs and G-RNMs missing if the estimated CFER is not available.[14] We also discard very few month-stock IV surface observations if either calls' or puts' strikes recorded in the OM VS file are not strictly monotonic in deltas. Such a case may occur when the interpolated IV surface exhibits a very steep skew. See Appendix 2.B.4 for a detailed description of the calculation of RNMs.

Finally, we calculate two IV slope measures as follows. First, we calculate the *original* XZZ measure as $XZZ^o = IV^p(-0.2) - IV^c(0.5)$, where $IV^c(d)$ ($IV^p(d)$) stands for the standard call (put) IV at option's delta equal to d. We obtain 30-day-to-maturity IVs from the OM VS file. Next, we calculate the *robust* XZZ measure based on the same formulae, yet we use the robust IVs, that is, $XZZ^r = IV^p_{rob}(-0.2) - IV^c_{rob}(0.5)$. To obtain the robust IVs, we convert implied option prices recorded in the OM VS file and then we use equation (2.42) to calculate the robust IVs.

2.5.3 The estimated CFER: Do U.S. stocks violate MR?

2.5.3.1 S&P 500 index

Figure 2.4a depicts the daily estimated 30-day-to-maturity CFER of the S&P 500 index. We can see that the estimated CFER is highly volatile; the proportion of CFER being positive is about 48% in our sample, that is, the estimated CFER takes positive and negative values with almost the same frequency. This suggests that the $\mathbb{Q}^*$-expected return of the S&P 500 index is greater or smaller than the risk-free rate with almost the same probability. This result is in contrast to the results obtained by the studies which test MR via the implied stock approach. These studies find that $S_t^* > S_t$ for more than 90% of samples which would imply a positive CFER for most of the time; this correspondence between CFER and the implied stock price approach stems from equation (2.15), which holds under the key assumption of the implied stock approach. As we have shown in Section 2.3.3, the result in the previous literature is likely to reflect a negative IV skew rather than a violation of MR.

Given the high volatility of the daily estimated CFER, to visualise the pattern of

[14] This situation occurs when the OM VS file contains the data, yet the OM SO file does not. Since there is only one such month-stock observation, this treatment is unlikely to affect our analysis.

the dynamics of CFER, Figure 2.4b plots the 21-trading day moving average of the absolute value of the estimated CFER, which measures the degree of the departure of the index from MR (see equation (2.11)). Two remarks can be drawn. First, CFER takes extreme values periods of market distress such as the collapse of the IT bubble (around October 2000) and the market meltdown over the recent financial crisis (around October 2008). This is expected given that in these periods markets frictions and hence violations of MR intensify. This pattern is consistent with the findings in HS regarding the CFER of individual stocks; they find that individual stocks tend to take more extreme CFER values during periods where the market is distressed.

Second, apart from the sporadic extreme values of CFER, the S&P 500 index's CFER exhibits a salient downward trend in absolute value terms. In particular, until 2003, the absolute value of CFER lies between 1% to 3% per year, whereas after 2003, it decreased to less than 1% for most of times. This pattern is in line with HS who show theoretically that the absolute value of CFER is less than the round-trip transaction costs, implying that the decline in the absolute CFER value is related to a decline in transaction costs. Indeed, transaction costs decreased during early 2000s due to various institutional reforms.[15]

[Figure 2.4 about here.]

To assess the statistical significance of the violation of MR, we test the statistical significance of the time-series mean of the *absolute* CFER by means of a 95% ($\pm 2\sigma$) confidence interval; the absolute value of CFER measures the magnitude of the deviation of the risk-neutral expected return from the risk-free rate (equation (2.11)), that is, it measures the magnitude of the violation of MR.[16]

[15]For example, Green et al. (2017) argue that "*the adoption of Reg. FD in October 2000 and the decimalization of quotes in January 2001*" reduced "*effective spreads, price impact, and trading costs*." They also point out that the introduction of the autoquoting software by NYSE between January and May 2003 "*led to dramatic reductions in trading frictions and costs*" (Green et al., 2017, p.4424). Similarly, French (2008) documents that the per transaction "trading costs" of passive investment backed out from securities firms' commissions and trading gains data decreased by the half from mid 1990s to 2000s.

[16]We do not examine the mean of the time-series of the raw estimated CFER. This is because a negligible time-series mean CFER does not imply that the magnitude of CFER is always negligible and hence the stock satisfies MR. For example, the mean of the raw estimated CFER can be close to zero if it takes large positive and negative CFER values yet with similar frequency. The absolute value of our estimated CFER (the scaled deviations from put-call parity) is similar in spirit to

The 95% confidence interval of the mean value of the absolute CFER is 1.27% $\pm$ 0.08% per year, when we use the full sample of CFER values. Given the aforementioned salient decline in the absolute CFER value, we also split the observations into 1996 to 2002 and 2003 to 2017. The $\pm 2\sigma$ confidence interval of the former period is 1.98% $\pm$ 0.1%, whereas that of the latter period is 0.94% $\pm$ 0.06%. Therefore, the mean absolute CFER is statistically different from zero in both periods and it is also economically non-negligible (e.g., compared to the U.S. equity premium, which Fama and French, 2002 estimate as 4.32% per year).

2.5.3.2 Individual equities

Next, we investigate the estimated CFER of individual equities. We estimate CFER for 12.2 million stock-day observations from 1996 to 2017 for 6,824 distinct stocks. This yields, on average, estimated CFER for about 2,200 stocks on each trading day. Table 2.3, Panel A, reports the summary statistics of the daily estimated CFER. We can see that the estimated CFER exhibits large variations; the standard deviation is 15.9% per year and the range from fifth to 95th percentile points reaches 31.0% per year. This suggests that the estimated CFER can take economically large value. Compared to the magnitude of variation, the mean and median of the estimated CFER are close to zero (-1.1% and -0.6% per year, respectively), implying that the estimated individual stock CFER take positive and negative values with almost the same probability. Regarding the absolute value of CFER, which reflects whether MR holds for individual stocks, the mean and median are 6.8% and 2.8% per year, respectively. This implies that market frictions have an economically non-negligible impact on the expected return of individual stocks and MR is violated for these stocks. Panel B reports the summary statistics of the estimated CFER over the end-of-month trading dates. We can see that the summary statistics are qualitatively the same with those obtained from the all daily estimated CFER.

[Table 2.3 about here.]

Similar to the S&P 500 index case, we test for each stock whether MR holds by assessing whether its time-series mean absolute CFER is statistically different from

the market dislocation index proposed by Pasquariello (2014) in that it is calculated based on the absolute value of deviations from the law of one price.

zero. We adopt a critical value of three for the t-test to decide on the significance. In addition, we record the number of stocks which violate MR and at the same time their respective absolute CFER is greater than a reference point. We consider four alternative values for the reference point, zero, 1%, 2%, and 3% per year (the zero reference point corresponds to testing whether the MR is violated).

The left part of Table 2.3, Panel C, reports the result of this counting exercise based on the daily estimated CFER, where there reference value is specified in the first column. We conduct this counting exercise over the following three groups: stocks with at least 21 non-missing CFER observations, those with at least 1,200 non-missing observations, and those with 2,500 non-missing observations. From the second to fifth rows in Panel C, we can see that the mean absolute CFER is significantly larger than 1% per year for virtually all stocks, regardless of the number of non-missing CFER observations. Moreover, about 75% of four stocks with at least 21 valid observations have the mean absolute CFER larger than 3% per year. In short, this result suggests that almost all stocks have significantly positive CFER (i.e., violating MR) and for a large part of stocks, the magnitude of the violation of MR is economically sizable.

We repeat the same exercise using only the end-of-month estimated CFER and the result is reported in the right part of Panel C. Compared to the daily observations case, the proportion of the stocks which have statistically significant positive mean absolute CFER decreases. However, this is a mechanical result; the monthly dataset contains far less observations and hence the standard deviations would be larger. Nevertheless, the implication is the same with the analysis using the daily observations. Around 90% stocks have the mean absolute CFER greater than 1% and about 40% of stocks have the mean absolute CFER greater than 3%.

These results together with the findings in HS strongly suggest that individual stocks' expected returns are affected by market frictions, that is, individual stocks' CFER is not negligible and hence individual stocks frequently violate MR.

2.5.4 Generalized RNMs: Summary statistics

As a preliminary step, for any given RNM, we provide summary statistics and compute the correlation between O-RNM and G-RNM to assess the effect of the violation of MR in the underlying on the estimated RNMs. We report the summary statistics of end-of-month observations since we use only end-of-month RNMs in the subsequent monthly-rebalancing portfolio analysis. Our untabulated results confirm that there are no qualitative differences in the summary statistics if we use all daily observations.

Table 2.4, Panel A, reports the descriptive statistics of CFER and the estimated RNMs. We also report the difference between O-RNMs and G-RNMs (e.g., ΔMFIV equals O-MFIV minus G-MFIV). This difference can be interpreted as the bias in the estimated O-RNMs due to ignoring the violation of MR. We estimate CFER and RNMs for 582,796 month-stock observations over 264 months period from January 1996 to December 2017. Therefore, on average, the estimated CFER and RNMs are available for about 2,200 stocks in each month. The median of ΔMFIV, ΔRNS and ΔRNK are close to zero (-0.02, -0.01, and 0.01, respectively). This implies that the bias in the O-RNMs takes positive and negative sign by almost the same probability. Next, we focus on the ratio of the standard deviation of ΔRNM and that of the respective G-RNM. By viewing the decomposition O-RNM = G-RNM + ΔRNM as the decomposition of O-RNM into the true RNM and the bias term, this ratio of standard deviations can be interpreted as a rough measure of the "noise-to-signal" ratio. The ratio of the MFIV is about 6% ($\sim 1.52/26.68$), that of RNS is about 48% ($\sim 0.23/0.48$) and that of RNK is 34% ($\sim 0.39/1.15$). Therefore, the impact of bias in the estimated O-RNMs is the biggest for RNS.

Table 2.4, Panel B, reports the pairwise Pearson and Spearman correlations between O-RNMs and G-RNMs as well as those between ΔRNM and CFER. We can see that O-RNMs and G-RNMs are almost perfectly correlated for MFIV (1.00 for both Pearson and Spearman coefficients) and for RNK (0.94 and 0.96 for Pearson and Spearman coefficients, respectively) and thus the CFER-adjustment has a negligible impact for the estimated MFIV and RNK. On the other hand, in the case of RNS, the two correlation coefficients are 0.90 and 0.89, respectively. Even though these values are still high, in Section 2.6, we show that the difference between O-RNS and G-RNS creates significant differences in their predictive power for the cross-section of future stock returns.

We can also see that the correlation between CFER and RNMs takes its highest values for O-RNS (0.36 and 0.34), whereas it is almost zero for G-RNS. Moreover, the correlation of CFER with ΔRNM takes its highest value for ΔRNS (0.82 and

0.97). This suggests that the correlation between CFER and O-RNS stems from the correlation between CFER and the "bias" term in the O-RNS due to the violation of MR, ΔRNS.

[Table 2.4 about here.]

2.6 Predictive power of the risk-neutral skewness

In Section 2.5.4, we have found that the violation of MR in the underlying affects mostly RNS among the three RNMs. In this Section, we investigate further whether the difference between G-RNS and O-RNS is economically significant, too. This will shed light on whether the violation of MR are economically significant for the purposes of computing RNS. In Section 2.6.1, we document that O-RNS predicts the cross-section of future stock returns, whereas G-RNS does not. In Section 2.6.2, we confirm these findings by conducting further robustness tests. The ability of O-RNS to predict future stock returns is of particular interest; as we have seen in Section 2.2, there is no unanimous agreement on the sign of this predictive relation and the mechanism of the predictive power of O-RNS. In Section 2.6.3, we show that the violation of MR (i.e., non-zero CFER) provides an explanation for the ability of O-RNS to predict stock returns. This is because CFER is correlated with the bias term ΔRNS caused by the violation of MR and thus it is correlated with O-RNS. Then, in Section 2.6.4, we argue that our CFER-based mechanism of the predictive power of O-RNS encompasses the limits-of-arbitrage story. Finally, in Section 2.6.5, we investigate the predictive power of an alternative measure of the risk-neutral skewness, Xing et al.'s (2010) IV slope measure.

2.6.1 Predictive power of O- and G-RNS: Baseline result

We examine whether O-RNS and G-RNS predicts stock returns cross-sectionally by means of portfolio sort analysis. At the end of each month, we sort stocks into decile portfolios based on the estimated O-RNS and G-RNS, separately. For each estimator of RNS, we form equally-weighted and value-weighted decile portfolios and the long-short spread portfolios, where we go long in the portfolio with the highest RNS and go short in the portfolio with the lowest RNS. Table 2.5 reports the average post-ranking returns of the equally- and value-weighted decile portfolios formed based on the estimated O-RNS and G-RNS, separately, as well as the average post-ranking returns and the risk-adjusted returns (alpha) of the high-minus-low spread portfolios with respect to the CAPM, the Fama and French (1993) three-factor model (FF3), the Carhart (1997) four-factor model (FFC), the Fama and French (2018) five- and six-factor models (FF5 and FF6), the Hou et al. (2015) q-factor model (q4), and the Stambaugh and Yuan (2017) mispricing factor model (SY).

Regarding the results based on O-RNS, we can see a clear increasing monotonic relation between O-RNS and both the equally- and value-weighted decile portfolio returns. As a result, the average returns of the equally- and value-weighted long-short portfolios are 114 bps and 79 bps per month, respectively. These values are statistically highly significant (the t-statistics are greater than 3.5). Moreover, the alphas of the spread portfolios are significant and positive; their t-statistics are greater than three in most of the cases, which is the recommended threshold by Harvey et al. (2016), with the only exception being α_{FF5} and α_{q4} of the value-weighted spread portfolio. The positive sign of the alphas is in line with the results of recent studies (e.g., Stilger et al., 2017; Borochin and Zhao, 2018; Chordia et al., 2019), that is, stocks with lower RNS underperform.

[Table 2.5 about here.]

On the other hand, in the case where we sort stocks based on the G-RNS, the evidence for the predictive power of G-RNS is much weaker. The t-statistics of the average returns and alphas do not exceed three (with the exception of α_{SY} of the equally weighted portfolio), which is the threshold value for t-statistics set by Harvey et al. (2016). Moreover, all value-weighted alphas are insignificant at a 1% level and they are insignificant even at a 5% level except α_{FFC}. Given the recent literature on the data snooping concerns which recommends a high threshold t-statistics value (e.g., Harvey et al., 2016) and a value-weighted portfolio construction (e.g., Bali et al., 2016; Hou et al., 2018), these results suggest that the G-RNS does not predict future stock returns.

Table 2.6 reports the factor loadings obtained from estimating various factor models on the returns of the value-weighted spread portfolios sorted by O-RNS and G-

RNS, separately. We can see that there is no qualitative change in the economic magnitude of the factor loadings between the O-RNS-sorted and G-RNS-sorted spread portfolios. This suggests that the significant change in alphas (i.e., the intercepts) of the long-short portfolio when switching from the O-RNS to the G-RNS sorting criterion is not caused by the change in the exposure of the two portfolios to standard risk-factors.

[Table 2.6 about here.]

The difference in the predictive ability of O-RNS and G-RNS may stem from two sources: the difference in the formula used to compute RNS and/or the difference in the IVs used at the interpolation and extrapolation stage. We disentangle the effect from these two sources by estimating RNS in four different ways: OS-, OR-, GS-, and GR-RNS. The first letter of the prefix stands for the type of the formula ("Original" or "Generalized" BKM formula). The second letter of the prefix stands for the type of the IVs ("Standard" IVs or "Robust" IVs).[17] Figure 2.5 shows the t-statistics of the average return and the t-statistics of the alphas of the long-short spread portfolios constructed by sorting stocks into decile portfolios based on these four alternative RNS estimates. We can see that the magnitude of t-statistics decrease from OS-, OR-, GS-, to GR-RNS. This pattern suggests that both channels of the modification (the formula and the IV) contribute to the decrease in the predictive power of RNS in line with our simulation result in Section 2.4.3. These findings suggest that the previously documented ability of O-RNS (OS-RNS) to predict stock returns arises because the violation of MR had not been taken into account in both the RNS formula as well as in the computation of IVs.

[Figure 2.5 about here.]

2.6.2 Predictive power of O- and G-RNS: Robustness results

Next, we conduct further robustness tests to assess whether the predictive power of RNS ceases to exist once the violation of MR is taken into account. First, we run the Fama and MacBeth (1973) (FM) regressions. At the end of each month, we conduct a cross-sectional regression where we regress the stock returns during the

[17] Therefore, O-RNS (G-RNS) investigated above is labeled OS-RNS (GR-RNS) here.

succeeding month on the estimated RNS and other firms' and stocks' characteristics. Then, we report the time-series average of the estimated cross-sectional intercepts and the coefficients on the characteristic variables together with their respective t-statistics. We conduct two types of regressions for the cross-sectional regression at any point in time: the ordinary least square (OLS) regression and the value-weighted least square (VWLS) regressions employed by Green et al. (2017). OLS- and VWLS-based FM regressions can be viewed as the counterparts of the equally-weighted and value-weighted portfolio in the portfolio analysis, respectively. Therefore, given the recent data-snooping concerns, it is desirable to examine the statistical significance of O-RNS and G-RNS under the VWLS regression (see also a related discussion in Hou et al., 2018).

Table 2.7 reports the results from the FM regressions. Columns (1) to (4) show the results including O-RNS as a regressor. The results from the two univariate regressions of O-RNS, Columns (1) and (2), show that O-RNS is a statistically significant predictor of stock returns. Moreover, Columns (3) and (4) show that this predictive power remains even after controlling for standard characteristic variables. These results corroborate our portfolio sort result based on O-RNS as well as the finding in the previous literature (e.g., Stilger et al., 2017; Borochin and Zhao, 2018).

Next, we examine the predictive power of G-RNS in the FM setting. Columns (5) and (6) show that G-RNS does not predict stock returns in the univariate regression given Harvey et al.'s (2016) criterion for the t-statistic threshold, in line with our finding from the portfolio sort analysis. The results in the multivariate regressions are mixed; G-RNS becomes significant under the OLS regression (Column (7)), whereas it remains insignificant under the VWLS regression (Column (8)). These results are analogous to our portfolio sort result; we find several alphas of equally-weighted G-RNS spread portfolio are significant, whereas those of value-weighted G-RNS spread portfolio are not. Therefore, given the recent data snooping literature, these FM regression results suggests that G-RNS does not predict future stock returns.

[Table 2.7 about here.]

Second, we repeat the portfolio sort analysis of Section 2.6.1 by forming quintile portfolios instead of forming decile portfolios. Each quintile portfolio contains twice as many stocks as a decile portfolio, and as a result, quintile portfolios are more diversified and hence less affected by possible outliers. Table 2.8 reports the results. We can see that the O-RNS-sorted spread portfolios earn positive and significant average return and alphas, whereas the G-RNS-sorted spread portfolios do not earn significant returns just as it was in the case where decile portfolios were formed. This suggests that our baseline result is not driven by possible outliers.

[Table 2.8 about here.]

Finally, we examine whether our treatment of dividend payments in the estimation of RNS affects its predictability. So far, we have found that O-RNS with dividends being taken into accounts, predicts stock returns just as the previous literature had found using the original BKM formulae, whose derivation assumes non-dividend paying stocks. To further confirm that the (lack of the) predictability of O-RNS (G-RNS) is robust to the treatment of dividend payments, we estimate RNS by ignoring the dividend payment adjustment; we re-estimate the O- and G-RNS by setting $\widetilde{D}_{t,T} = 0$ (i.e., assuming as if all stocks pay no dividends) and repeat the portfolio sort analysis. Even though this is a theoretically imprecise treatment, this exercise would expedite the comparison between our results and those in the previous literature which does not consider dividend payments in the calculation of RNS.

Table 2.9 reports the results. As we can see, there are no qualitative changes in the predictive power of the O-RNS and G-RNS; the O-RNS predicts future returns whereas the G-RNS does not. This result suggests that our findings on the (lack of the) predictability of O-RNS (G-RNS) as well as the findings in the previous literature are not affected by whether dividend payments are taken into account for the estimation of RNS.

[Table 2.9 about here.]

We also estimate the RNS of the *ex-dividend return* using Proposition 2.4.2. Then, we repeat the portfolio sorting analysis done in Section 2.6.1 by using the estimated O- and G-RNS of the ex-dividend returns. Our untabulated result shows that the results do not differ from the results obtained from the baseline analysis. This is sensible because, under the deterministic dividend assumption, the distribution of the ex-dividend return coincides with that of the cum-dividend return shifted by a deterministic amount, and hence the higher moments do not change.[18]

2.6.3 CFER and the predictive power of O-RNS

So far, we have confirmed that O-RNS predicts future returns whereas G-RNS does not. We conjecture that O-RNS predicts future returns because its estimation error, ΔRNS, arising from the violation of MR, predicts stock returns. The rationale is that the violation of MR and hence the estimation error occurs because of market frictions, that is, because of a non-zero CFER. The estimation bias in O-RNS caused by the violation of MR makes it a noisy measure of CFER, of which HS document strong return predictive power. This conjecture is in line with the lack of the predictability in G-RNS since G-RNS is free from the bias caused by the violation of MR and G-RNS is uncorrelated with CFER.

We confirm our conjecture by means of three sets of empirical results. First, we repeat the portfolio sorting analysis, where we sort stocks in decile portfolios using ΔRNS as a sorting variable. Given the almost perfect rank correlation between ΔRNS and CFER (Table 2.4), we effectively sort stocks based on CFER, which HS document to have strong predictive power.[19] Table 2.10 shows the results. In line with the predictive ability of CFER, we can see that ΔRNS predicts stock returns. Moreover, its predictive power is stronger than that of O-RNS; t-statistics of the equally-weighted spread returns and alphas reach nine and those of the value-weighted ones are about four to five. Combined with the finding of the lack of the predictive power of G-RNS, this result means that the predictive power of O-RNS is "condensed" in the ΔRNS part.

[Table 2.10 about here.]

Next, we investigate the relation between the predictive power of O-RNS and the degree of the violation of MR. If the predictive power of O-RNS stems from the

[18] Note that the moments of the ex-dividend and cum-dividend *log* returns may differ due to the non-linearity of the logarithm function. However, this effect is negligible given a typical dividend-to-stock price ratio.

[19] HS document that the alphas of the CFER-sorted value-weighted decile spread portfolio are about 180 bps per month and their t-statistics are greater than five. This result suggests that the return predictive power of CFER is stronger than that of O-RNS reported in Table 2.5.

violation of MR, we expect that O-RNS does not show predictive power among stocks which do not severely violate MR (i.e., CFER is close to zero), while we expect that the predictive power of O-RNS will be stronger among stocks which severely violate MR. We examine these conjectures by means of a bivariate portfolio sort using the absolute value of CFER and O-RNS. We use the absolute value of CFER for controlling the degree of the violation of MR. In particular, we conduct a dependent bivariate portfolio sort, where we first sort stocks into quartile portfolios based on the absolute value of CFER. Then, within each one of four subgroups, we form quartile portfolios by sorting stocks based on O-RNS and we calculate the post-ranking average return of the spread portfolios (highest O-RNS minus lowest O-RNS).

Table 2.11 reports the average returns and alphas of the highest O-RNS minus lowest O-RNS spread portfolios. The results confirm our conjectures. The average returns and alphas are highest for the highest |CFER| bin and lowest for the lowest |CFER| bin. Therefore, O-RNS exhibits stronger return predictive power among stocks subject to severer violation of MR. Specifically, the results for the lowest |CFER| bin (the first and fifth Columns) indicate that O-RNS does not predict future stock returns among the lowest |CFER| bin; all returns and alphas but α_{SY} of the equally-weighted portfolio are insignificant at 5% level (α_{SY} of the equally-weighted portfolio is insignificant at a 1% level).[20] In line with our conjecture, this result suggests that O-RNS does not predict stock returns when MR is not (severely) violated.

[Table 2.11 about here.]

Finally, we provide direct evidence that the predictive power of O-RNS stems from CFER by employing the *CFER-adjusted return regression* proposed by HS. They show that their CFER-asset pricing model has the following testable implication for linear factor models; in the case where the IMRS m^* is described by a linear combination of risk factors $\boldsymbol{f}$, then the intercept α of the following regression of the CFER-adjusted excess return on risk factors,

$$R_{t,T} - R^f_{t,T} - CFER_{t,T} = \alpha + \boldsymbol{\beta}'\boldsymbol{f}_T + \varepsilon_T, \tag{2.48}$$

should be zero. We examine this hypothesis for each one of the CFER-adjusted

[20]We also conduct three-by-three and five-by-five double sorting exercises and we confirm that results do not change qualitatively.

returns of the O-RNS-sorted value-weighted decile portfolios and the associated long-short spread portfolio. If we fail to reject this hypothesis, this would suggest that the significant alphas of the CFER-*non*-adjusted returns of the O-RNS-sorted portfolios stem from the presence of market frictions, that is, a non-zero CFER.

Table 2.12 shows the results, where we regress the CFER-adjusted returns of each one of the O-RNS sorted value-weighted decile portfolios and the CFER-adjusted return of the long-short spread portfolio on a set of factors. In line with the positive correlation between CFER and O-RNS, the first row shows that the portfolio average CFER is monotonically increasing. However, the alphas (intercepts) of the CFER-adjusted excess returns of decile portfolios are almost always insignificant. In addition, the spread portfolio's alphas are now insignificant even though the non-CFER adjusted case (Table 2.5, Panel B) exhibits significant alphas. The fact that the CFER-adjustment removes the significant predictive power OF O-RNS implies that the predictive power of the O-RNS is driven by the CFER component, that is, the effect of market frictions on the expected returns, and it is not driven by (omitted) risk factors. This corroborates our previous finding that the predictability of the O-RNS stems from its correlation with CFER.

[Table 2.12 about here.]

2.6.4 Conjectures proposed by previous studies: Discussion

In this subsection, we discuss our contributions to the ongoing debate on the mechanism behind the predictive power of O-RNS, especially to the two main existing explanations reviewed in Section 2.2: the *limits-of-arbitrage* explanation and the *informed option trading* explanation.

First, our findings, especially those in Section 2.6.3, support the limits-of-arbitrage explanation as the mechanism behind the predictive ability of O-RNS. The predictive power of O-RNS stems from its estimation bias component caused by the violation of MR (non-zero CFER); O-RNS proxies CFER via its bias component and hence it signals the effect of market frictions on the expected stock returns. Note that our discussion is distinct from the existing limits-of-arbitrage explanation in the following sense. Our theoretical and empirical results clarify that O-RNS suffers from estimation biases when the MR is violated. This point is largely overlooked in the previous studies based on the limits-of-arbitrage explanation, even though this strand of studies in effect conjecture that MR may be violated.

Second, our CFER-based story can explain both the positive (higher O-RNS predicting future outperformance) and negative (lower O-RNS predicting future underperformance) informational contents in a unified framework. This is in contrast to the existing literature which finds it challenging to explain the positive and negative informational contents in a single type of market frictions. For example, the explanation based on the short-sale constraints employed by Stilger et al. (2017) can explain only the negative informational contents of O-RNS. whereas Gkionis et al.'s (2018) explanation based on the downside risk can explain only the positive informational contents of O-RNS. HS show that CFER subsumes the aggregate effect of any relevant market frictions on the expected stock returns including short-sale constraints, margin constraints and transaction costs among others. Consequently, the bias in the O-RNS which is determined by CFER signals the aggregate effect of various types of market frictions on the expected stock returns. In other words, our CFER-based story encompasses all explanations based on limits-of-arbitrage caused by a specific type of market frictions.

Third, we shortly document that the predictive power of O-RNS is more pertaining to the market frictions than the informed option trading explanation. We conduct a subsample analysis in which we divide stocks into those with option trading and those without option trading, then we repeat the portfolio sort exercise using O-RNS within each subsample. Under the informed option trading story, we expect that the non-traded option prices should have inferior informational contents (i.e., lower predictive power) compared to the traded options. This is because their prices (and hence RNS) would not reflect any effects from informed option trading. To examine this prediction, we obtain the daily aggregate option trading volume data from OM. Then, for each end-of-month trading day, we group stocks into two subgroups: stocks whose options have non-zero trading volume on that day and remaining zero option trading stocks.[21]

[21]We use the aggregate option trading volume across all strikes and maturities of both call and put as the trading activity indicator for several reasons. First, our option prices and IVs data come from the interpolated volatility surfaces and hence there are no trading volume data for each option price and IV. Second, OM estimate volatility surfaces using all strikes and maturities. Therefore, the aggregate trading volume is a suitable characteristic to measure the overall estimation quality of

Table 2.13 reports the results. We can see that there are qualitatively no differences between the traded options subgroup and non-traded options subgroup; the average return and the alphas of the O-RNS-sorted spread portfolio are generally significant, whereas those of the G-RNS-sorted spread portfolio are generally insignificant. The informed option trading and option trading demand stories cannot fully explain this estimated pattern; we cannot find visible differences between the informativeness of options with non-zero trading volume and zero trading volume. On the other hand, the friction-based story can explain the predictive power among non-option trading stocks, because our untabulated result indicates that these stocks tend to be smaller stocks which face a considerable degree of market frictions. This finding is in line with Goncalves-Pinto et al. (2019), who show that their option-based return predictor, DOTS, does not exhibit the different level of predictability among stocks with zero and non-zero option trading volume. Since the OM IV data is calculated based on the option bid and ask quotes, the findings in this Chapter and Goncalves-Pinto et al. (2019) imply that bid and ask quotes set by option market-makers contain similar informational content regardless of the trading demand from option end-users.

[Table 2.13 about here.]

Two more remarks are in order regarding the conjectures for the mechanism of the predictive ability of O-RNS. First, admittedly, the validity of the limits-of-arbitrage story for the O-RNS predictive power mechanism is not accepted universally. For example, Chordia et al. (2019) conduct depending bivariate portfolio sort exercises, where they first control for either Amihud's (2002) illiquidity measure or the idiosyncratic volatility (IVOL) and then sorting stocks based on O-RNS. They find that O-RNS retains predictive power even in the lowest Amihud and IVOL bins, that is, among stocks which face low degree of the limits-of-arbitrage. From this result, they conclude that the degree of the limits-of-arbitrage cannot fully account for the RNS-return relation. However, our result in Table 2.11 suggests the opposite story. Our bivariate portfolio sort analysis there resembles those in Chordia et al. (2019), except that we use the absolute value of CFER to control for the degree of the limits-of-arbitrage. In our analysis, we find that O-RNS does not predict stock returns in

volatility surfaces. Finally, we focus on a short-term horizon and hence the aggregate trading volume is a good proxy of the horizon of our interest, since short-term options are relatively heavily traded.

the lowest limits-of-arbitrage bin. We obtain a different result from Chordia et al. (2019) and corroborate the limits-of-arbitrage story because we use a more direct and theoretically-founded proxy of the degree of the limits-of-arbitrage.

Second, our results in Table 2.13, albeit supportive for the limits-of-arbitrage story, do not exclude the possibility that other empirical predictive patterns are well explained by the informed option trading. Nevertheless, the following discussion suggests that the existence of market frictions is still necessary to justify the informed option trading story. To begin with, the theoretical model of Easley et al. (1998) predicts that informed traders trade in the option market if the leverage effect of options is large enough (implying that there is leverage and margin constraints), or the stock market liquidity is low; their model implies that informed traders trade options only when there are sufficiently big frictions to trade stocks (relative to trade options). On the contrary, informed option trading activity would be quickly reflected in the underlying price through the put-call parity relation *unless* there are market frictions which prevent arbitrageurs from trading the underlying stock to exploit deviations from put-call parity. Deviations from put-call parity is an essential element to justify the predictive power of O-RNS. This is because we estimate CFER based on the scaled deviations from put-call parity as proposed by HS, and non-zero CFER (i.e., the violation of MR) drives the predictive power of O-RNS.

2.6.5 Predictive power of implied volatility slope

We have seen in Section 2.4.4 that the IV slope calculated based on the standard IV does not correctly proxy RNS under the violation of MR. We interpret the XZZ IV slope measure calculated based on the robust IV curve as the counterpart of G-RNS in the sense that both the robust IV and G-RNS take the possible violation of MR by the underlying into account. Specifically, our discussion in Section 2.4.4 shows that the robust IV-based XZZ measure is not affected by non-zero IVS and the zero slope of the robust IV curve implies that the underlying distribution is log-normal. In the rest of this subsection, we document that the predictive power of the XZZ measure vanishes once it is calculated based on the robust IVs, corresponding to our finding that G-RNS does not have the return predictive power.

We start with reporting the summary statistics of the XZZ slope measures. Table

2.14, Panel A, reports the summary statistics of CFER, the original and robust XZZ measure estimated at each end of month, and their difference $\Delta XZZ = XZZ^o - XZZ^r$. We can see that CFER and the XZZ measures are calculated for about 2,200 stocks on average on each end-of-month trading day. The mean and median of the two XZZ measures are both positive, implying that the IV curve exhibits a negative skew more often than a positive skew, regardless of using the original IV or the robust IVs. Table 2.14, Panel B, reports the pairwise Pearson and Spearman rank correlations. Although the correlation between the original and robust XZZ measures is fairly high (0.78 and 0.8), we will show that the outcome of the portfolio sort analysis crucially depends on the choice of the XZZ measure. CFER and the original XZZ measure are negatively correlated, while CFER and the robust XZZ measure do not show strong correlation. These patterns are in line with our discussion in Section 2.4.4; the original XZZ measure is mechanically negatively correlated with IVS and CFER (equations (2.45) and (2.46)), whereas the robust XZZ measure is not affected by IVS because the robust IV does not exhibit IVS.

[Table 2.14 about here.]

Next, we sort stocks in ascending order based on either XZZ^o and XZZ^r separately, and form value-weighted decile portfolios. Then, we construct the long-short spread portfolios where we go long in the stocks with the highest measures and short in the stocks with the lowest measures. Table 2.15 reports the result. First, consistent with Xing et al. (2010), the original XZZ measure negatively predicts future stock returns; the decile portfolio returns are generally decreasing in the level of the XZZ measure. The average return and alphas of the long-short spread portfolio are highly significant; t-statistics are above 5.34 (3.89) in magnitude for the equally-weighted portfolio (value-weighted portfolio). On the other hand, XZZ^r does not predict stock returns. The average returns and alphas are all insignificant at a 5%-significance level for the value-weighted portfolio and mostly insignificant for the equally-weighted portfolio.

[Table 2.15 about here.]

We repeat the CFER-adjusted regression considered in Table 2.12 for the XZZ-sorted portfolios to see whether the predictive power of XZZ^o stems from the CFER channel. Table 2.16 reports the result, where we regress the CFER-adjusted returns of the XZZ^o-sorted value-weighted decile portfolios and their long-short spread portfolio. First, the average portfolio CFER of decile portfolios are monotonically decreasing in the level of the XZZ^o in line with its mechanical relation to CFER, equation (2.45). Then, the alpha (intercept) of the CFER-adjusted excess returns are overall insignificant even though the non-CFER adjusted case (Table 2.15, Panel B) exhibits significant alphas. Similar to the O-RNS case, the switch from the significant alphas to insignificant alphas implies that the predictive power of the XZZ^o is driven by the CFER component.

[Table 2.16 about here.]

Two final remarks are in order to summarize our findings in this Section. First, option-implied skewness measures lose their return predictive power once we account for the possible violation of MR in the estimation of the option-implied measures. This suggests that the (more) correctly measured RNS does not contain information on the future stock returns. Second, our results suggest that O-RNS and the XZZ slope measure do not contain additional information over CFER. In particular, the return predictive power of O-RNS and XZZ^o stems from their mechanical correlation with CFER.

2.7 Conclusions

The violation of the martingale restriction (MR), that is, the deviation of the expected risk-neutral return from the risk-free rate, implies the presence of market frictions. Surprisingly, the implications of the violation of MR for the computation of option-implied qualities has not been explored.

We propose a novel way to test MR and explore the implications of the violation of MR for the risk-neutral moments (RNMs) under a unified setting. Our approach to test MR is based on the Hiraki and Skiadopoulos (2019) (HS) asset pricing model with market frictions. We show that the risk-neutral expected excess asset return equals the *contribution of frictions to expected returns* (CFER) and hence we can test MR by testing whether CFER is zero (i.e., the risk-neutral expected excess return is zero).

Since HS show that CFER can be reliably estimated from a properly scaled deviations from put-call parity, this CFER-based approach provides a robust way to test MR. On the other hand, we show that the implied stock price approach for testing MR proposed by Longstaff (1995) may spuriously reject MR even the underlying satisfies MR. We apply our testing methodology to the S&P 500 and the U.S. individual equities and find that both the index and individual stocks frequently violate MR. This finding validates our motivation to theoretically and empirically investigate how the violation of MR results in biases in the estimated RNMs.

We theoretically show that, under the violation of MR, the Bakshi et al. (2003) (BKM) formulae do not correctly estimate RNMs due to their implicit assumption that the underlying asset satisfies MR. To remedy this drawback, we propose generalized formulae to estimate RNMs which account for the possible violation of MR. Our empirical analysis shows that the difference between the estimated RNS based on the original BKM formula (O-RNS) and our generalized BKM formula (G-RNS) is economically significant; O-RNS predicts the cross-section of future stock returns in line with the literature (e.g., Stilger et al., 2017), whereas G-RNS does not. Furthermore, we document that the predictive power of O-RNS stems from its estimation bias caused by the violation of MR.

While we focused on the estimation of RNMs, especially RNS, our theoretical result on testing MR is relevant to any studies on the informational content of option prices which relies on the martingale assumption, including the estimation of risk-neutral distributions, pricing kernels, risk-aversion parameters, various portfolio allocation applications among others. These topics are best left for future research.

2.A Proofs

2.A.1 Proof of Lemma 2.3.1

By applying the change of measure formula to equation (2.8) and substituting the definition of CFER, equation (2.10), we obtain

$$S_t = \frac{1}{R^f_{t,T}} \mathbb{E}^{\mathbb{Q}^*}_t [S_T + \widetilde{D}_{t,T}] - \frac{S_t}{R^f_{t,T}} CFER_{t,T}. \tag{2.A.1}$$

Multiplying $R^f_{t,T}/S_t$ to the both sides proves equation (2.11). □

2.A.2 Proof of Proposition 2.3.1

Let m and σ be the mean and the volatility (annualized standard deviation) parameters of the true log-normal distribution, respectively, that is, $S_T \sim \text{Lognormal}(m, \sigma^2\tau)$. On the other hand, the BS benchmark model assumption suggests that the benchmark model-based distribution of S_T is given by

$$\text{Lognormal}\left(\log(S^*_t - e^{-r_f\tau}\widetilde{D}_{t,T}) + r_f\tau - \frac{(\sigma^*)^2\tau}{2}, (\sigma^*)^2\tau\right). \tag{2.A.2}$$

When there are two or more observed option prices in the minimization problem, equation (2.14), the two parameters of the true distribution, m and σ, can be correctly identified and satisfies

$$m = \log(S^*_t - e^{-r_f\tau}\widetilde{D}_{t,T}) + r_f\tau - \frac{(\sigma^*)^2\tau}{2} \tag{2.A.3}$$

and $\sigma = \sigma^*$.

In this case, the expected stock price under the true measure can be calculated as

$$\mathbb{E}^{\mathbb{Q}^*}_t[S_T] = \exp\left(m + \frac{\sigma^2\tau}{2}\right) = e^{r_f\tau}S^*_t - \widetilde{D}_{t,T} \tag{2.A.4}$$

due to equation (2.A.3). Equation (2.A.4) means that $S^*_t = e^{-r_f\tau}\mathbb{E}^{\mathbb{Q}^*}_t[S_T + \widetilde{D}_{t,T}]$ and hence the first term in the right-hand side of equation (2.16) equals zero. This proves equation (2.17). □

2.A.3 Proof of Proposition 2.4.1

Recall the following result by Carr and Madan (2001) that, for an arbitrary positive number F, any twice differentiable payoff function $f(S_T)$ satisfies

$$f(S_T) = f(F) + f'(F)(S_T - F) + \int_F^\infty f''(K)(S_T - K)^+ dK + \int_0^F f''(K)(K - S_T)^+ dK. \tag{2.A.5}$$

The power-n log return contract functions satisfy $f(S_t - \widetilde{D}_{t,T}) = f'(S_t - \widetilde{D}_{t,T}) = 0$. Hence, for these payoff functions, it follows that

$$f(S_T) = \int_{S_t - \widetilde{D}_{t,T}}^{\infty} f''(K)(S_T - K)^+ dK + \int_0^{S_t - \widetilde{D}_{t,T}} f''(K)(K - S_T)^+ dK. \quad (2.A.6)$$

Under the assumption that option prices satisfy MR, we obtain the following equation by multiplying both sides by $e^{-r_f\tau}$ and taking the risk-neutral expectation:

$$e^{-r_f\tau}\mathbb{E}_t^{\mathbb{Q}^*}[f(S_T)] = \int_{S_t - \widetilde{D}_{t,T}}^{\infty} f''(K)C_t(K,T)dK + \int_0^{S_t - \widetilde{D}_{t,T}} f''(K)P_t(K,T)dK. \quad (2.A.7)$$

By calculating the second derivatives for $f(S_T) = \log(S_T + \widetilde{D}_{t,T}/S_t)^n$, we obtain equations (2.26) and (2.27).

Next, note that $e^{r_f\tau} + CFER_{t,T} = \mathbb{E}_t^{\mathbb{Q}^*}[R_{t,T}] = \mathbb{E}_t^{\mathbb{Q}^*}[\exp(r_{t,T})]$ holds. Therefore, by expanding the last expression by the forth-order Taylor series approximation, we obtain

$$e^{r_f\tau} + CFER_{t,T} = 1 + \mathbb{E}_t^{\mathbb{Q}^*}[r_{t,T}] + \frac{\mathbb{E}_t^{\mathbb{Q}^*}[(r_{t,T})^2]}{2} + \frac{\mathbb{E}_t^{\mathbb{Q}^*}[(r_{t,T})^3]}{6} + \frac{\mathbb{E}_t^{\mathbb{Q}^*}[(r_{t,T})^4]}{24} + o((r_{t,T})^4). \quad (2.A.8)$$

In line with BKM, we rearrange and ignore the higher-order residual term to obtain

$$\widetilde{\mu}_{t,T} := \mathbb{E}_t^{\mathbb{Q}^*}[r_{t,T}] = e^{r_f\tau} + CFER_{t,T} - 1 - \frac{e^{r_f\tau}}{2}M(2)_{t,T} - \frac{e^{r_f\tau}}{6}M(3)_{t,T} - \frac{e^{r_f\tau}}{24}M(4)_{t,T}. \quad (2.A.9)$$

This proves equation (2.25). Then, equations (2.22) to (2.24) are obvious. □

2.A.4 Proof of Proposition 2.4.2

Since the payoff functions change to $f(S_T) = \log(S_T/S_t)^n$ for the ex-dividend case, we now have $f(S_t) = f'(S_t) = 0$. Therefore, the boundary value for the two integrals in equations (2.A.6) and (2.A.7) changes from $S_t - \widetilde{D}_{t,T}$ to S_t. Moreover, by calculating the second derivative of the ex-dividend power-n log return functions, we obtain the expressions (2.28) and (2.29). Next, note that the following relation holds.

$$\mathbb{E}_t^{\mathbb{Q}^*}[\exp(r_{t,T}^{ex})] = \mathbb{E}_t^{\mathbb{Q}^*}\left[\frac{S_T}{S_t}\right] = \mathbb{E}_t^{\mathbb{Q}^*}[R_{t,T}] - \frac{\widetilde{D}_{t,T}}{S_t} = e^{r_f\tau} + CFER_{t,T} - \frac{\widetilde{D}_{t,T}}{S_t}. \quad (2.A.10)$$

By expanding the left hand side of equation (2.A.10) by a Taylor series approximation similar to equation (2.A.8) and rearranging, we obtain equation (2.30). □

2.A.5 Proof of Proposition 2.4.3

First, we prove (a) for the call option case. Let $c(S_t, K, \tau, \sigma, r_f, \widetilde{D}_{t,T}, r_S)$ be the call pricing function, equation (2.34), under the distribution (2.32). Our purpose is to find the expression of $IV^c(K; r_S)$ which solves

$$c(S_t, K, \tau, \sigma, r_f, \widetilde{D}_{t,T}, r_S) = BS_{call}(S_t, K, \tau, r_f, \widetilde{D}_{t,T}, IV^c(K; r_S)).$$

In what follows, we suppress the arguments of the c function and IV^c other than r_S.

The first-order Taylor series approximation of the (standard) BS call function with respect to the volatility parameter around the true volatility parameter σ yields

$$BS_{call}(S_t, K, \tau, r_f, \widetilde{D}_{t,T}, IV^c(r_S)) \approx BS_{call}(S_t, K, \tau, r_f, \widetilde{D}_{t,T}, \sigma) + \mathcal{V}_{BS}(IV^c(r_S) - \sigma), \tag{2.A.11}$$

where $\mathcal{V}_{BS}$ is the BS vega evaluated at $(S_t, K, \tau, r_f, \widetilde{D}_{t,T}, \sigma)$. On the other hand, the first-order Taylor series approximation of the c function with respect to the r_S-argument around $r_S = r_f$ yields

$$c(r_S) \approx c(r_f) + \frac{\partial c}{\partial r_S}\bigg|_{r_S = r_f}(r_S - r_f). \tag{2.A.12}$$

The left hand side of equations (2.A.11) and (2.A.12) are the same by the definition of the IV, $c(r_S) = C_t(K, T) = BS_{call}(S_t, K, \tau, r_f, \widetilde{D}_{t,T}, IV^c(r_S))$. Moreover, the first term in the right-hand side of equations (2.A.11) and (2.A.12) are also the same because $c(r_f) = BS_{call}(S_t, K, \tau, r_f, \widetilde{D}_{t,T}, \sigma)$ follows from equation (2.34). Therefore, equations (2.A.11) and (2.A.12) yield

$$IV^c(r_s) \approx \sigma + \frac{r_S - r_f}{\mathcal{V}_{BS}} \frac{\partial c}{\partial r_S}\bigg|_{r_S = r_f}. \tag{2.A.13}$$

A straightforward calculation of the partial derivative $\partial c/\partial r_s$ yields

$$\begin{aligned}\frac{\partial c}{\partial r_S} &= e^{(r_S-r_f)\tau}\left(\tau BS_{call}(S_t,K,\tau,r_S,\widetilde{D}_{t,T},\sigma)+\frac{\partial BS_{call}(S_t,K,\tau,r,\widetilde{D}_{t,T},\sigma)}{\partial r}\Big|_{r=r_S}\right)\\ &= e^{(r_S-r_f)\tau}\tau S_t\Phi(d_1),\end{aligned} \tag{2.A.14}$$

because the *rho* of the BS function with discrete dividends evaluated at $r = r_S$ is given by $\tau(S_t\Phi(d_1) - BS_{call}(S_t, K, \tau, r_S, \widetilde{D}_{t,T}, \sigma))$. Therefore, equation (2.A.13) reduces to

$$IV^c(r_S) \approx \sigma + (r_S - r_f)\tau\frac{S_t\Phi(d_1)}{\mathcal{V}_{BS}} = \sigma + (r_S - r_f)\tau\frac{S_t}{S_t - e^{-r_f\tau}\widetilde{D}_{t,T}}\frac{\Phi(d_1)}{\sqrt{\tau}\phi(d_1)}, \tag{2.A.15}$$

because the BS vega is given by $\mathcal{V}_{BS} = \left(S_t - e^{-r_f\tau}\widetilde{D}_{t,T}\right)\phi(d_1)\sqrt{\tau}$. Since

$$CFER_{t,T}/R^f_{t,T} = \mathbb{E}^{\mathbb{Q}}_t[R_{t,T}]/R^f_{t,T} - 1 = e^{(r_S-r_f)\tau} - 1 \approx (r_S - r_f)\tau$$

holds, we prove equation (2.39). A similar calculation for the put option proves equation (2.40).

Next, we prove (b). The statement about the inequality relation between σ and IV^c or IV^p is obvious because $\Phi(x)/\phi(x)$ is always positive. To prove the remaining claim, it suffices to show $\Phi(x)/\phi(x)$ is increasing in x. Since the derivative $(\Phi(x)/\phi(x))'$ is given by

$$\frac{\phi^2(x) - \Phi(x)\phi'(x)}{\phi^2(x)} = \frac{\phi^2(x) + x\Phi(x)\phi(x)}{\phi^2(x)} = \frac{\phi(x) + x\Phi(x)}{\phi(x)}, \tag{2.A.16}$$

it suffices to show $\phi(x) + x\Phi(x) > 0$ for any x. This is obvious for $x \geq 0$. For $x < 0$, we prove $x > -\phi(x)/\Phi(x)$ instead. In this case, $\phi(x)/\Phi(x)$ (the inverse Mills ratio of the standard normal distribution) is known to be equal to the negative of the truncated mean $-\mathbb{E}[Z|Z < x]$, where Z is a standard normal random variable (see e.g., Theorem 22.2 of Greene, 2003). Then, $x > \mathbb{E}[Z|Z < x]$ is obvious. □

2.A.6 Proof of Proposition 2.4.4

The observed put and call prices satisfy the following relation.

$$
\begin{aligned}
& e^{-(r_S-r_f)\tau}C_t(K,T) - e^{-(r_S-r_f)\tau}P_t(K,T) = e^{-(r_S-r_f)\tau}e^{-r_f\tau}\mathbb{E}_t^{\mathbb{Q}}[S_T-K] \\
&= e^{-r_S\tau}\mathbb{E}_t^{\mathbb{Q}}[S_T-K] = e^{-r_S\tau}S_t\mathbb{E}_t^{\mathbb{Q}}[R_{t,T}] - e^{-r_S\tau}\widetilde{D}_{t,T} - e^{-r_S\tau}K \\
&= (S_t - e^{-r_S\tau}\widetilde{D}_{t,T}) - e^{-r_S\tau}K.
\end{aligned}
\tag{2.A.17}
$$

Therefore, it follows that $IV^c_{rob} = IV^p_{rob} = IV_{rob}$ if and only if

$$
\begin{aligned}
& BS_{call}(S_t, K, \tau, r_S, \widetilde{D}_{t,T}, IV_{rob}) - BS_{put}(S_t, K, \tau, r_S, \widetilde{D}_{t,T}, IV_{rob}) \\
&= (S_t - e^{-r_S\tau}\widetilde{D}_{t,T}) - e^{-r_S\tau}K.
\end{aligned}
\tag{2.A.18}
$$

Equation (2.A.18) holds because of the property of the BS formulae with deterministic dividends. □

2.A.7 Proof of Lemma 2.4.1

For simplicity, we assume that the stock pays no dividends. The weighting function $\eta(K; S_t, n)$ in equation (2.26) equals

$$
\eta(K; S_t, n) = \frac{n(\log(K/S_t))^{n-2}}{K^2}\left[(n-1) - \log(K/S_t)\right]. \tag{2.A.19}
$$

For the strike $K/S_t > e^{n-1}$, equation (2.A.19) is negative regardless of whether n is odd or even. This proves the respective last inequalities in equations (2.43) and (2.44).

For the strike $K/S_t < e^{n-1}$, the square bracket term in equation (2.A.19) is positive. Therefore, the sign of the η function equals the sign of $(\log(K/S_t))^{n-2}$. This term is always non-negative when n is an even number, proving equation (2.44). On the other hand, when n is an odd number, this term is non-negative for $K/S_t \geq 1$ and negative for $K/S_t < 1$. This proves equation (2.43). □

2.B Description of dataset and variables

In this Appendix, we provide a detailed explanation for the data sources and the calculation procedures for the characteristic variables, CFER, and RNMs. We also explain how we link the three databases we use: OM, CRSP and Compustat.

2.B.1 Firms' and stocks' characteristic variables

Beta: In each month, we regress daily stock excess returns over past 12 months on the daily excess market return to obtain the beta. We require there are at least 200 non-missing observations. Stock return data are obtained from the CRSP database. We use the excess market return provided at Kenneth French's website.

SIZE: Size is the natural logarithm of the market equity. The market equity is calculated as the product of the number of outstanding share with the price of the stock at the end of each month. Data are obtained from the CRSP database.

Book-to-Market equity (B/M): We follow Davis et al. (2000) to measure book equity as stockholders' book equity, plus balance sheet deferred taxes and investment tax credit (Compustat annual item TXDITC) if available, minus the book value of preferred stock. Stockholders' equity is the value reported by Compustat (item SEQ), if it is available. If not, we measure stockholders' equity as the book value of common equity (item CEQ) plus the par value of preferred stock (item PSTK), or the book value of assets (item AT) minus total liabilities (item LT). Depending on availability, we use redemption (item PSTKRV), liquidating (item PSTKL), or par value (item PSTK) for the book value of preferred stock. From June of each year t to May of $t+1$, the book-to-market equity (B/M) is calculated as the ratio of the book equity for the fiscal year ending in calendar year $t-1$ to the market equity at the end of December of year $t-1$. We treat non-positive B/M data as missing.

Profitability: We follow Fama and French (2015) to measure profitability as revenues (Compustat annual item REVT) minus cost of goods sold (item COGS) if available, minus selling, general, and administrative expenses (item XSGA) if available, minus interest expense (item XINT) if available all divided by (non-lagged) book equity. From June of year t to May of $t+1$, we assign profitability for the fiscal year ending in calendar year $t-1$.

Investment: We follow Fama and French (2015) to measure investment as the change in total assets (Compustat annual item AT) from the fiscal year ending in year $t-1$ to the fiscal year ending in t, divided by $t-1$ total assets. From June of year t to May of $t+1$, we assign investment for the fiscal year ending in calendar year $t-1$.

2.B.2 Linking OM, CRSP, and Compustat databases

In our analysis, we use three databases, OM, CRSP, and Compustat. To link CRSP and Compustat, we use the SAS linking macro provided by the Wharton Research Data Services (WRDS) and match PERMNO (the CRSP identifier) and GVKEY (the Compustat identifier). To link CRSP and OM, we link SECID (the OM identifier) and PERMNO by the SAS linking macro provided by the WRDS. This WRDS macro links SECID and PERMNO based on the CUSIP code, the similarity in the issuer's (company) name recorded in OM and CRSP, and partly the security ticker. The macro also provides the score of the strength of the link based on a 0–6 integer scale (0 stands for the strongest link, where the eight-digit CUSIP exactly matches as well as the issuer's name matches). We keep SECID-PERMNO link with the score less than or equal to three. After linking SECID, GVKEY and PERMNO, we keep only U.S. common stocks (SHRCD 10 or 11) traded in NYSE/Amex/NASDAQ (EXCHCD 1, 2, or 3).

2.B.3 Estimation of CFER

HS show that the underlying stocks' CFER satisfies the following approximate relation:

$$CFER_{t,T} \approx \frac{R^0_{t,T}}{S_t}(\widetilde{S}_{t,T}(K) - S_t), \tag{2.B.20}$$

where $\widetilde{S}_{t,T}(K)$ is the *synthetic stock* price defined as

$$\widetilde{S}_{t,T}(K) = C_t(K,T) - P_t(K,T) + \frac{K + \widetilde{D}_{t,T}}{R^f_{t,T}}. \tag{2.B.21}$$

Their result, equation (2.B.20), shows that CFER is accurately proxied by the observable *scaled* deviations from put-call parity.

In the actual estimation process of CFER, we use an alternative equivalent expression of $\widetilde{S}_{t,T} - S_t$ in the right-hand side of equation (2.B.20).

$$\widetilde{S}_{t,T}(K) - S_t = BS_{call}(S_t, K, \tau, r, \widetilde{D}_{t,T}, IV_t^c(K)) - BS_{call}(S_t, K, \tau, r, \widetilde{D}_{t,T}, IV_t^p(K)), \tag{2.B.22}$$

where BS_{call} is the Black-Scholes European call option function with deterministic dividend payment (equation (2.19)) and $IV_t^c(K)$ ($IV_t^p(K)$) is the IV of the call (put) option. We estimate the right-hand side of equation (2.B.20) by calculating deviations from put-call parity based on equation (2.B.22).

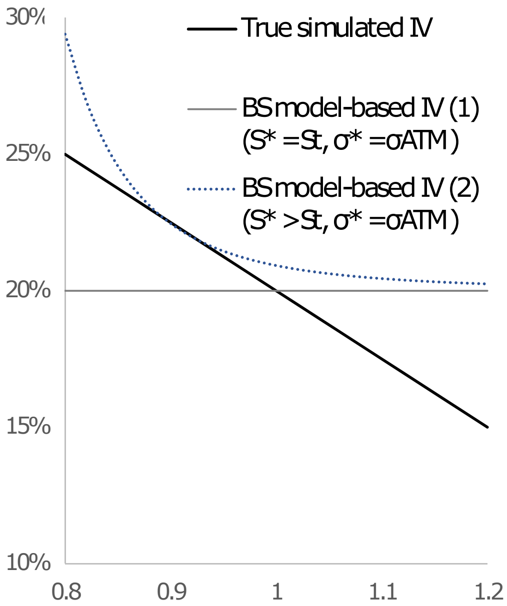
We use the 30-day forward ATM call and put IVs provided by the OM SO file for the calculation of equation (2.B.22), that is, $K = F_{t,T} = e^{r_f \tau} S_t - \widetilde{D}_{t,T}$ and $\tau = 30$ days. The underlying stock price S_t and the risk-free rate r_f are also obtained from the OM database. We interpolate the OM zero yield data to obtain the risk-free rate for the specified day-to-maturity. Regarding the dividend payment $\widetilde{D}_{t,T}$, we back it out from the forward price in the OM SO file as $\widetilde{D}_{t,T} = e^{r_f \tau} S_t - F_{t,T}$.

2.B.4 Estimation of risk-neutral moments

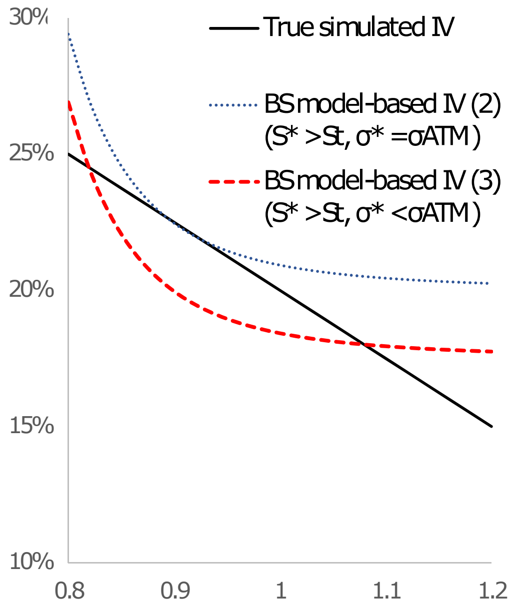
We estimate risk-neutral moments (RNMs) based on the formulae shown in Section 2.4. The option price data across a wide range of strike comes from the OM VS file. The OM VS file contains the smoothed IVs for standardized maturities and deltas. In particular, it contains IVs and associated option prices and strike prices at delta equals 0.2, 0.25, . . . , 0.8 for call options and -0.8, -0.75, . . . , -0.2 for put options for the day-to-maturity equals 30, 60, 91, 122, 152, 182, 273, 365, 547 and 730 calendar days. We use 30-day-to-maturity data for the purpose of estimating risk-neutral moments. Therefore, this file allows us to use call and put option data for 13 different strikes for standardized time-to-maturities. The OM volatility surface file provides IVs, implied strike prices, implied European option prices at each standardized delta for call and put options, separately. We discard day-stock observations and hence do not estimate the RNMs when either the strike prices provided by the OM of calls or puts are not monotonic in deltas. Such a situation may occur when the interpolated volatility surface exhibits an extremely steep skew.

We follow Stilger et al. (2017) to interpolate and extrapolate IVs as follows. We interpolate call IVs and put IVs separately in the strike-IV dimension based on the cubic piecewise Hermite polynomial interpolation. Then, we extrapolate horizontally the IV beyond the highest and the lowest strikes. For the estimation of O-RNMs, we directly interpolate the IVs provided by OM. For the estimation of G-RNMs, first we calculate the robust IVs from the implied option prices provided by OM and interpolate these robust IVs. Once the continuous IV curve is obtained, we calculate option prices at 1001 equally-spaced strike points over the moneyness from 1/3 to 3. Finally, we estimate O-RNMs and G-RNMs by numerically calculating the integrals in the original BKM formulae and Proposition 2.4.1, respectively.

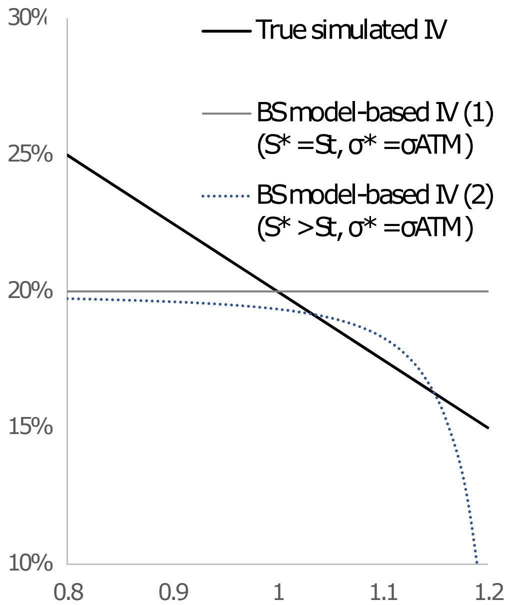
The usage of the OM VS file has several advantageous points. First, it provides the IVs at a standardized maturity and for a unified strike range (0.2 to 0.8 deltas in the absolute value) for all stocks. This helps us to avoid any issues arising from mixing estimated RNMs from different maturities and different moneyness ranges covered by traded options. Second, this dataset has been used in the estimation of RNMs (e.g., DeMiguel et al., 2013; Borochin and Zhao, 2018; Chordia et al., 2019), thus making possible the comparison of our results to theirs.



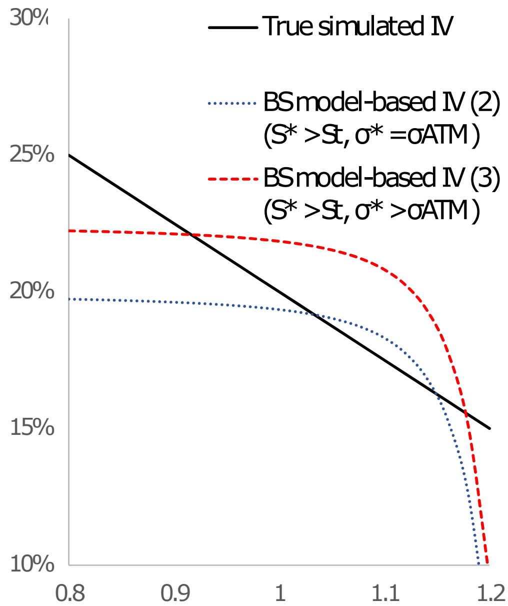
(a) Call IVs (1): the effect of a higher S_t^*



(b) Call IVs (2): the effect of a lower σ^*



(c) Put IVs (1): the effect of a higher S_t^*



(d) Put IVs (2): the effect of a higher σ^*

Figure 2.1. Implied stock price approach: Schematic graphs when the true BS-IV curve is negatively skewed

These four Figures schematically illustrate how the BS model-based implied stock price approach chooses the implied stock price S_t^* and the BS volatility parameter σ^*. The black thick line in each Figure depicts the true simulated BS-IV curve, which is given as a function of moneyness $IV(K/S_t) = \sigma_{ATM} - (K/S_t - 1)/4$. The BS model-based implied stock price approach chooses S_t^* and σ^* so that the BS-IV curve of the BS model-based option prices is the closest to the given true simulated BS-IV curve. Figures (a) and (c) depict the effect of choosing a higher value for S_t^* (from the gray thin line to the blue dotted line), for the call option and put option cases, respectively. Figures (b) and (d) depict the effect of choosing a lower (higher) value of σ^* (from the blue dotted line to the red dashed line) for the call (put) option case, respectively. The x-axis is moneyness and the y-axis is the IV in percent.

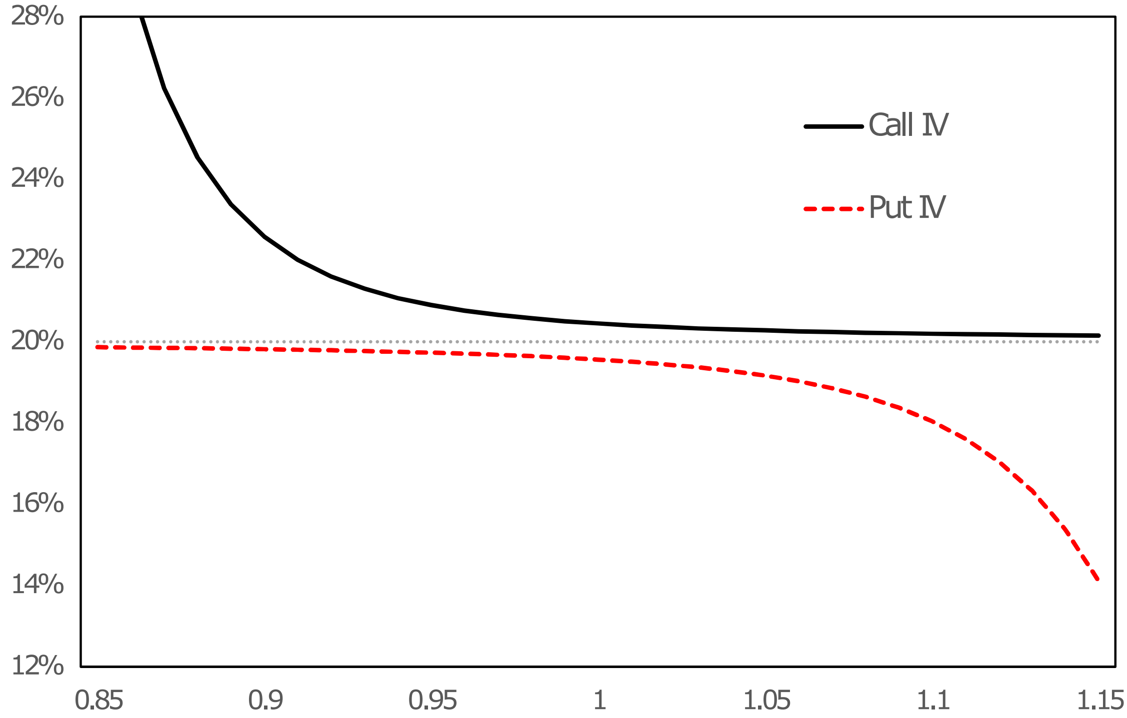


Figure 2.2. Implied volatility and the violation of MR

This Figure depicts the call and put BS-IV curves based on the equations (2.39) and (2.40), respectively. The time-to-maturity is $\tau = 1/8$, the risk-free rate is $r_f = 3\%$, the volatility parameter is $\sigma = 20\%$, the deterministic dividend-to-stock price is $\widetilde{D}_{t,T}/S_t = 0.5\%$ and the annualized CFER is $CFER_{t,T}/\tau = 1\%$. The x-axis is the moneyness and the y-axis is the implied volatility in percentage.

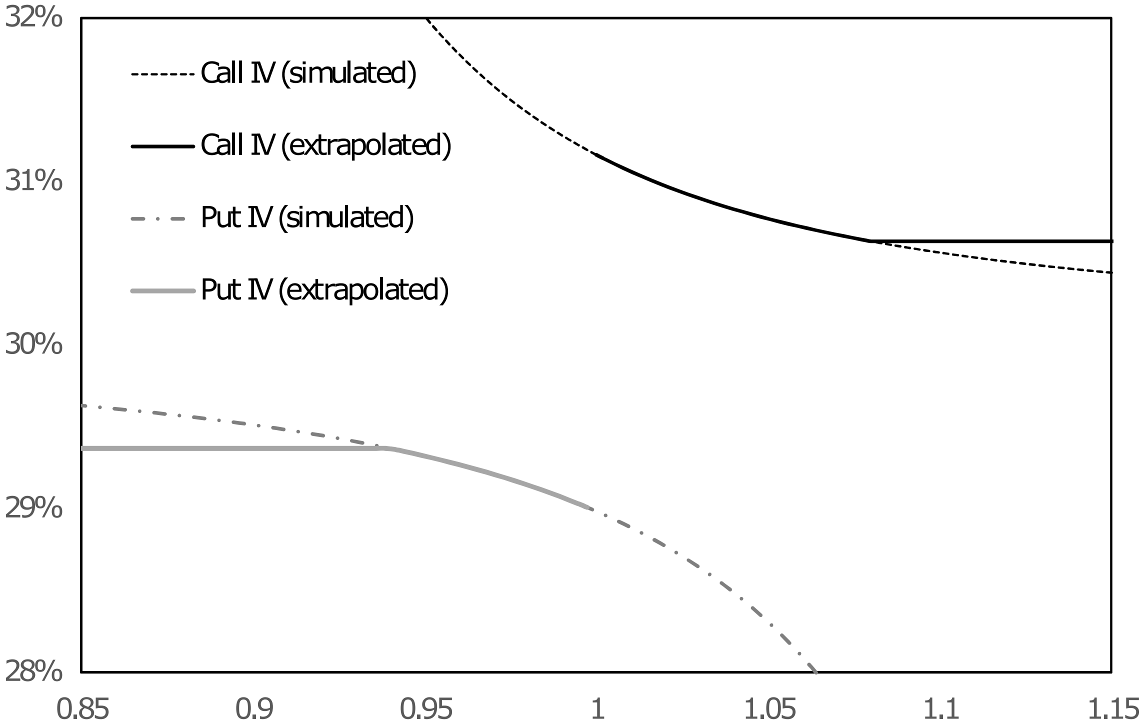
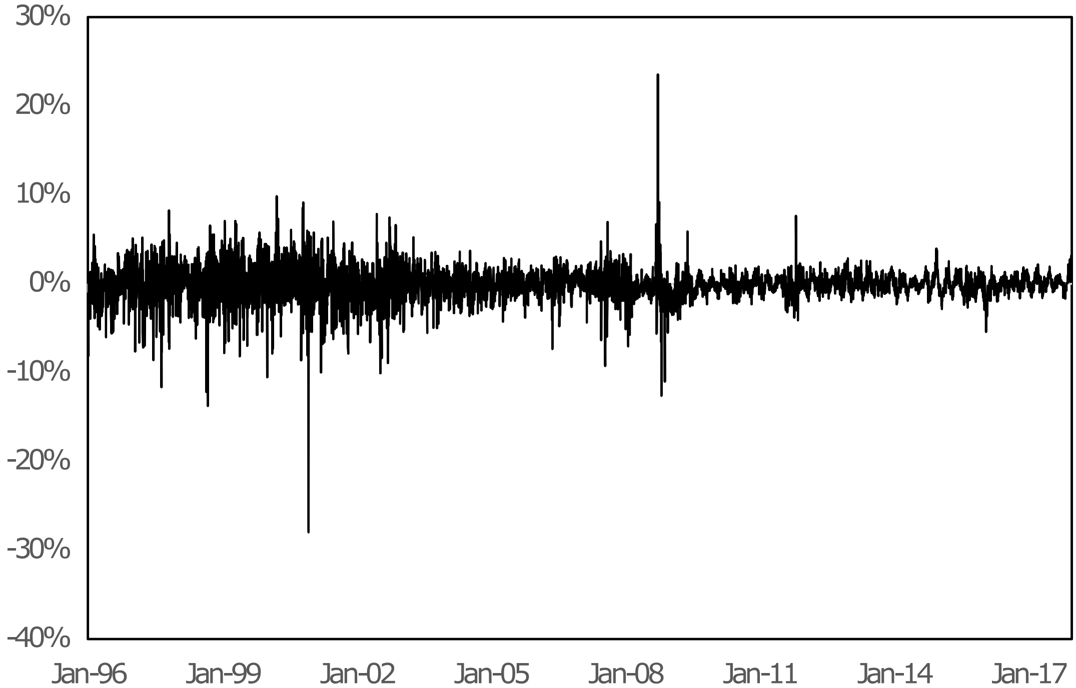
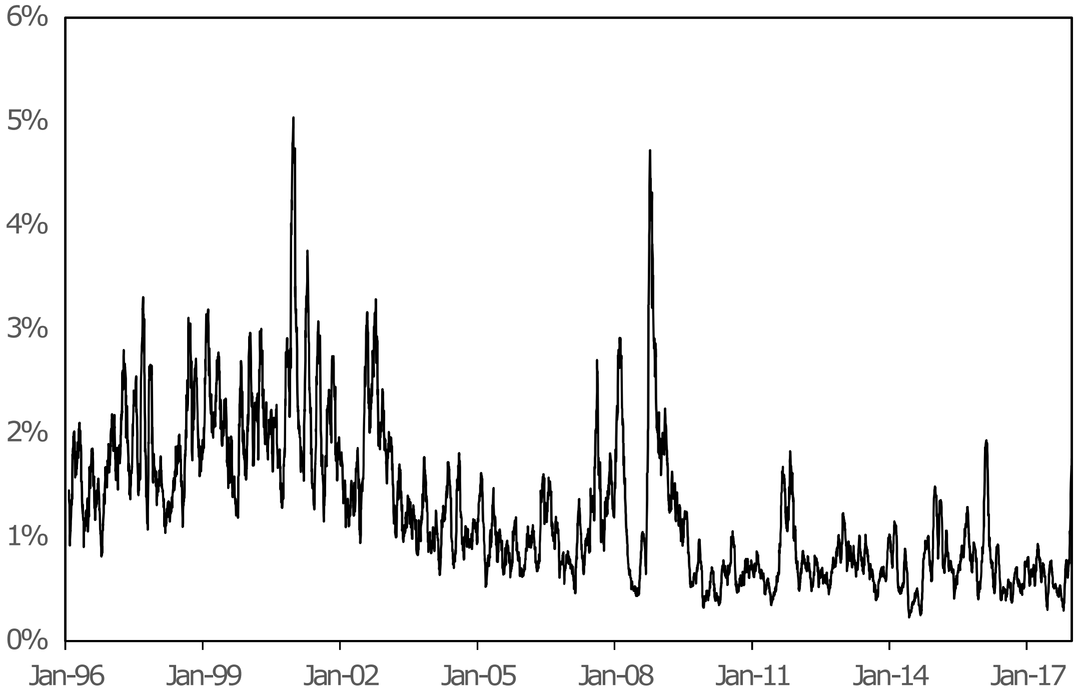


Figure 2.3. Horizontally extrapolated standard IVs in the RNM estimation

This Figure depicts the call and put standard IV curves based on the equations (2.39) and (2.40), respectively (dotted lines) as well as the horizontally extrapolated IV lines. The time-to-maturity is $\tau = 1/12$, the risk-free rate is $r_f = 3\%$, the volatility parameter is $\sigma = 30\%$ and the expected stock return is $r_S = r_f + 3\%$. We assume that options trade over moneyness range from 0.94 to 1.08. Out-of-the-money call and put option IVs are interpolated separately by the cubic Hermite polynomial interpolation and extrapolated horizontally beyond the observed moneyness range. The x-axis is the moneyness and the y-axis is the implied volatility in percent.



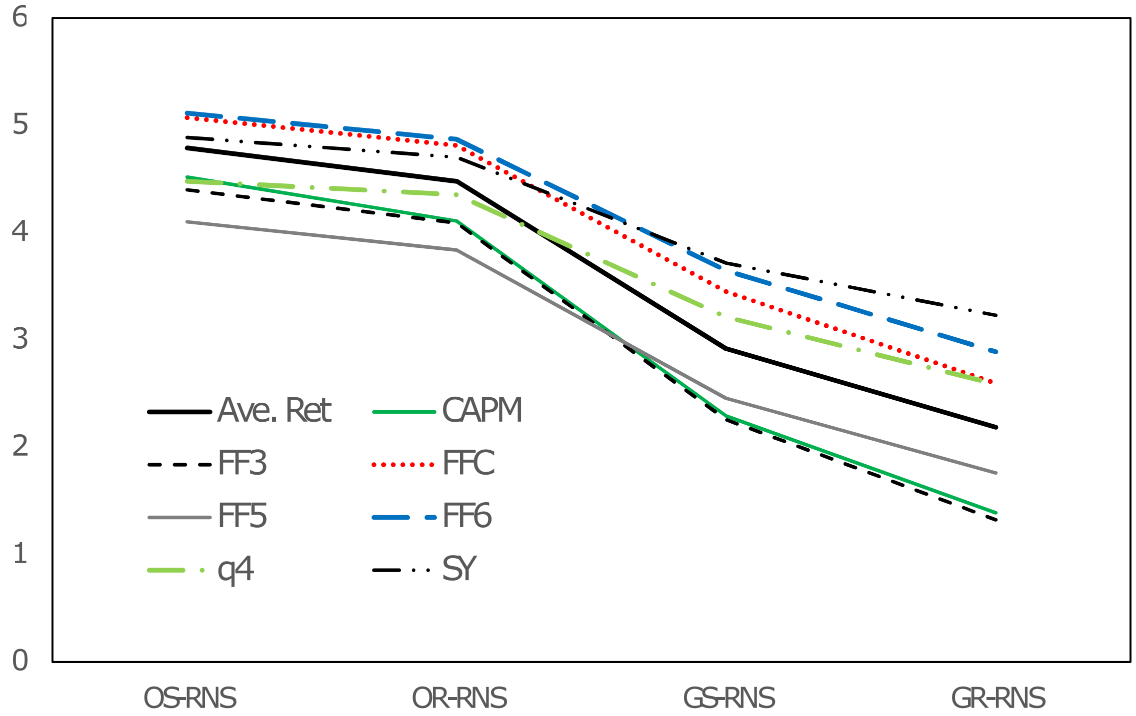
(a) Estimated daily 30-day CFER of S&P 500



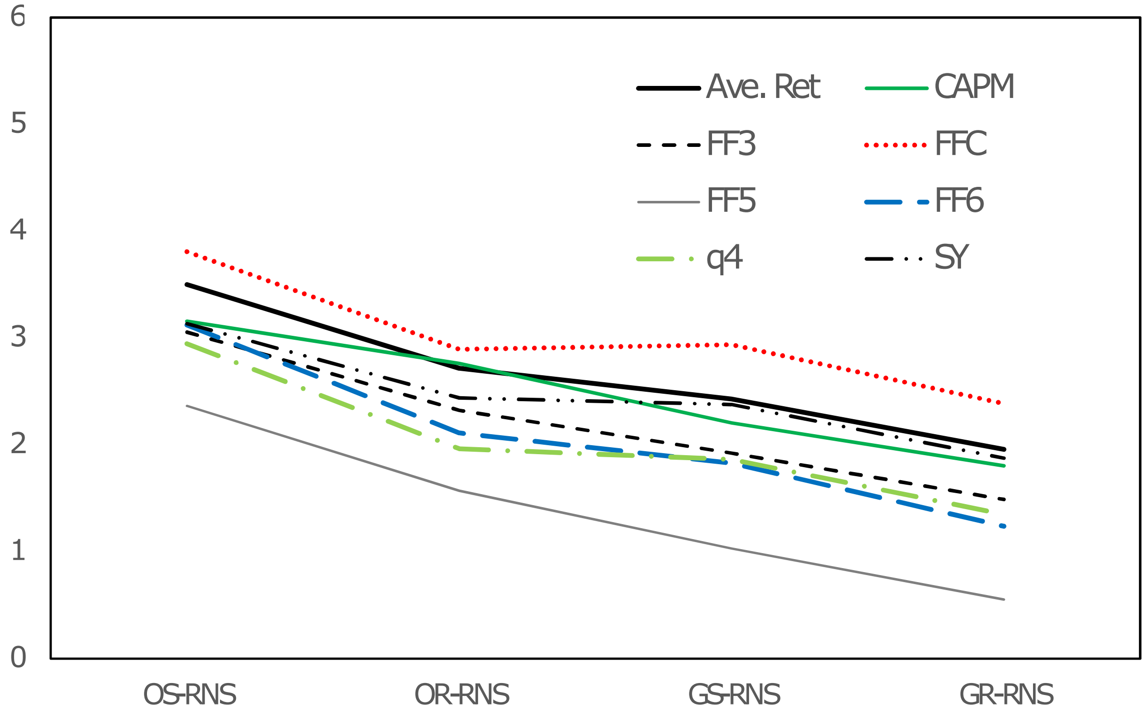
(b) Estimated daily 30-day CFER of S&P 500 (21-day moving average of the absolute value)

Figure 2.4. Estimated daily 30-day CFER of the S&P 500

Figure (a) depicts the daily 30-day CFER of the S&P 500 index estimated from the 30-day forward at-the-money call and put option prices recorded in the OptionMetrics Standardized Options file. Figure (b) depicts the 21-day moving average of the absolute value of the daily estimated 30-day CFER. The estimation period spans from the beginning of January 1996 to the end of December 2017. The unit of y-axis in the both Figures is the percentage value of CFER per year.



(a) Equally-weighted portfolios



(b) Value-weighted portfolios

Figure 2.5. The t-statistics of the average returns and alphas of the RNS-sorted long-short spread portfolios based on the four alternative RNS estimation specifications

The x-axis stands for the estimation specification of the RNS. The first letter of the prefix denotes the formula used to estimate the RNS (either Original or Generalized). The second letter of the prefix denotes the IVs used for the inter-/extrapolation (either Standard IV or Robust IV). The y-axis stands for the value of the t-statistics.

Table 2.1. Limitations of the implied stock price approach: Simulation results

This table reports the simulation results of the implied stock price approach using the BS model to calculate the implied stock price. We generate simulated call and put option prices based on $S_t = 100$, $\tau = 1/6$, $r_f = 4\%$, $\widetilde{D}_{t,T} = 0.5$ for strikes $K = 80, 85, \ldots, 120$. The true implied volatility (IV) curve is modeled as a function of moneyness, $IV(K/S_t) = \sigma_{ATM} + k(K/S_t - 1)$ with $\sigma_{ATM} = 20\%$. We consider five alternative values for k; $k = -1/2, -1/4, 0, 1/4, 1/2$. Given these assumptions, we generate a set of simulated option prices using the Black and Scholes (1973) (BS) function with deterministic dividend payments. Then, the two parameters S_t^* and σ^* are estimated by minimizing the sum of squared errors between the simulated option prices and the theoretical option prices based on the BS function $BS(S_t^*, K, T, r, q, \sigma^*, \widetilde{D}_{t,T})$. We use a set of call options and a set of put options, separately, as the observed option prices. The estimation results are reported in the columns under Call (Put) options, respectively. Given the estimate of S_t^*, we calculate the corresponding annualized estimated CFER, $\widehat{CFER} = \Delta_t e^{r_f \tau}/\tau$, which would prevail if the assumption of the implied stock approach (equation (2.12)) holds.

	Call options			Put options		
k	S_t^*	$\widehat{CFER}$	σ^*	S_t^*	$\widehat{CFER}$	σ^*
-1/2	100.38	2.3%	17.9%	100.35	2.1%	21.1%
-1/4	100.17	1.0%	19.0%	100.21	1.2%	20.6%
0	100	0.0%	20.0%	100	0.0%	20.0%
1/4	99.86	-0.9%	21.1%	99.71	-1.7%	19.1%
1/2	99.72	-1.7%	22.2%	99.35	-4.0%	18.0%

Table 2.2. Estimation of RNMs and the violation of MR: Simulation results

Entries report a simulation-based estimated RNMs based on four alternative estimation specifications. We generate simulated option prices based on the log-normal stock price distribution, equation (2.32). The parameters are set to $S_t = 100$, $r_f = 3\%$, $\tau = 1/12$ and $\sigma = 30\%$. We consider three cases of the risk-neutral expected stock return r_S: $r_S = r_f$ (zero CFER case), $r_S = r_f + 3\%$ (positive CFER case) and $r_S = r_f - 3\%$ (negative CFER case). Under this setup, the true MFIV, RNS, and RNK are 30%, zero, and three, respectively, regardless of the value of r_S. The first column specifies the RNM estimation specifications. The first letter stands for the type of the formula, either the "Original" BKM formula or the "Generalized" BKM formula. The second letter stands for the type of the IVs used as an input to the interpolation and extrapolation, either the "Standard" IVs or the "Robust" IVs.

	Zero CFER			Positive CFER			Negative CFER		
	MFIV	RNS	RNK	MFIV	RNS	RNK	MFIV	RNS	RNK
OS	30.0%	0.00	3.00	30.0%	0.11	3.00	30.0%	-0.11	3.00
OR	30.0%	0.00	3.00	30.0%	0.09	3.00	30.0%	-0.09	3.00
GS	30.0%	0.00	3.00	30.0%	0.02	2.99	30.0%	-0.02	3.01
GR	30.0%	0.00	3.00	30.0%	0.00	2.99	30.0%	0.00	3.00

Table 2.3. Summary statistics of CFER and the significance of mean absolute CFER value

Entries in Panels A and B report the summary statistics of the estimated CFER and its absolute value calculated from daily and end-of-month CFER estimates, respectively. We estimate CFER for all available U.S. common stocks traded at NYSE/Amex/NASDAQ for each trading day. P5 and P95 stands for the fifth and 95-th percentile points of the pooled stock-day observations, respectively. The unit of CFER summary statistics is % per year. Entries in Panel C report the number of individual stocks whose time-series mean absolute CFER is greater than zero, 1%, 2%, and 3%, respectively. We count stocks for having a higher than the reference point mean absolute CFER if the *t*-statistic is greater than three, where the null hypothesis is that the mean absolute CFER equals the reference point. The columns labeled "Daily" are based on the all daily observations, while the columns labeled "End-of-month" are based on the end-of-month observations only. We count stocks over three groups of stocks. For the daily case, we consider the groups of stocks with at least 21 non-missing CFER observations, 1200 non-missing observations, and 2500 non-missing observations. For the end-of-month exercise case, we consider the groups of stocks with at least 12 valid CFER observations, 60 valid observations, and 120 valid observations. The first row reports the number of stocks within each one of the groups specified by the minimum valid CFER observations. The estimation period spans from January 1996 to December 2017.

Panel A: Summary statistics: Daily estimated CFER						
	Obs	Mean	St. dev.	P5	Median	P95
CFER	12,469,967	-1.1	15.9	-17.0	-0.6	14.0
|CFER|	12,469,967	6.8	14.4	0.2	2.8	24.8
Panel B: Summary statistics: End-of-month estimated CFER						
	Obs	Mean	St. dev.	P5	Median	P95
CFER	596,679	-0.9	16.2	-17.1	-0.4	14.4
|CFER|	596,679	7.0	14.7	0.2	2.8	25.2

Panel C: Number of stocks with significant mean absolute CFER						
	Daily			End-of-month		
	Valid observations			Valid observations		
	$\geq$ 21	$\geq$ 1200	$\geq$ 2500	$\geq$ 12	$\geq$ 60	$\geq$ 120
Number of stocks	6791	3357	1836	6219	3277	1833
mean(|CFER|) > 0	6692	3337	1827	5748	3177	1797
	[98.5]	[99.4]	[99.5]	[92.4]	[96.9]	[98.0]
mean(|CFER|) > 1%	6631	3308	1807	5290	2986	1689
	[97.6]	[98.5]	[98.4]	[85.1]	[91.1]	[92.1]
mean(|CFER|) > 2%	5968	2809	1390	3989	2139	1075
	[87.9]	[83.7]	[75.7]	[64.1]	[65.3]	[58.6]
mean(|CFER|) > 3%	5139	2265	1019	2899	1470	668
	[75.7]	[67.5]	[55.5]	[46.6]	[44.9]	[36.4]

Table 2.4. Summary statistics of the estimated RNMs

Entries in Panel A report the summary statistics of CFER, the three estimated risk-neutral moments (RNMs), model-free implied volatility (MFIV), risk-neutral skewness (RNS), and risk-neutral kurtosis (RNK), computed by the original BKM formulae (prefix "O") and our generalized BKM formulae (prefix "G"). We also compute the difference between the estimated O-RNMs and the corresponding G-RNMs (with prefix "Δ"). We estimate these option-implied measures at the end of each month from January 1996 to December 2017 (264 months). P5 and P95 denote the fifth and 95th percentile point, respectively. The unit of CFER is % per 30-days, that of MFIV is % (per year). Entries in Panel B report the Pearson and Spearman pairwise correlations between CFER and RNMs, CFER and ΔRNM, and O-RNM and G-RNM.

	CFER	O-MFIV	G-MFIV	O-RNS	G-RNS	O-RNK	G-RNK	ΔMFIV	ΔRNS	ΔRNK
Panel A: Summary statistics										
Mean	-0.07	52.11	51.99	-0.20	-0.19	3.64	3.62	0.12	-0.01	0.02
St. dev.	1.33	26.99	26.68	0.51	0.48	1.16	1.16	1.52	0.23	0.39
P5	-1.40	21.54	21.60	-0.87	-0.80	2.88	2.88	-0.47	-0.30	-0.14
Median	-0.03	45.74	45.71	-0.24	-0.21	3.28	3.26	-0.02	-0.01	0.01
P95	1.18	103.58	103.08	0.58	0.50	5.72	5.70	0.89	0.27	0.21
N	582,796	582,796	582,796	582,796	582,796	582,796	582,796	582,796	582,796	582,796
Panel B: Pairwise correlations										
		CFER and RNMs						CFER and ΔRNM		
		O-MFIV	G-MFIV	O-RNS	G-RNS	O-RNK	G-RNK	ΔMFIV	ΔRNS	ΔRNK
Pearson		-0.11	-0.07	0.36	-0.01	0.13	0.04	-0.72	0.82	0.29
Spearman		-0.08	-0.07	0.34	-0.01	0.07	0.09	-0.78	0.97	-0.11
		O-RNM and G-RNM								
		MFIV		RNS		RNK				
Pearson		1.00		0.90		0.94				
Spearman		1.00		0.88		0.96				

Table 2.5. Decile portfolio sort based on the estimated RNS

Entries report the average post-ranking returns of the equally- and value-weighted decide portfolios formed based on the estimated O-RNS and G-RNS as well as the average post-ranking returns and the risk-adjusted returns (α) of the high-minus-low (10 minus 1) spread portfolios. We estimate α's with respect to the CAPM, the Fama and French (1993) three-factor model (FF3), the Carhart (1997) four-factor model (FFC), the Fama and French (2018) five- and six-factor models (FF5 and FF6), the Hou et al. (2015) q-factor model (q4), and the Stambaugh and Yuan (2017) mispricing factor model (SY). At the end of each month t, stocks are sorted in ascending order based on the estimated O-RNS or G-RNS and then equally- and value-weighted decile portfolios are formed. We then calculate the post-ranking return of these portfolios and the highest minus lowest (10 minus 1) spread portfolios in the succeeding month-($t+1$). The post-ranking return period spans February 1996 to December 2017 (263 months). t-statistics are adjusted for heteroscedasticity and autocorrelation and reported in parentheses. The unit of the average returns and alphas is % per month. N in the last row of Panel A is the average number of stocks in each decile portfolio in each month.

	Equally-weighted				Value-weighted			
	O-RNS		G-RNS		O-RNS		G-RNS	
Panel A: Average return of the decile portfolios								
1 (lowest)	0.36	(1.12)	0.78	(2.60)	0.44	(1.58)	0.67	(2.54)
2	0.59	(1.69)	0.79	(2.28)	0.70	(2.50)	0.74	(2.65)
3	0.73	(1.93)	0.82	(2.23)	0.86	(2.98)	0.82	(2.75)
4	0.76	(1.92)	0.96	(2.40)	0.87	(2.74)	0.95	(3.14)
5	0.81	(1.97)	0.90	(2.23)	1.06	(3.38)	1.04	(3.37)
6	0.86	(2.05)	0.86	(1.98)	1.07	(3.35)	1.02	(2.93)
7	1.03	(2.33)	0.78	(1.75)	1.22	(3.71)	0.94	(2.65)
8	1.02	(2.25)	0.80	(1.69)	0.91	(2.56)	0.89	(2.48)
9	1.29	(2.67)	1.01	(2.04)	1.24	(3.43)	0.99	(2.43)
10 (highest)	1.50	(3.30)	1.25	(2.83)	1.23	(3.78)	1.07	(3.38)
N	220.8		220.8		220.8		220.8	
Panel B: Average returns and alphas of the spread portfolios								
Ave. Ret	1.14	(4.79)	0.48	(2.19)	0.79	(3.50)	0.40	(1.96)
α_{CAPM}	0.90	(4.52)	0.25	(1.39)	0.74	(3.16)	0.38	(1.80)
α_{FF3}	0.90	(4.40)	0.22	(1.32)	0.65	(3.06)	0.27	(1.49)
α_{FFC}	1.14	(5.07)	0.47	(2.60)	0.81	(3.81)	0.41	(2.39)
α_{FF5}	1.06	(4.10)	0.38	(1.76)	0.53	(2.36)	0.11	(0.55)
α_{FF6}	1.21	(5.11)	0.54	(2.89)	0.65	(3.12)	0.22	(1.24)
α_{q4}	1.37	(4.48)	0.65	(2.59)	0.73	(2.95)	0.29	(1.35)
α_{SY}	1.46	(4.89)	0.79	(3.23)	0.78	(3.14)	0.39	(1.88)

Table 2.6. Factor loadings obtained from regressions of the RNS-sorted value-weighted spread portfolios on factors

Entries report the factor loadings obtained from regressing the RNS-sorted value-weighted decile spread portfolios returns on factors. The factor models are the CAPM (MKT), the Fama and French (1993) three-factor model (MKT, SMB, HML), the Carhart (1997) four-factor model (MKT, SMB, HML, UMD), the Fama and French (2018) five-factor model (MKT, SMB(FF5), HML, RMW, CMA) and six-factor model (MKT, SMB(FF5), HML, RMW, CMA, UMD), the Hou et al. (2015) q-factor model (MKT(q4), ME, IA, ROE) and the Stambaugh and Yuan (2017) mispricing-factor model (MKT, SMB(SY), MGMT, PERF). The columns labeled "O" ("G") use O-RNS (G-RNS) as a sorting variable. The post-ranking return period spans February 1996 to December 2017 (263 months). t-statistics are adjusted for heteroscedasticity and autocorrelation and reported in parentheses.

	CAPM		FF3		FFC		FF5		FF6		q4		SY	
	O	G	O	G	O	G	O	G	O	G	O	G	O	G
MKT	0.09	0.04	0.10	0.05	0.00	-0.04	0.15	0.12	0.08	0.05	0.07	0.04	0.04	-0.03
	(1.30)	(0.55)	(1.58)	(0.85)	(0.04)	(-0.82)	(2.20)	(1.81)	(1.40)	(0.98)	(1.11)	(0.66)	(0.57)	(-0.43)
SMB			0.13	0.17	0.17	0.20	0.19	0.25	0.24	0.29			0.16	0.20
			(1.30)	(1.47)	(2.15)	(2.19)	(1.86)	(2.69)	(3.08)	(4.26)			(1.42)	(1.47)
HML			0.31	0.38	0.22	0.30	0.15	0.19	-0.02	0.04				
			(3.00)	(3.81)	(2.20)	(3.03)	(1.20)	(1.65)	(-0.19)	(0.35)				
UMD					-0.24	-0.22			-0.26	-0.24				
					(-3.91)	(-3.76)			(-4.41)	(-4.51)				
RMW							0.12	0.18	0.19	0.25				
							(0.77)	(1.20)	(1.67)	(1.99)				
CMA							0.23	0.23	0.32	0.32				
							(1.13)	(1.51)	(2.11)	(2.61)				
ME											0.06	0.11		
											(0.53)	(0.74)		
IA											0.47	0.56		
											(2.92)	(3.89)		
ROE											-0.26	-0.21		
											(-1.88)	(-1.97)		
MGMT													0.18	0.18
													(1.52)	(1.42)
PERF													-0.17	-0.17
													(-1.90)	(-1.98)
Alpha	0.74	0.38	0.65	0.27	0.81	0.41	0.53	0.11	0.65	0.22	0.73	0.29	0.78	0.39
	(3.16)	(1.80)	(3.06)	(1.49)	(3.81)	(2.39)	(2.36)	(0.55)	(3.12)	(1.24)	(2.95)	(1.35)	(3.14)	(1.88)
Adj. R^2	0.01	-0.01	0.08	0.12	0.19	0.22	0.10	0.15	0.23	0.27	0.11	0.14	0.07	0.08

Table 2.7. Robustness analysis: Fama-MacBeth regressions

Entries report the results from the Fama and MacBeth (1973) regression of the stock returns on the estimated RNS, firms' and stocks' characteristics, and CFER. Columns (1) to (4) report the results using O-RNS as a regressor, and Columns (5) to (8) report those using G-RNS as a regressor. We conduct two alternative types of regressions for the cross-sectional regression at any given month: the ordinary least square (OLS) and the value-weighted least square (VWLS). The last row indicates the type of the cross-sectional regression. Beta is the market beta, SIZE is the log market capitalization, log(B/M) is the log book-to-market, Profit is the operational profit, Invest is the annual growth rate of the total asset, Momentum (MOM) is the past 11 months cumulative return $R_{t-12,t-1}$ and Reversal is the previous month return $R_{t-1,t}$. Adj. R^2 is the time-series average of the adjusted R^2 of the cross-sectional regressions. N is the time-series average of the number of stocks in the cross-sectional regressions. The return period spans February 1996 to December 2017 (263 months). t-statistics are adjusted for heteroscedasticity and autocorrelation and reported in parentheses.

FM regression of stock returns on RNS and characteristic variables								
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
O-RNS	0.73	0.71	0.70	0.43				
	(3.94)	(3.61)	(7.15)	(3.16)				
G-RNS					0.39	0.49	0.36	0.19
					(1.86)	(2.02)	(3.96)	(1.29)
BETA			-0.07	-0.11			-0.07	-0.10
			(-0.25)	(-0.32)			(-0.24)	(-0.30)
SIZE			0.03	-0.05			0.01	-0.05
			(0.46)	(-0.89)			(0.15)	(-1.07)
log(B/M)			0.11	0.01			0.11	0.01
			(1.25)	(0.14)			(1.28)	(0.10)
Profit			0.13	0.08			0.13	0.08
			(2.85)	(1.92)			(2.86)	(1.89)
Invest			-0.05	-0.03			-0.05	-0.03
			(-1.44)	(-0.86)			(-1.48)	(-0.86)
MOM			0.13	0.26			0.11	0.25
			(0.48)	(0.94)			(0.42)	(0.89)
Reversal			-0.90	-1.42			-1.04	-1.54
			(-1.63)	(-2.08)			(-1.89)	(-2.27)
Intercept	1.06	1.13	0.65	1.78	0.99	1.05	0.89	1.84
	(2.46)	(3.49)	(0.60)	(1.90)	(2.30)	(3.21)	(0.83)	(1.95)
Adj. R^2	0.25%	-3.34%	7.16%	3.07%	0.20%	-3.36%	7.11%	3.06%
N	2208	2208	2057	2057	2208	2208	2057	2057
Reg. type	OLS	VWLS	OLS	VWLS	OLS	VWLS	OLS	VWLS

Table 2.8. Robustness analysis: Quintile portfolios sorted by the estimated RNS

Entries report the average post-ranking returns of the equally- and value-weighted quintile portfolios formed based on the estimated O-RNS and G-RNS as well as the average post-ranking returns and the risk-adjusted returns (α) of the high-minus-low (5 minus 1) spread portfolio. We estimate α's with respect to the CAPM, the Fama and French (1993) three-factor model (FF3), the Carhart (1997) four-factor model (FFC), the Fama and French (2018) five- and six-factor models (FF5 and FF6), the Hou et al. (2015) q-factor model (q4), and the Stambaugh and Yuan (2017) mispricing factor model (SY). At the end of each month t, stocks are sorted in ascending order based on the estimated O-RNS or G-RNS and then equally- and value-weighted quintile portfolios are formed. We then calculate the return of these portfolios and the long-short (5 minus 1) spread portfolios in the succeeding month-($t+1$). The post-ranking return period spans February 1996 to December 2017 (263 months). t-statistics are adjusted for heteroscedasticity and autocorrelation and reported in parentheses. The unit of the average returns and alphas is % per month.

	Equally-weighted				Value-weighted			
	O-RNS		G-RNS		O-RNS		G-RNS	
Panel A: Average return of the quintile portfolios								
1 (lowest)	0.48	(1.43)	0.78	(2.45)	0.57	(2.11)	0.71	(2.67)
2	0.74	(1.94)	0.89	(2.33)	0.86	(2.94)	0.87	(2.97)
3	0.83	(2.02)	0.88	(2.12)	1.07	(3.48)	1.03	(3.27)
4	1.03	(2.30)	0.79	(1.73)	1.08	(3.23)	0.90	(2.61)
5 (highest)	1.40	(3.00)	1.13	(2.43)	1.23	(3.66)	1.02	(2.87)
Panel B: Average returns and alphas of the spread portfolios								
Ave. Ret	0.92	(4.20)	0.35	(1.66)	0.66	(3.50)	0.31	(1.48)
α_{CAPM}	0.69	(3.75)	0.11	(0.68)	0.56	(3.08)	0.20	(1.01)
α_{FF3}	0.68	(3.81)	0.09	(0.57)	0.48	(2.98)	0.10	(0.62)
α_{FFC}	0.92	(4.63)	0.34	(2.04)	0.63	(3.98)	0.26	(1.60)
α_{FF5}	0.86	(3.67)	0.28	(1.26)	0.44	(2.37)	0.07	(0.38)
α_{FF6}	1.01	(4.82)	0.44	(2.41)	0.54	(3.32)	0.18	(1.06)
α_{q4}	1.15	(4.08)	0.54	(2.08)	0.66	(3.14)	0.23	(1.11)
α_{SY}	1.23	(4.48)	0.66	(2.73)	0.67	(3.56)	0.27	(1.41)

Table 2.9. Decile portfolios sorted by the dividend non-adjusted RNS

Entries report the average post-ranking returns of the equally- and value-weighted decide portfolios formed based on the estimated dividend non-adjusted O-RNS and G-RNS as well as the average post-ranking returns and the risk-adjusted returns (α) of the high-minus-low (10 minus 1) spread portfolios. The dividend non-adjusted RNS is estimated by setting $\widetilde{D}_{t,T} = 0$, in Proposition 2.4.1. We estimate α's with respect to the CAPM, the Fama and French (1993) three-factor model (FF3), the Carhart (1997) four-factor model (FFC), the Fama and French (2018) five- and six-factor models (FF5 and FF6), the Hou et al. (2015) q-factor model (q4), and the Stambaugh and Yuan (2017) mispricing factor model (SY). At the end of each month t, stocks are sorted in ascending order based on the estimated O-RNS or G-RNS and then equally- and value-weighted decile portfolios are formed. We then calculate the return of these portfolios and the long-short (10 minus 1) spread portfolios in the succeeding month-($t + 1$). The post-ranking return period spans February 1996 to December 2017 (263 months). t-statistics are adjusted for heteroscedasticity and autocorrelation and reported in parentheses. The unit of the average returns and alphas is % per month.

	Equally-weighted				Value-weighted			
	O-RNS		G-RNS		O-RNS		G-RNS	
Panel A: Average return of the decile portfolios								
1 (lowest)	0.43	(1.41)	0.75	(2.63)	0.53	(1.97)	0.74	(2.88)
2	0.57	(1.68)	0.81	(2.43)	0.71	(2.64)	0.75	(2.78)
3	0.71	(1.89)	0.90	(2.48)	0.82	(2.82)	0.84	(2.73)
4	0.78	(1.99)	0.92	(2.33)	0.88	(2.74)	0.94	(2.98)
5	0.82	(2.00)	0.87	(2.14)	1.06	(3.23)	1.07	(3.27)
6	0.83	(1.97)	0.83	(1.89)	1.05	(3.17)	0.89	(2.55)
7	0.97	(2.17)	0.81	(1.79)	1.21	(3.46)	1.12	(3.07)
8	1.05	(2.28)	0.84	(1.77)	0.97	(2.62)	0.89	(2.32)
9	1.27	(2.60)	0.99	(1.96)	1.17	(3.18)	0.90	(2.16)
10 (highest)	1.50	(3.23)	1.24	(2.72)	1.30	(4.00)	1.06	(3.25)
Panel B: Average returns and alphas of the spread portfolios								
Ave. Ret	1.06	(4.16)	0.49	(2.04)	0.77	(3.73)	0.32	(1.64)
α_{CAPM}	0.78	(3.72)	0.20	(1.04)	0.67	(3.06)	0.24	(1.24)
α_{FF3}	0.78	(3.72)	0.18	(1.09)	0.59	(3.11)	0.14	(0.83)
α_{FFC}	1.03	(4.56)	0.44	(2.47)	0.72	(3.83)	0.29	(1.87)
α_{FF5}	0.97	(3.79)	0.40	(1.87)	0.52	(2.63)	0.08	(0.45)
α_{FF6}	1.13	(4.80)	0.56	(3.07)	0.61	(3.36)	0.19	(1.18)
α_{q4}	1.29	(4.24)	0.66	(2.64)	0.70	(3.20)	0.22	(1.11)
α_{SY}	1.37	(4.61)	0.80	(3.29)	0.73	(3.38)	0.32	(1.68)

Table 2.10. Decile portfolio sort based on ΔRNS

Entries report the average post-ranking returns of the equally- and value-weighted decile portfolios formed based on the ΔRNS, which is the difference between O-RNS and G-RNS, as well as the average post-ranking returns and the risk-adjusted returns (α) of the high-minus-low (10 minus 1) spread portfolios. We estimate α's with respect to the CAPM, the Fama and French (1993) three-factor model (FF3), the Carhart (1997) four-factor model (FFC), the Fama and French (2018) five- and six-factor models (FF5 and FF6), the Hou et al. (2015) q-factor model (q4), and the Stambaugh and Yuan (2017) mispricing factor model (SY). At the end of each month t, stocks are sorted in ascending order based on the estimated ΔRNS and then equally- and value-weighted decile portfolios are formed. We then calculate the return of these portfolios and the long-short (10 minus 1) spread portfolios in the succeeding month-$(t+1)$. The post-ranking return period spans February 1996 to December 2017 (263 months). t-statistics are adjusted for heteroscedasticity and autocorrelation and reported in parentheses. The unit of the average returns and alphas is % per month. N in the last row of Panel A is the average number of stocks in each decile portfolio.

	ΔRNS-sorted portfolios			
	Equally-weighted		Value-weighted	
Panel A: Average return of the decile portfolios				
1 (lowest)	-0.02	(-0.04)	0.22	(0.66)
2	0.44	(1.10)	0.59	(2.07)
3	0.62	(1.59)	0.65	(2.07)
4	0.80	(2.01)	0.73	(2.31)
5	0.92	(2.21)	0.82	(2.62)
6	0.92	(2.20)	1.01	(3.05)
7	1.01	(2.47)	1.02	(3.22)
8	1.19	(2.90)	0.90	(2.81)
9	1.40	(3.39)	1.14	(3.70)
10 (highest)	1.66	(3.84)	1.41	(4.36)
N	220.8		220.8	
Panel B: Average returns and alphas of the spread portfolios				
Ave. Ret	1.68	(9.70)	1.19	(5.69)
α_{CAPM}	1.59	(9.23)	1.11	(4.80)
α_{FF3}	1.63	(9.56)	1.14	(4.71)
α_{FFC}	1.71	(9.41)	1.27	(5.07)
α_{FF5}	1.71	(8.71)	1.12	(4.62)
α_{FF6}	1.76	(8.85)	1.21	(4.97)
α_{q4}	1.91	(8.56)	1.35	(4.84)
α_{SY}	1.82	(7.39)	1.33	(4.59)

Table 2.11. Dependent bivariate sort: First by $|CFER|$ and then by O-RNS

Entries report the result of the bivariate dependent sort, where we first sort stocks based on the absolute value of the estimated CFER and then by O-RNS. At the end of each month, we first sort stocks into four subgroups based on the absolute value of the estimated CFER, and then within each absolute CFER group, we further sort stocks into quartile portfolios by the O-RNS criterion. Each column corresponds to the level of the first sorting variable and each row reports the average return and alphas of the quartile spread portfolios with respect to the second sorting variable, O-RNS. The post-ranking return period spans February 1996 to December 2017 (263 months). t-statistics are adjusted for heteroscedasticity and autocorrelation and reported in parentheses. The unit of the average returns and alphas is % per month.

	Equally-weighted				Value-weighted			
$\|CFER\|$	1 (Lo)	2	3	4 (Hi)	1 (Lo)	2	3	4 (Hi)
Highest O-RNS minus Lowest O-RNS spread portfolios								
Ave. Ret	0.26	0.37	0.57	1.51	0.36	0.18	0.63	1.15
	(1.43)	(1.94)	(2.67)	(6.18)	(1.85)	(0.78)	(2.93)	(4.58)
α_{CAPM}	0.07	0.17	0.39	1.28	0.29	0.05	0.56	1.06
	(0.45)	(1.00)	(2.02)	(5.88)	(1.48)	(0.22)	(2.85)	(4.17)
α_{FF3}	0.02	0.16	0.41	1.33	0.23	0.04	0.53	1.07
	(0.16)	(1.07)	(2.21)	(6.00)	(1.23)	(0.17)	(2.62)	(4.25)
α_{FFC}	0.19	0.31	0.66	1.53	0.31	0.09	0.68	1.24
	(1.36)	(1.88)	(3.44)	(6.19)	(1.68)	(0.44)	(3.44)	(4.36)
α_{FF5}	0.11	0.31	0.59	1.53	0.18	0.06	0.50	1.17
	(0.61)	(1.57)	(2.45)	(5.56)	(0.89)	(0.25)	(2.05)	(3.98)
α_{FF6}	0.22	0.40	0.75	1.65	0.24	0.09	0.60	1.27
	(1.42)	(2.22)	(3.71)	(6.11)	(1.21)	(0.43)	(2.81)	(4.16)
α_{q4}	0.29	0.45	0.86	1.84	0.26	0.18	0.75	1.42
	(1.59)	(2.03)	(3.02)	(6.06)	(1.25)	(0.81)	(3.09)	(4.24)
α_{SY}	0.36	0.56	1.02	1.85	0.27	0.17	0.71	1.38
	(2.10)	(2.80)	(3.75)	(5.73)	(1.33)	(0.75)	(3.04)	(3.73)

Table 2.12. CFER-adjusted regressions of the O-RNS-sorted portfolios

Entries report the intercepts of the regressions of CFER-adjusted excess returns of the value-weighted decide portfolios formed based on the estimated O-RNS, equation (2.48), on a set of risk factor(s) of the CAPM, the Fama and French (1993) three-factor model (FF3), the Carhart (1997) four-factor model (FFC), the Fama and French (2018) five- and six-factor models (FF5 and FF6), the Hou et al. (2015) q-factor model (q4), and the Stambaugh and Yuan (2017) mispricing factor model (SY), respectively. At the end of each month t, stocks are sorted in ascending order based on the estimated O-RNS and then and value-weighted decile portfolios are formed. We then calculate the CFER-adjusted return of these portfolios and the spread portfolio in the succeeding month-$(t+1)$ and regress it on the risk-factors. The post-ranking return period spans February 1996 to December 2017 (263 months). t-statistics are adjusted for heteroscedasticity and autocorrelation and reported in parentheses. The unit of CFER is % per 30 days and that of the alphas is % per month.

	O-RNS-sorted value-weighted decile portfolios										Spread
	1 (L)	2	3	4	5	6	7	8	9	10 (H)	10-1
CFER	-0.27	-0.11	-0.06	-0.04	-0.01	0.01	0.04	0.09	0.18	0.57	0.84
	(-16.54)	(-10.03)	(-7.84)	(-4.41)	(-1.35)	(0.74)	(5.02)	(8.77)	(11.99)	(9.86)	(12.54)
α_{CAPM}	-0.04	0.02	0.06	0.02	0.19	0.16	0.27	-0.11	0.14	-0.14	-0.10
	(-0.34)	(0.22)	(0.81)	(0.20)	(1.86)	(1.40)	(2.20)	(-0.81)	(0.85)	(-0.74)	(-0.44)
α_{FF3}	-0.06	0.02	0.06	0.03	0.23	0.16	0.28	-0.12	0.07	-0.25	-0.19
	(-0.51)	(0.22)	(0.83)	(0.31)	(2.41)	(1.61)	(2.31)	(-0.89)	(0.45)	(-1.46)	(-0.93)
α_{FFC}	-0.05	0.03	0.05	0.05	0.23	0.18	0.32	-0.04	0.22	-0.07	-0.02
	(-0.44)	(0.40)	(0.60)	(0.51)	(2.38)	(1.77)	(2.60)	(-0.25)	(1.49)	(-0.41)	(-0.09)
α_{FF5}	-0.01	0.00	0.04	0.02	0.24	0.17	0.29	-0.11	0.07	-0.31	-0.30
	(-0.09)	(0.01)	(0.55)	(0.16)	(2.37)	(1.65)	(2.32)	(-0.73)	(0.41)	(-1.59)	(-1.36)
α_{FF6}	-0.01	0.01	0.03	0.03	0.24	0.17	0.32	-0.06	0.17	-0.19	-0.18
	(-0.06)	(0.15)	(0.42)	(0.30)	(2.34)	(1.74)	(2.54)	(-0.37)	(1.13)	(-1.08)	(-0.88)
α_{q4}	-0.03	-0.04	0.03	0.06	0.29	0.24	0.38	0.07	0.24	-0.15	-0.11
	(-0.27)	(-0.54)	(0.37)	(0.50)	(2.68)	(2.08)	(2.97)	(0.48)	(1.26)	(-0.65)	(-0.46)
α_{SY}	0.01	-0.01	0.03	0.05	0.31	0.23	0.38	0.04	0.27	-0.03	-0.04
	(0.09)	(-0.18)	(0.35)	(0.43)	(2.92)	(1.97)	(2.97)	(0.24)	(1.62)	(-0.14)	(-0.16)

Table 2.13. Subsample analysis: traded versus non-traded options

Entries report the average post-ranking returns of the value-weighted decide portfolios formed based on the estimated O-RNS and G-RNS as well as the average post-ranking returns and the risk-adjusted returns (α) of the high-minus-low (10 minus 1) spread portfolios. We split our stock universe into two subgroups based on the aggregate option trading volume on each sorting date. The "traded" subgroup contains stocks whose options have non-zero trading volume, whereas the "non-traded" subgroup contains stocks whose options have zero trading volume. We estimate α's with respect to the CAPM, the Fama and French (1993) three-factor model (FF3), the Carhart (1997) four-factor model (FFC), the Fama and French (2018) five- and six-factor models (FF5 and FF6), the Hou et al. (2015) q-factor model (q4), and the Stambaugh and Yuan (2017) mispricing factor model (SY). At the end of each month t, stocks are sorted in ascending order based on the estimated O-RNS or G-RNS within each one of the subgroups and then and value-weighted decile portfolios are formed. We then calculate the return of these portfolios and the long-short (10 minus 1) spread portfolios in the succeeding month-($t+1$). The post-ranking return period spans February 1996 to December 2017 (263 months). t-statistics are adjusted for heteroscedasticity and autocorrelation and reported in parentheses. The unit of the average returns and alphas is % per month. N in the last row of Panel A is the average number of stocks in each decile portfolio.

	O-RNS				G-RNS			
	Traded		Non-traded		Traded		Non-Traded	
Panel A: Average return of the decile portfolios								
1 (lowest)	0.53	(1.98)	0.67	(2.53)	0.72	(2.86)	0.93	(3.42)
2	0.68	(2.48)	0.97	(3.07)	0.75	(2.72)	1.08	(3.46)
3	0.87	(3.00)	0.96	(3.08)	0.84	(2.75)	1.08	(3.61)
4	0.90	(2.77)	1.10	(3.30)	0.92	(2.88)	1.10	(3.27)
5	0.99	(2.97)	1.13	(3.33)	1.02	(3.16)	0.93	(2.77)
6	1.09	(3.24)	1.12	(3.10)	0.89	(2.62)	1.10	(2.81)
7	1.11	(3.26)	1.41	(4.03)	1.13	(3.05)	1.35	(3.60)
8	0.98	(2.57)	1.29	(3.39)	0.93	(2.45)	1.14	(2.94)
9	1.24	(3.45)	1.61	(4.16)	0.73	(1.79)	1.37	(3.69)
10 (highest)	1.21	(3.45)	1.37	(4.03)	1.07	(2.99)	1.26	(3.82)
N	174.4		46.8		174.2		46.6	
Panel B: Average returns and alphas of the spread portfolios								
Ave. Ret	0.68	(3.11)	0.69	(3.58)	0.35	(1.61)	0.33	(2.02)
α_{CAPM}	0.53	(2.46)	0.59	(3.14)	0.21	(1.04)	0.27	(1.48)
α_{FF3}	0.47	(2.36)	0.54	(2.84)	0.12	(0.66)	0.22	(1.28)
α_{FFC}	0.62	(3.16)	0.71	(3.42)	0.29	(1.68)	0.39	(2.20)
α_{FF5}	0.44	(2.08)	0.49	(2.79)	0.12	(0.61)	0.22	(1.37)
α_{FF6}	0.55	(2.85)	0.62	(3.61)	0.23	(1.33)	0.33	(2.16)
α_{q4}	0.66	(2.90)	0.76	(3.34)	0.27	(1.32)	0.35	(1.71)
α_{SY}	0.68	(2.98)	0.73	(2.88)	0.33	(1.60)	0.45	(1.99)

Table 2.14. Summary statistics of the estimated IV skew measures

Entries in Panel A show the summary statistics of CFER, the XZZ IV slope measures calculated based on the standard IV (XZZ^o) and based on the robust IV (XZZ^r) as well as the difference between the two XZZ measures. P5 and P95 stand for the fifth and 95th percentile, respectively. The unit of CFER is % per 30-day. The unit for other columns is the volatility in percentage. N stands for the average observation per each end-of-month trading day. Entries in Panel B report the pairwise correlations between CFER and the XZZ slope measures, and XZZ^o and XZZ^r. We omit samples below the 0.1 percentile point and above 99.9 percentile point in the estimated XZZ^r because outliers may affect the calculation of correlations. We estimate monthly end-of-month IV slope measures from January 1996 to December 2017 (264 months).

	CFER	XZZ^o	XZZ^r	ΔXZZ
Panel A: Summary statistics				
Mean	-0.07	6.84	6.29	0.53
St. dev.	1.34	14.84	10.06	8.98
P5	-1.41	-6.79	-2.44	-8.09
Median	-0.03	4.95	4.21	0.26
P95	1.18	27.31	23.10	9.59
N	2214	2218	2209	2213
Panel B: Pairwise correlations				
		CFER vs XZZ measures		
		XZZ^o	XZZ^r	ΔXZZ
Pearson		-0.65	-0.07	-1.00
Spearman		-0.51	-0.06	-0.99
		XZZ^o vs XZZ^r		
Pearson		0.78		
Spearman		0.80		

Table 2.15. Decile portfolio sort based on the XZZ IV slope measures

Entries report the average post-ranking returns of the equally- and value-weighted decide portfolios formed based on the estimated XZZ^o and XZZ^r as well as the average post-ranking returns and the risk-adjusted returns (α) of the high-minus-low (10 minus 1) spread portfolios. We estimate α's with respect to the CAPM, the Fama and French (1993) three-factor model (FF3), the Carhart (1997) four-factor model (FFC), the Fama and French (2018) five- and six-factor models (FF5 and FF6), the Hou et al. (2015) q-factor model (q4), and the Stambaugh and Yuan (2017) mispricing factor model (SY). At the end of each month t, stocks are sorted in ascending order based on the estimated XZZ^o or XZZ^r and then equally- and value-weighted decile portfolios are formed. We then calculate the post-ranking return of these portfolios and the long-short (10 minus 1) spread portfolios in the succeeding month-($t+1$). The post-ranking return period spans February 1996 to December 2017 (263 months). t-statistics are adjusted for heteroscedasticity and autocorrelation and reported in parentheses. The unit of the average returns and alphas is % per month. N in the last row of Panel A is the average number of stocks in each decile portfolio in each month.

	Equally-weighted				Value-weighted			
	XZZ^o		XZZ^r		XZZ^o		XZZ^r	
Panel A: Average return of the decile portfolios								
1 (lowest)	1.53	(2.87)	1.15	(2.17)	1.75	(4.25)	1.03	(2.77)
2	1.36	(3.25)	1.07	(2.40)	1.18	(4.06)	1.25	(4.22)
3	1.14	(3.17)	0.88	(2.38)	1.10	(4.34)	0.90	(3.49)
4	1.00	(2.83)	0.90	(2.47)	0.95	(3.57)	0.82	(3.18)
5	0.91	(2.58)	0.90	(2.50)	0.75	(2.81)	0.82	(3.08)
6	0.86	(2.27)	0.83	(2.20)	0.68	(2.23)	0.79	(2.61)
7	0.66	(1.69)	0.79	(2.05)	0.70	(2.23)	0.66	(2.20)
8	0.66	(1.56)	0.80	(1.97)	0.54	(1.48)	0.88	(2.61)
9	0.61	(1.44)	0.82	(1.90)	0.60	(1.51)	0.78	(2.02)
10 (highest)	0.20	(0.44)	0.80	(1.92)	0.30	(0.64)	0.65	(1.55)
N	221.3		220.7		221.3		220.7	
Panel B: Average returns and alphas of the spread portfolios								
Ave. Ret	-1.33	(-6.25)	0.49	(1.80)	-1.45	(-4.67)	-0.38	(-1.46)
α_{CAPM}	-1.18	(-5.66)	0.16	(0.72)	-1.49	(-4.50)	-0.44	(-1.86)
α_{FF3}	-1.24	(-6.29)	0.14	(0.69)	-1.53	(-4.69)	-0.44	(-1.89)
α_{FFC}	-1.35	(-5.93)	0.41	(1.99)	-1.53	(-4.37)	-0.42	(-1.69)
α_{FF5}	-1.34	(-5.76)	0.37	(1.48)	-1.30	(-4.10)	-0.23	(-1.04)
α_{FF6}	-1.41	(-5.88)	0.55	(2.51)	-1.31	(-3.94)	-0.22	(-0.97)
α_{q4}	-1.63	(-6.09)	0.65	(2.23)	-1.58	(-4.19)	-0.35	(-1.44)
α_{SY}	-1.54	(-5.34)	0.78	(2.86)	-1.39	(-3.89)	-0.16	(-0.66)

Table 2.16. CFER-adjusted regression of the XZZ-sorted portfolios

Entries report the intercepts of the regressions of CFER-adjusted excess returns of the value-weighted decide portfolios formed based on the estimated XZZ^o, equation (2.48), on a set of risk factor(s) of the CAPM, the Fama and French (1993) three-factor model (FF3), the Carhart (1997) four-factor model (FFC), the Fama and French (2018) five- and six-factor models (FF5 and FF6), the Hou et al. (2015) q-factor model (q4), and the Stambaugh and Yuan (2017) mispricing factor model (SY), respectively. At the end of each month t, stocks are sorted in ascending order based on the estimated XZZ^o and then and value-weighted decile portfolios are formed. We then calculate the CFER-adjusted return of these portfolios and the long-short (10 minus 1) spread portfolio in the succeeding month-($t+1$) and regress it on the risk-factors. The post-ranking return period spans February 1996 to December 2017 (263 months). t-statistics are adjusted for heteroscedasticity and autocorrelation and reported in parentheses. The unit of CFER is % per 30 days and that of the alphas is % per month.

	XZZ-sorted value-weighted decile portfolios										Spread
	1 (Lo)	2	3	4	5	6	7	8	9	10 (Hi)	10-1
CFER	12.86	2.19	0.74	0.02	-0.52	-0.99	-1.61	-2.51	-3.60	-8.51	-21.37
	(11.53)	(8.09)	(6.23)	(0.32)	(-7.57)	(-14.18)	(-17.20)	(-18.95)	(-17.52)	(-13.92)	(-12.96)
α_{CAPM}	-0.29	0.18	0.28	0.19	0.02	-0.07	-0.04	-0.20	-0.10	-0.01	0.28
	(-1.30)	(1.53)	(3.13)	(1.69)	(0.18)	(-0.81)	(-0.36)	(-1.68)	(-0.63)	(-0.02)	(0.94)
α_{FF3}	-0.33	0.17	0.27	0.19	0.01	-0.10	-0.05	-0.22	-0.17	-0.08	0.25
	(-1.47)	(1.38)	(3.29)	(1.78)	(0.06)	(-1.12)	(-0.47)	(-1.84)	(-1.17)	(-0.36)	(0.84)
α_{FFC}	-0.11	0.20	0.25	0.11	-0.03	-0.11	-0.05	-0.19	-0.07	0.13	0.24
	(-0.45)	(1.56)	(2.85)	(1.11)	(-0.32)	(-1.20)	(-0.49)	(-1.54)	(-0.42)	(0.68)	(0.71)
α_{FF5}	-0.24	0.08	0.21	0.02	-0.06	-0.12	-0.03	-0.07	0.00	0.20	0.44
	(-0.94)	(0.60)	(2.40)	(0.20)	(-0.79)	(-1.43)	(-0.25)	(-0.64)	(-0.02)	(0.99)	(1.38)
α_{FF6}	-0.10	0.10	0.19	-0.02	-0.08	-0.13	-0.03	-0.06	0.06	0.33	0.43
	(-0.40)	(0.80)	(2.17)	(-0.19)	(-1.01)	(-1.45)	(-0.28)	(-0.50)	(0.38)	(1.74)	(1.24)
α_{q4}	0.06	0.17	0.19	0.06	-0.06	-0.13	-0.03	-0.10	0.03	0.26	0.20
	(0.20)	(1.17)	(2.18)	(0.57)	(-0.68)	(-1.38)	(-0.23)	(-0.74)	(0.17)	(1.12)	(0.55)
α_{SY}	0.14	0.19	0.19	-0.08	-0.11	-0.11	-0.01	-0.06	0.14	0.46	0.33
	(0.50)	(1.33)	(2.12)	(-0.86)	(-1.38)	(-1.08)	(-0.12)	(-0.51)	(0.81)	(2.30)	(0.92)

Chapter 3

Option-Implied Expected Returns and the Construction of Mean-Variance Portfolios

3.1 Introduction

Despite its theoretical importance and elegance, the empirical performance of the Markowitz (1952) mean-variance portfolios using sample means and covariances has been known to be poor. Among the two ingredients of the mean-variance portfolio problem, the expected stock returns have been perceived as the harder one to estimate with precision (e.g., Merton, 1980). Most notably, Jagannathan and Ma (2003) argue that *"[t]he estimation error in the sample mean is so large that nothing much is lost in ignoring the mean altogether when no further information about the population mean"* (pp.1652–1653). In line with their argument, a large proportion of studies focus on the estimation of the covariance matrix in the minimum variance portfolio construction.[1] This leaves the research on the estimation of the expected stock returns relatively scarce.

This study fills this gap in the literature by proposing a novel way to utilize option-implied information on the population mean in the mean-variance portfolio optimization. To this end, we generalize Martin and Wagner's (2018) formula for

[1]Examples include the shrinkage covariance matrix estimators in Ledoit and Wolf (2003, 2004, 2017), the factor model-based estimator (e.g., Chan et al., 1999) and the imposing constraints on the portfolio weight (e.g., Jagannathan and Ma, 2003; DeMiguel et al., 2009a; Fan et al., 2012).

the expected stock return by allowing the existence of market frictions considered in Hiraki and Skiadopoulos (2019) (henceforth HS).[2] Roughly speaking, our generalized formula shows that the expected excess stock return in the presence of market frictions equals the sum of two components: the risk premium component which is expressed by the scaled risk-neutral simple stock variance of Martin (2017) and the *contribution of frictions to expected return* (CFER) component of HS. CFER represents the effect of market frictions on the expected return; it is part of the expected stock return distinct from the risk premium component. Since Martin (2017) and HS show that the risk-neutral simple variance and CFER, respectively, can be estimated from option price data, our formula allows us to obtain a real-time, forward-looking estimate of the expected stock returns. Based on this theoretical result, we examine three mean-variance portfolio strategies, for which we use CFER, the risk premium component implied by the risk-neutral simple variance, and the sum of these two components, respectively, as the estimate of the expected stock returns.

We examine the empirical performance of our option-based mean-variance portfolio strategies as well as a number of benchmark strategies (e.g., equally-weighted portfolio) using the U.S. individual stock data from 1996 to 2017. We find that the mean-variance portfolio strategy for which we use the estimated CFER as the estimate of the expected stock returns (the Q-CFER strategy; Q stands for "quantitative" estimate of the expected return) obtains the best performance, especially when constraints on portfolio weights are imposed. This result brings three important implications.

First, our findings indicate that utilizing the option-based estimate of the expected stock returns improves the performance of the mean-variance portfolio. The Q-CFER strategy outperforms other portfolio strategies which have been documented to have good performance, including the equally-weighted portfolio (DeMiguel et al., 2009b), minimum variance portfolios (e.g., Jagannathan and Ma, 2003), both before and after considering a moderate size of transaction costs. For example, in our baseline exercise over the member stocks of the S&P 500, the annualized (pre-transaction costs) Sharpe ratio of the Q-CFER strategy with constraints on portfolio weights is 0.98, while that of the equally-weighted portfolio and the minimum variance portfolio

[2] Chapter One of this thesis is based on Hiraki and Skiadopoulos (2019).

are 0.58 and 0.68, respectively. This may change the common perception in the literature that it is notoriously difficult to construct a mean-variance portfolio with good empirical performance. Moreover, the Q-CFER strategy outperforms previously proposed option-based portfolio strategies in DeMiguel et al. (2013) (DPUV), Kempf et al. (2015), and Martin and Wagner (2018). We shortly argue that this is because we utilize richer informational content of option prices compared to these studies.

Second, the outperformance of the Q-CFER strategy suggests that incorporating the effect of market frictions into the estimation of the expected return is of importance. On the other hand, we find that utilizing the option-implied estimate of the risk premium component does not improve the portfolio performance in general. This sharp difference between CFER and the risk premium component stems from the fact that the latter is a compensation for risk exposures, whereas the CFER component is not. Tilting portfolio weights toward stocks with a high CFER component is effective to improve the portfolio expected return *without* much increasing the portfolio variance. On the contrary, tilting portfolio weights toward stocks with a high risk premium component mechanically increases the portfolio variance. Therefore, it is not a priori clear whether the trade-off between a higher mean and a higher variance is improved.

Third, we find that imposing constraints on portfolio weights dramatically improves the performance of mean-variance portfolios. This result complements the empirically unexploited theoretical result in Jagannathan and Ma (2002, 2003) on the *mean*-variance portfolio construction.[3] Albeit less famous compared to their seminal result on the *minimum* variance portfolio, they theoretically show that portfolio weight constraints on the mean-variance portfolio problem can be interpreted to have a *shrinkage-like* effect. In particular, Jagannathan and Ma (2003, p.1660) discuss that

> [T]he effect of the lower bound on portfolio weights is to adjust the mean returns upward by an amount proportional to the Lagrange multiplier [of the weight constraints], and the effect of the upper bound on portfolio weights is a similar downward adjustment.

Nevertheless, they document that *"the sampling error in the estimated [histori-*

[3] Jagannathan and Ma (2002) is the working paper version of Jagannathan and Ma (2003), which provides more detailed analysis on the effect of constraints on the mean-variance portfolio construction.

cal sample] mean returns is too large for this shrinkage to be useful" (Jagannathan and Ma, 2002, p.24) and urge us to *"bring additional information about the population mean instead of relying on sample mean as an estimator of population mean"* (Jagannathan and Ma, 2003, p.1660). Our result shows for the first time that this shrinkage-like effect of the weight constraints on the mean-variance portfolio problem is useful once the option-implied expected return is employed. This result also implies that our option-based forward-looking estimate of the expected returns bring additional information about the population mean as requested by Jagannathan and Ma.

The study in this Chapter contributes to three strands of the literature. The first strand is the literature on the estimation of the expected stock returns as an input to the mean-variance portfolio problem. The previous literature proposes various ways to reduce measurement errors in estimates of the expected stock returns such as employing a Bayesian approach and factor-based expected stock return models (e.g., Black and Litterman, 1991; Garlappi et al., 2007; Lai et al., 2011, among others). Our option-based approach differs from these approaches in two important ways. The first difference is that our option-based expected stock return is forward-looking and we can estimation it in a model-free manner without relying on the historical data. The second difference is that our estimation approach incorporates the effect of market frictions on the expected returns. These features of our approach enable us to utilize additional information on the expected stock returns compared to the previous approaches.

Second, our study contributes to the literature on the application of option-implied information for asset allocations and portfolio selection. Aït-Sahalia and Brandt (2008) and Kostakis et al. (2011) use the option-implied distributional information for asset allocations, although they do not address the portfolio selection problem over the universe of individual stocks. Kempf et al. (2015) propose estimation approaches for the fully forward-looking covariance matrix. However, they examine only minimum variance portfolios and do not address how option-implied information can be used to construct the expected return input.

Among this strand of studies, our study is closely related to DPUV and Martin and Wagner (2018) in that they also document that option-implied information on the expected returns is useful to enhance the portfolio performance. DPUV document that four option-based return predicting characteristic variables, the model-free implied volatility (MFIV), the risk-neutral skewness (RNS), the implied volatility spread (IVS), and the volatility risk premium (VRP), are useful to improve the Sharpe ratio of portfolios.[4] Specifically, they utilize the informational content of these four variables based on the parametric portfolio approach of Brandt et al. (2009) and the characteristic-adjusted return approach. In the parametric portfolio approach adopted by DPUV, portfolio allocations are determined by a linear function of the option-implied characteristics where the loading of each variable is estimated using historical data. In the characteristic-adjusted return approach, expected stock returns take either one of three values depending on the cross-sectional ranking of an option-based characteristic variable. Martin and Wagner (2018) follow Asness et al. (2013) and construct a portfolio based on the cross-sectional ranking of the estimated individual stock risk premium components, even though their estimated stock risk premia can be used as an input to the mean-variance portfolio problem. Our Q-CFER strategy outperforms these previously proposed option-based strategies. This is because our option-based mean-variance portfolio approach utilizes richer informational content of options; DPUV rely on the option-based characteristics variables which are only *proxies* of future stock returns, while Martin and Wagner (2018) utilize only the cross-sectional ranking information of their estimated stock risk premia.[5]

Third, our findings on the effect of portfolio weight constraints is pertaining to a large number of studies which document that the mean-variance portfolio optimization has an *error maximization* property, that is, small measurement errors in the estimates of the expected returns and covariance matrix are amplified in the resulted optimal portfolio allocation (e.g., Michaud, 1989; Best and Grauer, 1991; Chopra and Ziemba, 1993). Even though our option-based estimate of the expected stock

[4]There is a voluminous amount of the literature on the return predictive power of option-based variables. For instance, Bali and Hovakimian (2009) and Cremers and Weinbaum (2010) find that the IVS predicts future stock returns. Bali and Hovakimian (2009) also document the return predictive ability of VRP. Conrad et al. (2013), Stilger et al. (2017) and Borochin and Zhao (2018) among others document the predictive ability of the risk-neutral skewness (RNS) estimated based on the Bakshi et al. (2003) formulae. See survey studies by Giamouridis and Skiadopoulos (2011) and Christoffersen et al. (2013) for more detail.

[5]DPUV also document that option-implied variance is useful to reduce the variance of minimum variance portfolios. We find that their result does not extend to general mean-variance portfolios; using the option-implied variance does not necessarily improve the performance of mean-variance portfolios.

returns is thought to be much less noisier than the historical sample mean, *the wrong constraints help* to improve the mean-variance portfolio performance, analogous to Jagannathan and Ma's (2003) discussion on the minimum variance portfolio. Moreover, this literature shows that the mean-variance weight is more sensitive to measurement errors in the expected returns than those in the covariance matrix, implying that the precision of the estimation of the expected returns is of first-order importance. In line with this result, we find that the expected return specification is more pertinent to the performance of portfolios compared to the choice of the covariance matrix specification.

The remaining part of this Chapter is organized as follows. In Section 3.2, we derive the option-based formula for the expected return under market frictions. Section 3.3 explains the data and the estimation procedures for option-implied variables. Section 3.4 explains the empirical portfolio strategies and performance measures. In Section 3.5, we report our main empirical results and discuss their implications. Section 3.6 concludes this study.

3.2 Theoretical model for the expected returns

3.2.1 Setup

Let us consider the market where N individual stocks and their equity index are traded. The market price and dividend payments of individual stocks at time t are denoted by $S_{i,t}$ and $D_{i,t}$, respectively, and those of the market index are denoted by $S_{M,t}$ and $D_{M,t}$, respectively. The gross return of the stock is denoted by $R_{i,t,t+1} = (S_{i,t+1} + D_{i,t+1})/S_{i,t}$, $i = 1, \dots, N$ or $i = M$. We assume that there are one-period European options (i.e., options expiring in the next period) written on individual stocks and the market index. A call (put) option stock i with strike K is denoted by $C_{i,t}(K)$ $(P_{i,t}(K))$. The gross risk-free rate from time t to $t+1$ is denoted by $R_{f,t,t+1}$.

We follow HS and assume that there is an agent who is a marginal investor in both the stock and the option market. The agent chooses her consumption c_t and asset allocation $\boldsymbol{\theta}_t$, which is a vector of allocations on traded assets, by maximizing her lifetime utility $\sum_{j\geq t}\beta^{j-t}\mathbb{E}_t^{\mathbb{P}}[u(c_j)]$, where β is the subjective discount factor and $\mathbb{E}_t^{\mathbb{P}}$ is the conditional expectation given the information up to time t. This maximization problem is subject to the following dynamics of the agent's wealth W_t,

$$
\begin{aligned}
c_t &= W_t - \theta_t^0 - \sum_{i=1}^{N}\left(\theta_{i,t}^S S_{i,t} + \sum_K (\theta_{i,t}^c(K) C_{i,t}(K) + \theta_{i,t}^p(K) P_{i,t}(K))\right) - TC_t(\Delta\boldsymbol{\theta}_t),\\
W_{t+1} &= \theta_t^0 R_{f,t,t+1} + \sum_{i=1}^{N} \theta_{i,t}^S (S_{i,t+1} + D_{i,t+1})\\
&\quad + \sum_{i=1}^{N}\sum_K \left[\theta_{i,t}^c(K)\max(S_{i,t+1} - K, 0) + \theta_{i,t}^p(K)\max(K - S_{i,t+1}, 0)\right],
\end{aligned}
$$

where TC_t is the transaction costs function with $\Delta\boldsymbol{\theta}_t = \boldsymbol{\theta}_t - \boldsymbol{\theta}_{t-1}$, as well as L constraints on the asset allocation,

$$
g_t^l(\boldsymbol{\theta}_t) \geq 0, \quad l = 1\ldots, L. \tag{3.1}
$$

With these constraints, the Bellman equation of the maximization problem is given by

$$
V_t(W_t, \theta_{1,t-1}^S, \ldots, \theta_{N,t-1}^S) = \max_{c_t, \boldsymbol{\theta}_t} \left\{u(c_t) + \beta \mathbb{E}_t^{\mathbb{P}}[V_{t+1}(W_{t+1}, \theta_{1,t}^S, \ldots, \theta_{N,t}^S)]\right\}.
$$

The allocation on the stocks at time $t-1$ is treated as state variables because it affects the transaction costs and hence the agent's decision making. On the other hand, since we assume that options and the risk-free bond are one-period assets, their past allocations do not enter as state variables. The first-order condition with respect to the allocation on i-th stock $\theta_{i,t}^S$ is given by

$$
S_{i,t} = \mathbb{E}_t^{\mathbb{P}}[m_{t,t+1}^*(S_{i,t+1} + D_{i,t+1})] + \sum_{l=1}^{L} \frac{\lambda_t^l}{u'(c_t)} \frac{\partial g_t^l}{\partial \theta_{i,t}^S} - \frac{\partial TC_t}{\partial \theta_{i,t}^S} + \frac{\beta}{u'(c_t)} \mathbb{E}_t^{\mathbb{P}}\left[\frac{\partial V_{t+1}}{\partial \theta_{i,t}^S}\right], \tag{3.2}
$$

where $m_{t,t+1}^* = \beta(\partial V_{t+1}/\partial W_{t+1})/u'(c_t)$ is the agent's intertemporal marginal rate of substitution (IMRS) between time t and $t+1$, and λ_t^l is the Lagrange multiplier of the constraint, equation (3.1). Throughout this Chapter, we assume that $m_{t,t+1}^*$ is almost surely positive. By transforming equation (3.2), HS derive the following asset pricing equation,

$$
\begin{aligned}
\mathbb{E}_t^{\mathbb{P}}[R_{i,t,t+1}] - R_{f,t,t+1} = -R_{f,t,t+1}Cov_t^{\mathbb{P}}(m^*_{t,t+1}, R_{i,t,t+1}) + CFER_{i,t,t+1}, \\
CFER_{i,t,t+1} = -\frac{R_{f,t,t+1}}{S_{i,t}}\left(\sum_{l=1}^{L}\frac{\lambda_t^l}{u'(c_t)}\frac{\partial g_t^l}{\partial \theta_{i,t}^S} - \frac{\partial TC_t}{\partial \theta_{i,t}^S} + \frac{\beta}{u'(c_t)}\mathbb{E}_t^{\mathbb{P}}\left[\frac{\partial V_{t+1}}{\partial \theta_{i,t}^S}\right]\right),
\end{aligned}
\tag{3.3}
$$

where $CFER_{i,t,t+1}$ is the *contribution of frictions to the expected return* of i-th stock, which represents the wedge between the expected stock return and its covariance risk premium term caused by constraints on the portfolio allocations (e.g., margin constraints) and the transaction costs.

Next, we define the "risk-neutral" probability measure $\mathbb{Q}^*$ by using $R_{f,t,t+1}m^*_{t,t+1}$ as the Radon-Nikodým derivative of the change of measure from the physical measure $\mathbb{P}$, so that the following change of measure relation holds:

$$
\mathbb{E}_t^{\mathbb{P}}[m^*_{t,t+1}X_{t+1}] = \frac{1}{R_{f,t,t+1}}\mathbb{E}_t^{\mathbb{Q}^*}[X_{t+1}], \tag{3.4}
$$

where X_{t+1} is any random variable measurable at time $t+1$. This is a unique well-defined probability measure because the IMRS is assumed to be positive almost surely. Note that we do not invoke the first fundamental theorem of asset pricing (FFTAP) to claim the existence of $\mathbb{Q}^*$. On the contrary, we cannot resort to FFTAP because we assume the market may not be frictionless. With this definition of the risk-neutral measure, equation (3.3) can be transformed into

$$
\mathbb{E}_t^{\mathbb{Q}^*}[R_{i,t,t+1}] - R_{f,t,t+1} = CFER_{i,t,t+1}. \tag{3.5}
$$

Equation (3.5) justifies to call $\mathbb{Q}^*$ the "risk-neutral" measure; equation (3.3) boils down to (3.5) in the case where the agent is risk-neutral in the sense that her IMRS is constant.

3.2.2 Formula for the Expected Return under market frictions

Equation (3.3) shows that the expected stock return is expressed as the sum of two terms: the covariance risk premium term and the CFER term. HS demonstrate that

CFER can be estimated reliably using an appropriately scaled deviations from put-call parity and the estimated CFER has a strong predictive power for the cross-section of stock returns. However, their study is silent about the covariance risk premium term. On the other hand, under the frictionless market assumption (hence absence of the CFER term), Martin and Wagner (2018) derive a formula in which the covariance risk premium term is expressed by the risk-neutral simple variance which can be estimated from the option price data.

In this Section, we generalize the Martin and Wagner (2018) model by allowing the presence of market frictions so that we can replace the covariance risk premium term in equation (3.3) by option-based variables.

Let us define $\xi_{t,t+1} = 1/m^*_{t,t+1}$, that is, $\xi_{t,t+1}$ is the inverse of the IMRS. Since $m^*_{t,t+1}$ is almost surely positive, $\xi_{t,t+1}$ is well-defined and positive almost surely. With this notation, the measure change formula, equation (3.4), yields

$$\begin{aligned}\mathbb{E}_t^{\mathbb{P}}[R_{i,t,t+1}] = \mathbb{E}_t^{\mathbb{P}}\left[m^*_{t,t+1}\frac{R_{i,t,t+1}}{m^*_{t,t+1}}\right] &= \frac{1}{R_{f,t,t+1}}\mathbb{E}_t^{\mathbb{Q}^*}\left[\xi_{t,t+1}R_{i,t,t+1}\right] \\ &= \frac{1}{R_{f,t,t+1}}\left(\mathbb{E}_t^{\mathbb{Q}^*}[\xi_{t,t+1}]\mathbb{E}_t^{\mathbb{Q}^*}[R_{i,t,t+1}] + Cov_t^{\mathbb{Q}^*}(\xi_{t,t+1}, R_{i,t,t+1})\right).\end{aligned}$$

Since $\mathbb{E}_t^{\mathbb{Q}^*}[\xi_{t,t+1}] = R_{f,t,t+1}\mathbb{E}_t^{\mathbb{P}}[m^*_{t,t+1}\xi_{t,t+1}] = R_{f,t,t+1}$ holds, the risk premium of i-th stock is given by

$$\mathbb{E}_t^{\mathbb{P}}[R_{i,t,t+1}] - \mathbb{E}_t^{\mathbb{Q}^*}[R_{i,t,t+1}] = \frac{1}{R_{f,t,t+1}}Cov_t^{\mathbb{Q}^*}(\xi_{t,t+1}, R_{i,t,t+1}). \tag{3.6}$$

Substituting equation (3.5) into (3.6) yields the following expression for the expected excess return

$$\mathbb{E}_t^{\mathbb{P}}[R_{i,t,t+1}] - R_{f,t,t+1} = CFER_{i,t,t+1} + \frac{1}{R_{f,t,t+1}}Cov_t^{\mathbb{Q}^*}(\xi_{t,t+1}, R_{i,t,t+1}). \tag{3.7}$$

Next, let us consider the following linear projection under $\mathbb{Q}^*$, that is, the projection of the random variable $R_{i,t,t+1}$ onto $\xi_{t,t+1}$ under $\mathbb{Q}^*$:

$$R_{i,t,s} = \alpha_i^* + \beta_i^*\xi_{t,t+1} + u_i^*. \tag{3.8}$$

Then, $\beta_i^* = Cov_t^{\mathbb{Q}^*}(\xi_{t,t+1}, R_{i,t,t+1})/Var_t^{\mathbb{Q}}(\xi_{t,t+1})$ and hence equation (3.7) can be transformed further into

$$\mathbb{E}_t^{\mathbb{P}}[R_{i,t,t+1}] - R_{f,t,t+1} = CFER_{i,t,t+1} + \beta_i^* \frac{Var_t^{\mathbb{Q}^*}(\xi_{t,t+1})}{R_{f,t,t+1}}. \tag{3.9}$$

Since $Var_t^{\mathbb{Q}^*}(\xi_{t,t+1})/R_{f,t,t+1}$ is the same across stocks, the cross-section of expected stock returns are determined by two elements: the CFER term and β_i^*.

Next, by taking the variance of equation (3.8), it follows that

$$Var_t^{\mathbb{Q}^*}(R_{i,t,t+1}) = (\beta_i^*)^2 Var_t^{\mathbb{Q}^*}(\xi_{t,t+1}) + Var_t^{\mathbb{Q}^*}(u_i^*). \tag{3.10}$$

Martin and Wagner (2018) employ a linear approximation for the squared beta term as $(\beta_i^*)^2 \approx 2\beta_i^* - 1$. With this approximation, we obtain

$$\beta_i^* Var_t^{\mathbb{Q}^*}(\xi_{t,t+1}) = \frac{1}{2}\left[Var_t^{\mathbb{Q}^*}(R_{i,t,t+1}) - Var_t^{\mathbb{Q}^*}(u_i^*) + Var_t^{\mathbb{Q}^*}(\xi_{t,t+1})\right] + v_{i,t}, \tag{3.11}$$

where $v_{i,t} = -(\beta_i^* - 1)^2 Var_t^{\mathbb{Q}^*}(\xi_{t,t+1})/2$ denotes the approximation error. Therefore, substituting equation (3.11) into (3.9) yields

$$\begin{aligned}\mathbb{E}_t^{\mathbb{P}}[R_{i,t,t+1}] - R_{f,t,t+1} = CFER_{i,t,t+1} + \frac{Var_t^{\mathbb{Q}^*}(R_{i,t,t+1})}{2R_{f,t,t+1}} + \frac{Var_t^{\mathbb{Q}^*}(\xi_{t,t+1})}{2R_{f,t,t+1}} \\ - \frac{Var_t^{\mathbb{Q}^*}(u_i^*)}{2R_{f,t,t+1}} + \frac{v_{i,t}}{R_{f,t,t+1}}.\end{aligned} \tag{3.12}$$

By introducing several notations as stated below, we obtain the following expression from equation (3.12), where we generalize Martin and Wagner's (2018) expected return expression by incorporating the effect of market frictions in terms of HS's CFER model.

Proposition 3.2.1. *Let $\omega_{i,t}^M$ be the value weight of i-th stock in the market index. Moreover, let us define $\overline{u}_t = \sum_i \omega_{i,t}^M Var_t^{\mathbb{Q}^*}(u_i^*)$ and $\phi_{i,t} = Var_t^{\mathbb{Q}^*}(u_i^*) - \overline{u}_t$. Similarly, let us define $\overline{v}_t = \sum_i \omega_{i,t}^M v_{i,t}$ and $\psi_{i,t} = v_{i,t} - \overline{v}_t$. Then, the expected excess stock return satisfies the following equation:*

$$\mathbb{E}_t^{\mathbb{P}}[R_{i,t,t+1}] - R_{f,t,t+1} = CFER_{i,t,t+1} + V_{i,t,t+1} + c_t + \varepsilon_{i,t}, \tag{3.13}$$

where

$$V_{i,t,t+1} = \frac{Var_t^{\mathbb{Q}^*}(R_{i,t,t+1})}{2R_{f,t,t+1}}, \tag{3.14}$$

c_t is constant across individual stocks and given by

$$c_t = \frac{Var_t^{\mathbb{Q}^*}(\xi_{t,t+1}) - \overline{u}_t + 2\overline{v}_t}{2R_{f,t,t+1}}, \tag{3.15}$$

and the idiosyncratic part $\varepsilon_{i,t,t+1}$ is given by

$$\varepsilon_{i,t} = -\frac{\phi_{i,t} - 2\psi_{i,t}}{2R_{f,t,t+1}}. \tag{3.16}$$

Equation (3.13) decomposes the expected excess return into four terms. The first CFER term represents the effect of frictions on the expected return. By comparing equations (3.3) and (3.13), we can see that the covariance risk premium term is expressed as the sum of the remaining three terms in equation (3.13). The first term among them, $V_{i,t,t+1}$, is a stock-specific yet observable part of the risk premium term. Indeed, the risk-neutral simple variance can be estimated based on the theory developed by Martin (2017) except that the estimation formula should be slightly modified in the presence of market frictions, as we will prove in Section 3.2.3. Next, c_t is part of the risk premium which is common across individual stocks. Finally, $\varepsilon_{i,t}$ is a stock-specific yet unobservable *residual* term; it is on average zero on value-weighted average basis because, by definition, both $\phi_{i,t}$ and $\psi_{i,t}$ are so (e.g., $\sum_i \omega_{i,t}^M \phi_{i,t} = 0$). Moreover, Martin and Wagner (2018) document that individual $\psi_{i,t}$ and $\phi_{i,t}$ terms are typically negligible. Given their findings, we ignore $\varepsilon_{i,t}$ throughout the remaining part of this Chapter.

Once the residual term $\varepsilon_{i,t}$ is ignored, the risk premium component is given by $V_{i,t,t+1} + c_t$, that is, individual stocks' risk premium component is given by the observable stock-specific term $V_{i,t,t+1}$ shifted by the constant c_t. Therefore, we call $V_{i,t,t+1}$ the *shifted risk premium* term. Since c_t is constant across stocks, equation (3.13) shows that cross-sectional variations in the expected stock returns are determined by variations in CFER and those in the shifted risk premium term. This means that we can estimate variations in the cross-section of expected stock returns from option price data in a model-free manner because both CFER and the shifted risk premium can be estimated from option price data.

Finally, we make three remarks regarding the difference between our results and those in Martin and Wagner (2018). First, they provide an option-based expression for the constant term c_t under an additional assumption. This enables them to express the *level* of the expected return by option-implied variables. On the other hand, we do not estimate the level of individual expected stock returns because, as we will show in Section 3.2.4, the information on the cross-sectional variation in the expected returns is sufficient for the mean-variance portfolio construction. Nonetheless, for completeness, we provide the following result which shows that Martin and Wagner's (2018) result on c_t is generalizable to our friction model.

Proposition 3.2.2. *If the inverse of the IMRS coincides with the market portfolio return, $\xi_{t,t+1} = R^M_{t,t+1}$, then, c_t is given by the following equation:*

$$c_t = \frac{Var_t^{\mathbb{Q}^*}(R_{M,t,t+1})}{R_{f,t,t+1} + CFER_{M,t,t+1}} + CFER_{M,t,t+1} - \sum_{i=1}^{N} \omega_{i,t}^M CFER_{i,t,t+1} - \sum_{i=1}^{N} \omega_{i,t}^M V_{i,t,t+1}. \tag{3.17}$$

Proof. See Appendix 3.A.1. □

Second, Martin and Wagner (2018) assume that the investor has a log utility in addition to the frictionless market assumption, whereas we do not assume the functional form of the IMRS. Under the log utility assumption, the inverse of the IMRS, $\xi_{t,t+1}$ coincides with the gross return of the growth optimal portfolio. This additional structure is not necessary for deriving Proposition 3.2.1 and hence for our subsequent mean-variance portfolio analysis.

Third, Martin and Wagner (2018) express the expected stock returns using the risk-neutral simple variance of the individual stocks as well as that of the market index. Therefore, their formula is applicable to the member stocks of the market index only. Our formula does not impose this constraint because the calculation of $CFER_{i,t,t+1}+V_{i,t,t+1}$ does not involve the option-implied variables of the market index. From the perspective of empirical analysis, this enables us to include stocks which do not belong to the major stock indices (e.g., the S&P 500 index) in the stock universe for mean-variance portfolio selection.

3.2.3 Estimation of the risk-neutral variance under frictions

In the frictionless setup, Martin (2017) develops a theory for estimating the risk-neutral simple variance or SVIX ("S" stands for "simple" indicating that it is the variance of the simple return $R_{i,t,t+1}$ instead of the log return). In particular, he shows that the risk-neutral simple variance discounted by the risk-free rate can be calculated from option prices as

$$\frac{Var_t^{\mathbb{Q}^*}(R_{i,t,t+1})}{R_{f,t,t+1}} = \frac{2}{S_{i,t}^2}\left[\int_0^{F_{i,t,t+1}} P_{i,t}(K,T)dK + \int_{F_{i,t,t+1}}^{\infty} C_{i,t}(K,T)dK\right], \tag{3.18}$$

where $F_{i,t,t+1} = R_{f,t,t+1}S_{i,t} - \widetilde{D}_{i,t,t+1}$ is the forward price of i-th stock with $\widetilde{D}_{i,t,t+1}$ being the time $t+1$ value of the dividend payments between t and $t+1$. We follow Martin (2017) and assume that the dividend payments between t and $t+1$ are known at time t. This is a plausible assumption at least for short-term horizon options (e.g., one month) which we use in the subsequent empirical analysis. In the case where market frictions are present, equation (3.18) should be modified as follows.

Proposition 3.2.3. *Let $F_{i,t,t+1} = R_{f,t,t+1}S_{i,t} - \widetilde{D}_{i,t,t+1}$ be the forward price of i-th stock. In the presence of market frictions, the risk-neutral simple variance divided by the risk-free rate is given by the following equation*

$$\frac{Var_t^{\mathbb{Q}^*}(R_{i,t,t+1})}{R_{f,t,t+1}} = \frac{2}{S_{i,t}^2}\left[\int_0^{F_{i,t,t+1}} P_{i,t}(K,T)dK + \int_{F_{i,t,t+1}}^{\infty} C_{i,t}(K,T)dK\right] - \frac{CFER_{i,t,t+1}^2}{R_{f,t,t+1}}. \tag{3.19}$$

Proof. See Appendix 3.A.2 □

Proposition 3.2.3 shows that the formula for the risk-neutral simple variance should be modified to include the squared CFER term. This modification stems from the fact that the mean of the $\mathbb{Q}^*$-expected stock return is different from the risk-free rate; it is now the sum of the risk-free rate and CFER (equation (3.5)). Since the mean changes, the variance (i.e., the central second moment) should also change. Moreover, when CFER is zero, equation (3.19) reduces to (3.18), meaning that equation (3.19) is the generalization of equation (3.18) to accommodate the possible presence of market frictions.

3.2.4 Shifted expected return and mean-variance optimization

Let us consider the following mean-variance portfolio problem,

$$\min_{\omega} \frac{\gamma}{2}\omega'\Sigma\omega - \omega'\mu, \quad s.t.\ \omega'e = 1, \tag{3.20}$$

where Σ and μ are the covariance matrix and the expected returns of the individual stocks, respectively, γ is the risk-aversion parameter, e is the vector of ones, and x' denotes the transpose of a vector or matrix x.

To solve this optimization problem, equation (3.20), we wish to use the option-based estimate of the expected stock return, $CFER_{i,t,t+1} + V_{i,t,t+1}$, as an estimate of μ_i. The following Lemma ensures that this is approach is valid, that is, we can ignore the constant shifting term c_t in the expected return and use the shifted expected return as an input to the mean-variance optimization problem.

Lemma 3.2.1. *Let ω^* be the solution to the minimization problem (3.20) given the covariance matrix Σ and the expected stock returns μ. Then, the same ω^* is the solution to the minimization problem (3.20) given the covariance matrix Σ and the parallel-shifted expected stock returns $\widetilde{\mu} = \mu + ke$, where k is an arbitrary constant.*

Proof. See Appendix 3.A.3. □

Lemma 3.2.1 implies that we will obtain the same optimal mean-variance portfolio weights regardless of using $\mu_i = CFER_{i,t,t+1}+V_{i,t,t+1}+c_t$ or $\widetilde{\mu}_i = CFER_{i,t,t+1}+V_{i,t,t+1}$. In other words, Lemma 3.2.1 allows us to ignore the cross-sectional constant term c_t when we construct an estimate of the expected return input for the mean-variance problem. Therefore, we can obtain sufficient information on the expected stock returns from only individual stocks' CFER and simple risk-neutral variance. This estimated (shifted) expected returns inherit the advantageous property of SVIX and CFER; it is a forward-looking, real-time measure of the (shifted) expected return and its estimation does not require any parameter estimations nor historical data.

3.3 Data and estimation of option-implied measures

In this Section, we describe the data on stocks and options we use in our subsequent empirical analysis. Then, we explain the estimation procedures for option-implied variables.

3.3.1 Data

We obtain the option data from the OptionMetrics Ivy DB (OM) via the Wharton Research Data Services (WRDS). Our sample period is from January 1996 to December 2017. We estimate CFER and the shifted risk premium as well as the model-free implied volatility (MFIV), risk-neutral skewness (RNS) and volatility risk premium (VRP) from the OM option data. We explain the estimation procedures for CFER, shifted risk premium, MFIV and RNS in Section 3.3.2. VRP is calculated as the difference between the 30-day MFIV and the 30-day historical volatility, where we obtain the historical volatility from the OM Historical Volatility file.

We obtain the stock data from the Center for Research in Security Prices (CRSP) via the WRDS. We construct size, log book-to-market (logBM) and the 12-month momentum (MOM) from the CRSP and Compustat. Size is calculated as the logarithm of the market capitalization, which is the product of the close stock price and the shares outstanding. We calculate logBM based on Davis et al. (2000) and we treat negative book-to-market samples as missing. Regarding MOM, we follow DPUV and calculate the MOM at time t as the cumulative return from day $t - 251 - 21$ to day $t - 21$. See Appendix 3.B.1 for detailed explanations on the variable construction procedures. Appendix 3.B.2 explains how we link CRSP, Compustat and OM. Finally, we obtain standard risk factors from Kenneth French's data library.

Our baseline stock universe for the subsequent mean-variance portfolio analysis is all the constituent stocks of the S&P 500 index on a particular end-of-month trading day for which we can estimate the option-implied measures (CFER, the risk-neutral simple variance and other option-implied measures). We obtain the information on the constituent stocks of the S&P 500 index from Compustat. In each month, there are on average 474 (at least 453) S&P 500 constituent stocks for which we can estimate the option-implied measures. For the purpose of estimating historical moments, we further restrict our universe to stocks with non-missing return data over the past 630 trading days (about 30 months). This window length is chosen in order to ensure the positive definiteness of historical covariance matrices while minimizing the number of stocks excluded due to missing past return data. Since the constituent stocks of the S&P 500 index are actively traded large stocks, this additional restriction does not have big impact; on average only 10 stocks are excluded from the S&P 500 universe each month due to missing past return data.

In addition, we consider an alternative stock universe of 500 randomly selected stocks. For each year y, first we find stocks which satisfy the following two criteria: (i) the option-implied measures are not missing for all the end-of-month trading days in year y, and (ii) on each end-of-month day in year y, the past 630 daily returns are not missing so that we can calculate the historical moments. Then, we randomly select 500 stocks from the pool of eligible stocks. We fix the selected 500 stocks throughout year y (i.e., the stock universe is updated in every January). On average, for each year, there are about 1,700 (at least 1,411) eligible stocks which satisfies the two screening criteria described above. This universe is constructed in the spirit similar to the randomly selected 500 CRSP stock universe considered in Jagannathan and Ma (2003) and DeMiguel et al. (2009a) among others, except that we impose an additional natural eligibility condition that option-implied measures are non-missing.

3.3.2 Estimation of option-implied measures

3.3.2.1 CFER

For the estimation of CFER, HS show that it can be reliably estimated by using a pair of call and put options at the same strike and maturity as follows:

$$CFER_{i,t,T} \approx \frac{R_{f,t,T}}{S_{i,t}} \left(\widetilde{S}_{i,t}(K) - S_{i,t} \right), \tag{3.21}$$

where $\widetilde{S}_{i,t}(K) = C_{i,t}(K,T) - P_{i,t}(K,T) + (K + \widetilde{D}_{i,t,T})/R_{f,t,T}$ is the price of the synthetic stock price and $\widetilde{D}_{i,t,T}$ is the time T value of dividend payments during the life of the options. We estimate CFER based on the *OM Standardized Option Price file*, which provides the call and put prices and respective implied volatilities (IVs) at the forward at-the-money (ATM) strike price (i.e., $K = F_{i,t,T} = R_{f,t,T}S_{i,t} - \widetilde{D}_{i,t,T}$) for standardized maturities such as 30, 60, 91, 182, and 365 days to maturities.[6] We use the 30-day-to-maturity forward ATM call and put option prices to calculate equation (3.21). See Appendix 3.B.3 for more detail.

[6]The OM calculates the forward prices based on its zero yield curve data (available in the OM Zero Curve File) and its projected dividend payments. Specifically, the forward prices recorded in the OM file are not actual market transaction or quote prices.

3.3.2.2 Risk-neutral simple variance and risk-neutral moments

For the estimation of the risk-neutral simple variance, we use equation (3.19). The estimation of MFIV and RNS is based on the Bakshi et al. (2003) formulae corrected for the dividend payments; as we have seen in Chapter Two, the original formulae in Bakshi et al. (2003) are derived under the assumption that the underlying pays no dividends. Specifically, we calculate risk-neutral moments as

$$MFIV_{t,T} = \sqrt{\frac{e^{r_f\tau}M(2)_{t,T} - \mu_{t,T}^2}{\tau}}, \tag{3.22}$$

$$RNS_{t,T} = \frac{e^{r_f\tau}M(3)_{t,T} - 3e^{r_f\tau}\mu_{t,T}M(2)_{t,T} + 2\mu_{t,T}^3}{[e^{r_f\tau}M(2)_{t,T} - \mu_{t,T}^2]^{3/2}}, \tag{3.23}$$

where $M(n)_{t,T}$ $(n = 2, 3, 4)$ is given by

$$M(n)_{t,T} = \int_{S_t-\widetilde{D}_{t,T}}^{\infty} \eta(K; S_t, n)C_t(K,T)dK + \int_0^{S_t-\widetilde{D}_{t,T}} \eta(K; S_t, n)P_t(K,T)dK, \tag{3.24}$$

$$\eta(K; S_t, n) = \frac{n}{(K+\widetilde{D}_{t,T})^2}\left[(n-1)\log\left(\frac{K+\widetilde{D}_{t,T}}{S_t}\right)^{n-2} - \log\left(\frac{K+\widetilde{D}_{t,T}}{S_t}\right)^{n-1}\right]. \tag{3.25}$$

and $\mu_{t,T} = R_{f,t,T} - 1 - R_{f,t,T}\,[M(2)_{t,T}/2 + M(3)_{t,T}/6 + M(4)_{t,T}/24]$.

To calculate the integrals in equations (3.19) and (3.24), a continuum of option prices with respect to the strike price is required. To this end, we interpolate the IVs provided by the *OM Volatility Surface* file. The Volatility Surface file provides call and put option prices and IVs at standardized maturities (such as 30, 60, 91, 182, and 365 day-to-maturity) and standardized delta-denominated strikes (deltas 0.2, 0.25, . . . , 0.8 for call options and -0.2, -0.25, . . . , -0.8 for put options). We follow the convention in the recent literature (e.g., Stilger et al., 2017; Borochin and Zhao, 2018) and separately interpolate the call and put IVs by the cubic Hermite polynomial in the moneyness-IV metric. We extrapolate the IV curve horizontally beyond the lowest and the highest moneyness to obtain 1,001 IVs at equally spaced strikes over the moneyness range from 1/3 to 3. Then, we convert these IVs back to European option prices and numerically calculate the integrals. The boundary value between the put and call integrals in equation (3.19) is $F_{i,t,t+1}$ and that in equation (3.24) is $S_t - \widetilde{D}_{t,T}$. We use the forward price recorded in the OM Standardized Option Price file for equation (3.19). Regarding the boundary value in equation (3.24), we back out $\widetilde{D}_{t,T}$ from the same OM forward price as $\widetilde{D}_{t,T} = R_{f,t,T}S_{i,t} - F_{i,t,T}$. See Appendix 2.B.4 for more detail.

Two remarks are in order regarding the option dataset we employ. First, our choice of the data source is in line with the previous literature; DPUV and Martin and Wagner (2018) estimate MFIV and RNS, and the stock risk premium, respectively, using the OM Volatility Surface file. Second, exchange listed individual U.S. equity options are American options, although the theoretical formulae for CFER, the simple variance, MFIV, and RNS are based on European option prices. Regarding this point, we follow the convention in the literature (e.g., Martin and Wagner, 2018) and regard IVs provided by OM as those of European options because OM calculate IVs by taking early exercise premium into account by relying on the binomial tree model.

3.3.3 CFER and the shifted risk premium: Estimation result

We estimate CFER and the shifted risk premium term implied by the risk-neutral simple variance, $V_{i,t,t+1}$, for each individual stock at the end of each month from January 1996 to December 2017.

Table 3.1 Panel A reports the summary statistics of the estimated CFER, where the unit of CFER is percent per year. It shows that the estimated CFER exhibits large variation. For example, in the all optionable stocks universe, the standard deviation and the inter-quartile range (IQR; the difference between 75th and 25th percentile points) are 15.9% and 5.58%, respectively. These values are economically large and hence the variation in the CFER component constitutes an essential element of the variation in the expected stock returns. The variation in CFER is large in the S&P 500 stocks universe as well (e.g., the standard deviation is 4.8%), even though the variation is milder compared to the all optionable stock universe. This pattern is sensible because larger stocks tend to face less market frictions. Nevertheless, the summary statistics suggest that the CFER component is of importance even for large S&P constituent stocks. The mean and median of the estimated CFER are close to zero yet slightly negative (-82 bps and -37 bps per year, respectively, in the all optionable stock universe). Therefore, CFER takes positive and negative value with roughly equal probability.

Table 3.1 Panel B shows the summary statistics of the estimated shifted risk premium. This term exhibits even larger variation compared to the CFER component; in the all optionable stock universe, its standard deviation and IQR is 22.0% and 14.8%, respectively. Similar to CFER, the variation in the S&P 500 universe is smaller compared to the all optionable stock universe yet non-negligible. In words, these summary statistics suggest that both the two component of the expected return, CFER and the risk premium term, exhibit large variation across individual stocks and over time.

[Table 3.1 about here.]

Next, we investigate how the two components of the expected return, CFER and the risk premium term, contribute to the variation in the expected returns. The cross-sectional variation in the expected returns can be decomposed into the variation in CFER, that in the risk premium term and their cross variation term, that is,

$$Var(CFER + V) = Var(CFER) + Var(V) + 2Cov(CFER, V).$$

To obtain the variance decomposition result, first for each month t, we conduct the variance decomposition, then we take the time-series average of each component.

Table 3.1 Panel C reports the result. We can see that the CFER component explains about one third of the total variation in the estimated expected returns and remaining two thirds of the variation is explained by the risk premium term. The cross variation term explains the total variation by at most a few percent points, implying that the variations in the CFER and the risk premium component are almost "orthogonal." The contribution of the CFER component is slightly smaller for the S&P 500 stock universe compared to the full optionable stock universe. This is sensible because larger stocks are subject to less market frictions and hence the role of the effect of market frictions is less pronounced. However, this result suggests that still non-negligible proportion of the variation in large stocks is attributable to the effect of market frictions.[7]

[7]This result may explain why Martin and Wagner (2018) find that their theory performs well only for estimated risk premia for longer term horizons (6-month or longer); the effect of frictions have non-negligible effect for short horizon expected returns.

3.4 Mean-variance portfolio analysis: Empirical setup

In this Section, we explain the mean-variance portfolio strategies we examine subsequently, as well as a number of benchmark portfolio strategies which do not rely on the mean-variance optimization. Then, we explain performance measures we employ for comparing the performance of various strategies.

3.4.1 Mean-variance portfolio: expected stock returns

A mean-variance portfolio strategy is specified by the combination of three elements: the expected stock returns, covariance matrix, and the constraint on the portfolio weights. We begin with our innovative part, the specification of the expected return vector, μ_t.

Based on our theoretical model in Section 3.2, we consider three forward-looking *quantitative* estimates of expected returns. For the first expected returns specification, we set $\mu_{i,t} = CFER_{i,t,t+1} + V_{i,t,t+1}$ based on Proposition 3.2.1 and Lemma 3.2.1. We call this return specification Q-FULL (Q stands for "quantitative") because this specification utilizes both the option-implied risk premium and the effect of frictions. We also consider two partial measures of the expected returns, $\mu_{i,t} = CFER_{i,t,t+1}$ and $\mu_{i,t} = V_{i,t,t+1}$. We call the former specification Q-CFER and the latter Q-RP (RP stands for "risk premium"). These two specifications would help us to better understand the role of the two components of the expected return (CFER and the risk premium) on the mean-variance portfolio performance by shutting down either one of the two components.

To expedite the comparison between our results and those in the previous literature, we also examine three alternative setting for the expected returns. First, we examine the global minimum variance portfolio (GMVP), that is, $\mu_{i,t} = 0$. For completeness, we also examine the historical sample mean as the second alternative choice. The third one is the *characteristic-adjusted returns* considered in DPUV. They construct characteristic-adjusted expected returns based on the following definition:

$$\mu_{i,t}^{k} = \mu_{\mathrm{B},t}\left(1 + \delta_t^k x_{i,t}^k\right),$$

where $\mu_{\mathrm{B},t}$ is the benchmark expected return, $x_{i,t}^k$ is an indicator variable which takes +1 (-1) when the k-th characteristic variable (e.g., RNS) belongs to the top (bottom) decile of time t cross-sectional values and takes zero otherwise, and δ_t^k controls the intensity of the effect of the characteristic k on the conditional mean. We follow DPUV to use the time t grand mean return across all individual historical returns as the benchmark expected return and $\delta_t^k = 0.1$ for any characteristic. Under this setting, the expected returns take one of the following three values depending on the cross-sectional rank of k-th characteristic:

$$\mu_{i,t}^k = \begin{cases} 1.1 \times \mu_{\mathrm{B},t} & \text{if } k\text{-th characteristic belongs to the top decile,} \\ 0.9 \times \mu_{\mathrm{B},t} & \text{if } k\text{-th characteristic belongs to the bottom decile,} \\ 1.0 \times \mu_{\mathrm{B},t} & \text{otherwise.} \end{cases}$$

We construct four characteristic-adjusted returns based on MFIV, RNS, VRP, and CFER. We label these specifications CA-MFIV, CA-RNS, CA-VRP, and CA-CFER, respectively, where CA stands for "characteristic-adjusted." Our choice of option-implied characteristics are largely in line with DPUV, who examine MFIV, RNS, VRP, and IVS (implied volatility spread). We use the estimated CFER instead of IVS because these two variables are approximately proportional and hence exhibit a nearly perfect rank correlation.[8]

3.4.2 Mean-variance portfolio: covariance matrix

We examine a number of covariance matrix specifications. We estimate the variance of stocks and the correlation separately and then form the covariance matrix Σ as $\Sigma = V'\Gamma V$, where V is the diagonal matrix of the estimated volatility and Γ is the correlation matrix.

[8]HS show that their estimate of CFER, the scaled deviations from put-call parity, is approximately equal to IVS times the Black-Scholes vega-to-stock price ratio. Since the vega-to-stock price ratio is roughly constant when the time-to-maturity and moneyness are the same, the estimated CFER calculated from the 30-day forward ATM call and put pair (as we do in this Chapter) and the IVS calculated from the 30-day ATM implied volatilities (as DPUV do) are roughly proportional in each cross-section. Note that this does not contradict the finding in HS that CFER has superior return predicting power compared to IVS. In HS and early studies on IVS (e.g., Bali and Hovakimian, 2009; Cremers and Weinbaum, 2010), they calculate CFER or IVS by aggregating the actual option quotes data with diverse maturities and moneyness recorded in the OM Option Price File. In this procedure, the predictive power of CFER and IVS may differ because the vega scaling factor is not constant across input option data with various strikes and maturities.

3.4.2.1 Estimation of the correlations

For the correlations, we examine the following two approaches. First, we simply use the historically estimated correlations (HC). As explained in Section 3.3.1, we use the past 630 trading days window for both the S&P 500 and the randomly selected 500 stock universe. Moreover, we follow DPUV and consider zero correlations (ZC) setup, where we set all the correlation coefficients to zero. Under the ZC specification, the resulted covariance matrix becomes a diagonal matrix of the estimated variances.[9]

3.4.2.2 Estimation of the variance

We examine the following three specifications for the variance of individual stocks. First, we use the historical variance (HV). The estimation window is the same as for the correlations (therefore, the combination HV-HC is the historical sample covariance matrix).

Next, we follow DPUV and construct the risk-adjusted option-implied variance (RIV). RIV is an option-implied forward-looking volatility estimate under the physical probability measure. We follow DPUV and "risk-adjust" MFIV to obtain the RIV as follows.[10] First, we calculate the historical volatility risk premium adjustment (HVRP) at time t for i-th stock as

$$\mathrm{HVRP}_{i,t} = \frac{\sum_{\tau=t-w+1}^{t} \mathrm{MFIV}_{i,\tau-30,\tau}}{\sum_{\tau=t-w+1}^{t} \mathrm{HV}_{i,\tau-30,\tau}},$$

where $\mathrm{MFIV}_{i,\tau-30,\tau}$ ($\mathrm{HV}_{i,\tau-30,\tau}$) denotes the 30-day MFIV (historical volatility) of i-th stock, that is, the implied (historical) volatility for the 30-day period ending at time τ. In short, we calculate the average value of 30-day implied and historical volatility over past w trading days and then calculate HVRP as the ratio of these two average volatility values. We follow DPUV and set $w = 252$. Given the estimated HVRP, we

[9]There are two alternative ways to estimate option-implied correlations. The first one is the heterogeneous implied correlation matrix corrected for the risk premium considered in DPUV. We do not examine this approach because they document that there are little gain from using this option-implied correlations. The second approach is the option-implied dispersion in Martin and Wagner (2018). They show that the ratio of the index SVIX to the value-weighted SVIX of the constituent stocks of the index provides an average option-implied correlation (dispersion), that is $\rho_t \approx \mathrm{SVIX}_t^2/\overline{\mathrm{SVIX}}_t^2$. We do not report this result neither because using this correlation does not improve the portfolio performances.

[10]We also calculate the risk-adjusted option-implied variance by using the risk-neutral simple variance instead of MFIV. However, we do not report this result because portfolio performance do not change qualitatively compared to those using the RIV based on MFIV.

calculate the predicted future realized volatility, $\widehat{\mathrm{RV}}_{i,t,t+30}$ as

$$\widehat{\mathrm{RV}}_{i,t,t+30} = \frac{\mathrm{MFIV}_{i,t,t+30}}{\mathrm{HVRP}_{i,t}}.$$

Finally, we also examine the shrinkage covariance matrix based on the Ledoit and Wolf (2004) method:

$$\widehat{\Sigma}^{shrink} = (1-\nu)\widehat{\Sigma} + \nu\Sigma^{target},$$

where ν is the shrinkage constant. We apply the shrinkage method to the historical covariance matrix (i.e., $\widehat{\Sigma}$ is set to the historical covariance matrix). We choose the scaled identity matrix $\Sigma^{target} = \overline{\sigma^2} I$ as the shrinkage target matrix, where $\overline{\sigma^2}$ is the average of the individual stocks' variances. We follow Ledoit and Wolf (2004) to calculate the shrinkage constant ν. We label this covariance matrix "LWI."

3.4.3 Mean-variance portfolio: constraints on portfolio weights

We examine the effect of constraints on portfolio weights in the mean-variance portfolio optimization. Specifically, for each stock weight ω_i, we examine the following form of constraints,

$$\omega^{LB} \leq \omega_i \leq \omega^{UB}, \tag{3.26}$$

with either one of the following choices for the lower and upper bound:

1. $(\omega^{LB}, \omega^{UB}) = (-\infty, \infty)$: unconstrained problem.
2. $(\omega^{LB}, \omega^{UB}) = (0, \infty)$: mo short-sale (NS) constraints.
3. $(\omega^{LB}, \omega^{UB}) = (0, 0.1)$: no short-sale and 10% maximum weight constraints (NS10).
4. $(\omega^{LB}, \omega^{UB}) = (-0.1, 0.1)$: 10% absolute portfolio weight constraints (ABS10).

The first setup corresponds to the unconstrained problem. The second case, short-sale constraints, are widely employed in the literature (e.g., Jagannathan and Ma, 2003; DeMiguel et al., 2013). The third specification, NS10, considers a realistic portfolio policy of mutual funds; they are not allowed to short-sell stocks. In addition, they are typically subject to the restriction on the maximum weight on each single stock, by their investment policy or law enforcement (see e.g., Hlouskova and Lee, 2001). The last setup considers a situation where short-sale is allowed, yet there is a restriction on the maximum absolute weight on each single stock. One may regard the last constraint as a restriction on a long-short equity fund which can take short positions yet not allowed to take an extreme position on a single stock.

3.4.4 Mean-variance portfolio: risk-aversion parameter

To solve the mean-variance portfolio problem, equation (3.20), we need to specify the risk-aversion parameter γ. To this end, we examine $\gamma = 2, 5, 10$. Since we obtain qualitatively similar results regardless of the choice of γ, we report the result of $\gamma = 10$ for brevity. Note that these risk-aversion parameter values are also examined in DeMiguel et al. (2009a) and DeMiguel et al. (2019).

3.4.5 Other benchmark strategies

For the sake of comparison, we consider a number of benchmark strategies examined in the literature. First, we examine the equally-weighted (EW) and the value-weighted (VW) portfolios. These two benchmarks are of importance. DeMiguel et al. (2009b) show that the EW portfolio has an excellent performance even though it does not involve any optimizations. The VW portfolio is also a natural benchmark since the VW portfolio approximately equals the market portfolio.[11]

Moreover, we consider the following two approaches which do not involve the mean-variance optimization, yet determine the portfolio weight by utilizing option-implied expected return information. The first approach is the rank-based weight approach. Namely, the weight of i-th stock at time t is given by

$$\omega_{i,t} = \frac{(\text{rank}[\mu_{i,t}])^a}{\Sigma_i(\text{rank}[\mu_{i,t}])^a}, \tag{3.27}$$

where $\text{rank}[\mu_{i,t}]$ denotes the rank of the i-th stock's expected return measure across the cross-section of time t expected stock returns. The rank is determined in the

[11] For example, the VW portfolio constructed within the S&P 500 universe is approximately equal to the S&P 500 index. This relation is approximate because there are a few S&P 500 member stocks for which we cannot estimate option-implied measures, even though the approximation is fairly accurate.

increasing order so that the stock with the highest expected return commands the largest weight. The parameter a controls the aggressiveness of the strategy; the higher a is, the more tilted the portfolio is toward stocks with higher expected returns. On the contrary, $a = 0$ corresponds to the EW portfolio. This rank-based approach is proposed by Asness et al. (2013) and Martin and Wagner (2018) apply this approach to utilized their option-based estimate of the stock risk premium. We follow Martin and Wagner (2018) and set $a = 2$. We consider the rank-based weights based on CFER, the shifted risk premium term, and the sum of these two. We label these strategies, RK-CFER, RK-RP, and RK-FULL, respectively.

The second approach we consider is the *parametric portfolio* approach with short-sale constraints. The parametric portfolio approach is developed by Brandt et al. (2009) and DPUV apply this approach with short-sale constraints to exploit information in option-based return predicting characteristics. The unconstrained parametric portfolio approach determines the asset allocation based on the value of characteristic variables:

$$\omega_{i,t}^{PP}(\boldsymbol{\varphi}_t) = \omega_{i,t}^{\mathrm{B}} + \frac{1}{N_t}\sum_{k=1}^{K} \varphi_t^k x_{i,t}^k, \tag{3.28}$$

where $\omega_{i,t}^{\mathrm{B}}$ is the benchmark allocation, N_t is the number of stocks in the time t investment universe, $x_{i,t}^k$ is the k-th characteristic variable of i-th stock at time t, φ_t^k is the loading on the characteristic k at time t, and $\boldsymbol{\varphi}_t$ is the vector of φ_t^k $(k = 1, \dots, K)$. The characteristic variables $x_{i,k}^k$ are standardized cross-sectionally to have zero mean and unit variance. See Brandt et al. (2009) for more detail. To impose the short-sale constraints, Brandt et al. (2009) proceed as follows. First, we calculate the parametric portfolio weight given by equation (3.28). Then, we truncate the negative weight and rescale the truncated weights so that they add up to one:

$$\omega_{i,t}^{PPNS}(\boldsymbol{\varphi}_t) = \frac{\max(0, \omega_{i,t}^{PP}(\boldsymbol{\varphi}_t))}{\sum_{i=1}^{N_t} \max(0, \omega_{i,t}^{PP}(\boldsymbol{\varphi}_t))}. \tag{3.29}$$

The optimal loadings $\boldsymbol{\varphi}_t$ for the short-constrained parametric portfolio are chosen by solving the following optimization problem

$$\max_{\boldsymbol{\varphi}} \frac{1}{w} \sum_{j=t-w}^{t-1} u\left(\sum_{i=1}^{N_j} \omega_{i,j}^{PPNS}(\boldsymbol{\varphi}) \cdot (R_{i,j,j+1} - 1)\right), \tag{3.30}$$

where w is the length of window period and $u(\cdot)$ is a utility function (e.g., the mean-variance utility and the power utility). We follow DPUV and set the benchmark portfolio to the EW portfolio (i.e., $\omega^{\mathrm{B}}_{i,t} = 1/N_t$), set the rolling window length w to 250 days, and set the utility function $u(\cdot)$ to the mean-variance utility. Regarding the characteristic variables, we start with the same set of variables considered in Brandt et al. (2009) and DPUV: size, logBM and MOM (FFM). Then, we follow DPUV and examine the parametric portfolios using either one of MFIV, RNS, VRP, or CFER in addition to the three FFM variables. We also consider the parametric portfolio strategy which includes all four option-implied variables and the three FFM variables. These six parametric portfolio strategies are labeled as PPNS-FFM, PPNS-MFIV, PPNS-RNS, PPNS-VRP, PPNS-CFER, and PPNS-ALL, respectively.

To wrap up, Table 3.2 summarizes the portfolio strategies we examine. Panel A shows the portfolio strategies which do not involve the mean-variance optimization, equation (3.20). Panel B lists the three components of the mean-variance portfolio problem, the expected return, covariance matrix, and constraints on portfolio weights, separately.

[Table 3.2 about here.]

3.4.6 Performance measures

We employ the following measures for the evaluation of the performance of portfolios. First, we calculate the annualized out-of-sample (OOS) excess return, volatility and the Sharpe ratio of portfolio returns. We use the risk-free rate provided by Kenneth French's data library for the calculation of excess returns. Moreover, although it is not common in the literature, we report risk-adjusted returns (alphas) of portfolios' OOS returns with respect to the Fama and French (1993) three-factor model and the Fama and French (2018) five-factor model.

We also report the certainty equivalent of the OOS returns,

$$CE = u^{-1}\left(\frac{\sum_{t=1}^{T} u(1 + r_{t-1,t})}{T}\right) - 1,$$

where $r_{t-1,t}$ is the net OOS return of the portfolio. We calculate the certainty equivalent based on the power utility with the relative risk aversion parameter being equal to five, that is, $u(x) = -x^{1-\gamma}/(1-\gamma)$ and $\gamma = 5$. Even though the mean-variance agent cares about only the first two moments, the certainty equivalent is a convenient way to incorporate the higher-order moments of the OOS returns into the performance evaluation.

Finally, we calculate the turnover of each portfolio as follows:

$$\text{TO} = \frac{1}{T}\sum_{t=1}^{T}\sum_{i=1}^{N}|\omega_{i,t+1} - \omega_{i,t+}|,$$

where $\omega_{i,t+}$ is the weight of i-th stock just before the rebalancing at time $t+1$. This weight differs from $\omega_{i,t}$ because the stock price fluctuations between t and $t+1$ affect the dollar weight on each stock. We define TO as the average turnover rate per rebalance. For example, TO = 1 means that an investor rebalances her portfolio by the same amount as their total investment amount while each rebalancing.

3.5 Estimation results

3.5.1 Baseline results: Choice of the expected return specification

In this subsection, we report our baseline results of the performance of the mean-variance portfolios. We focus on how the choice of the expected return specification affects the portfolio performance. Specifically, we consider nine specifications for the expected returns vector described in Section 3.4.1: the global minimum variance portfolio (GMVP, $\mu = 0$), the historical sample mean, three quantitative forward-looking expected returns based on CFER (Q-CFER), the shifted risk premium (Q-RP), and the sum of these two (Q-FULL), respectively, four characteristic adjusted returns based on MFIV (CA-MFIV), volatility risk premium (CA-VRP), risk-neutral skewness (CA-RNS), and CFER (CA-CFER). To highlight the effect of the specification of the expected returns, we use the historical sample covariance matrix throughout this subsection. We discuss the choice of the covariance matrix in Section 3.5.2.

Throughout this Section, the performance measures are calculated based on the OOS portfolio returns from January 1997 to December 2017 (252 months) because the first 12 months of our data period are dropped for the calculation of the risk-adjusted option-implied variance and the loadings of the parametric portfolio strategies.

3.5.1.1 Unconstrained strategies

First, we report the performance of unconstrained strategies. Table 3.3 Panel A reports the performance measures of EW, VW, and the three rank-based strategies. The Sharpe ratio of EW and VW strategies are 0.58 and 0.48, respectively. Moreover, the turnover rate of these two strategies are fairly low, 0.08 for EW and 0.02 for VW, respectively. These two strategies' returns do not yield a significant abnormal returns. The three rank-based strategies earn an OOS return and the Sharpe ratio similar to the EW portfolio, yet these strategies result in a higher turnover rate. This high turnover rates reflect the fact that both CFER and the risk premium term exhibit large time-series variations; the ranking of the expected returns also changes frequently over time.

[Table 3.3 about here.]

Table 3.3 Panel B reports the performance of the mean-variance portfolios based on the nine alternative specifications of the expected return vector. The first row of Panel B shows that GMVP during our sample period performs worse than EW and VW. The remaining eight strategies have a large OOS portfolio volatility, a huge negative certainty equivalent as well as an extremely high turnover rate. This is due to the well-documented empirical properties of mean-variance portfolio strategies that they tend to result in an extreme long-short portfolio and that the portfolio allocation is not stable over time. Specifically, the second row of Panel B shows that the mean-variance portfolio based on the historical mean sample performs poorly in line with the previous literature.

Albeit the high volatility and turnover, our result suggests that the estimated CFER is helpful to improve the OOS portfolio average return and the Sharpe ratio. The Q-CFER strategy earns a sufficiently high OOS portfolio return so that the Sharpe ratio comfortably surpasses that of the benchmark EW and VW portfolios, even though the OOS portfolio volatility is extremely high. The other two strategies which involve CFER, Q-FULL and CA-CFER, also earn the Sharpe ratio higher than the EW portfolio. On the other hand, using the shifted risk premium, MFIV, and

VRP do not improve the Sharpe ratio. Unfortunately, even for the Q-CFER strategy, the extremely high volatility and turnover rate make it impossible to execute this strategy in practice.

3.5.1.2 Constrained strategies

Next, we report the mean-variance portfolio results with constraints on portfolio weights. We consider the three types of constraints on portfolio weights: NS, NS10, and ABS10 (see Section 3.4.3). Table 3.4 Panel A reports the performance of the benchmark strategies which do not rely on the mean-variance optimization (EW, VW and the short-sale constrained parametric portfolios). The short-sale constrained parametric portfolios earn the Sharpe ratio and the certainty equivalent higher than the EW and VW portfolios. This result is in line with DPUV, who document that using option-implied information in the parametric portfolio approach enhances the portfolio performance.

Panels B to D report the performance of the mean-variance portfolios with NS, NS10, and ABS10 constraints, respectively. We can see from Panels B and C that imposing the short-sale constraints (and the upper bound for Panel C) on the portfolio weights improves the portfolio performance in various aspects. First, the OOS volatility decreases to the similar level of EW and VW portfolios except the two strategies which involve the scaled risk premium (i.e., Q-RP and Q-FULL). Second, as a result, the Sharpe ratio and the certainty equivalent improve drastically. Specifically, the Q-CFER strategy with NS10 constraint earns the Sharpe ratio 0.98 and the certainty equivalent 11.0%, which are much higher than the respective performance measures of EW and VW strategies. Third, the turnover rate decreases dramatically compared to the unconstrained portfolios, albeit still higher than those of EW and VW. We revisit the high turnover and related transaction costs issues in Section 3.5.5.

In the case where ABS10 constraint is imposed (Panel D), the resulted portfolio performance is mixed. On the one hand, the Sharpe ratio and the certainty equivalent of some strategies become even higher compared to the case where NS10 constraint is imposed. Specifically, the Sharpe ratio and the certainty equivalent of Q-CFER increase to 1.22 and 16.4%, respectively. On the other hand, albeit these impressive performance statistics, these strategies are not practically useful due to their high turnover.

[Table 3.4 about here.]

3.5.1.3 The best expected return specification: Discussion

We investigate whether our three mean-variance portfolio strategies based on the option-implied expected returns, Q-CFER, Q-RP and Q-FULL, statistically outperform other benchmark strategies in terms of the OOS return, Sharpe ratio and certainty equivalent. We follow DPUV and estimate the bootstrapped p-values, where the null hypothesis is that the performance metric of either one of Q-strategy is no better than that of the competing strategy. Therefore, a low p-value suggests that the Q-strategy outperforms competing strategies. We generate 10,000 bootstrapped return time-series to calculate the p-values. We consider EW, VW, short-constrained parametric portfolio strategies, and nine mean-variance portfolios constructed under NS10 constraint and using the historical covariance matrix.

Table 3.5 reports the result. The first three columns show the horse race result regarding the Q-CFER strategy. These results suggest that the Q-CFER strategy has superior performance compared to any other competing strategies. Specifically, the null hypothesis that Q-CFER has no better Sharpe ratio is rejected except only two (three) strategies at a 10% (5%) significance level. The result on the certainty equivalent is even more favorable to Q-CFER; all strategies have a p-value lower than 5%. This confirms that the Q-CFER strategy exhibits superior performance compared to not only the EW, VW and minimum variance portfolios but also the previously proposed option-based strategies in DPUV and Martin and Wagner (2018). This suggests that option-implied information on CFER enables us to achieve a notoriously difficult task to obtain the mean-variance portfolio which has superior empirical performance.

[Table 3.5 about here.]

On the other hand, one may be surprised to see that the option-implied risk premium component is not helpful for the mean-variance portfolio construction; the two strategies which utilize the estimated risk premium component, Q-RP and Q-FULL, do not beat other strategies in general. Why is the CFER component helpful to improve the portfolio performance whereas the risk premium component is not? This difference stems from the fact that the risk premium component is a compensation for risk exposures, whereas the CFER component is not. Equation (3.14) shows that a higher risk premium component means a higher (risk-neutral) variance. This suggests that tilting portfolio weights toward stocks with a high risk premium component mechanically increases the portfolio expected variance. Therefore, it is not a priori clear whether the trade-off between a higher mean and a higher variance is improved. On the other hand, the CFER component is not a compensation for risk exposures. As a consequence, tilting portfolio weights toward stocks with a high CFER component is effective to improve the portfolio expected return without much increasing the portfolio variance.

The alphas of the mean-variance portfolios are in line with the fact that the risk premium is a compensation for risk exposures and the CFER is not. We can see from Panels C and D of Table 3.4 that three strategies which involve CFER (Q-CFER, Q-FULL, and CA-CFER), specifically Q-CFER, earn significant abnormal returns. This result suggests that the outperformance of these strategies cannot be explained by standard risk factors, that is, their outperformance is not attributable to the risk premium component.[12] On the other hand, other portfolio strategies including Q-RP, EW, and VW do not earn significant alphas.

3.5.2 Choice of the covariance matrix

Now we investigate how the choice of the covariance matrix affects the performance of mean-variance portfolio strategies. To this end, we consider five alternative covariance matrix specifications: HV-HC, RIV-HC, HV-ZC, RIV-ZC, and LWI. For the first four covariance specifications, the first part of the model name denotes the specification of the volatility, either historical one (HV), or risk-adjusted option-implied volatility (RIV). The second part of the model name denotes the specification of the correlations, either historical one (HC) or zero correlations (ZC). LWI is the historical covariance matrix shrunk toward the scaled identity matrix based on Ledoit and Wolf (2004). For the specifications of the expected returns, we consider GMVP, Q-CFER, Q-FULL, and CA-CFER. We choose these four specifications because GMVP

[12]Note that CA-RNS also earns significant alphas under the ABS10 constraint. This is again in line with our CFER-based interpretation; Chapter Two document that RNS is correlated with CFER and the predictive power of RNS stems from its correlation with CFER.

is a natural benchmark choice which is widely employed in the literature while Q-CFER, Q-FULL, and CA-CFER are the top three strategies in terms of the Sharpe ratio among the unconstrained mean-variance portfolios. Therefore, we compare 20 mean-variance portfolio strategies in total to see how the specifications of the expected return and the covariance matrix matter for the mean-variance portfolio performance.

We start with the performance of unconstrained portfolio strategies reported in Table 3.6. The implication of the performance of GMVP reported in 3.6 Panel A is largely the same as DPUV; using RIV reduces the portfolio volatility and increases the Sharpe ratio. On the contrary, Panels B to D show that using RIV in conjunction with the option-implied expected returns worsens the portfolio performance due to a higher OOS volatility. This suggests that the benefit from incorporating better forward-looking volatility estimate via using RIV is present only for the minimum-variance portfolio construction. Regarding the choice of the correlation specification, for the mean-variance portfolios (Panels B to D), the ZC strategies always have a lower OOS portfolio volatility and a lower turnover rate than the corresponding HC strategies with the same expected return and volatility specifications. We will discuss the mechanism behind these findings in Section 3.5.3.

[Table 3.6 about here.]

Among 20 strategies shown in Table 3.6, the highest Sharpe ratio is obtained by HV-ZC (historical variance and zero correlation) covariance matrix combined with the Q-CFER expected returns vector, followed by the combination of HV-ZC covariance matrix and CA-CFER return vector. These Sharpe ratios are almost twice as high as those of the EW and VW, implying that option-implied expected returns are useful to improve the portfolio performance before transaction costs. On the other hand, since the turnover rate is extremely high for these strategies, it is not possible to exploit these seemingly high Sharpe ratios in practice as documented in Section 3.5.1.1.

Next, we examine the performance of the same 20 mean-variance portfolios under the NS10 constraint. Table 3.7 reports the result. We can see that the OOS Sharpe ratio is largely the same regardless of the covariance matrix specifications within the same expected return specifications. On the other hand, the choice of the expected return specification is more pertaining to the performance of portfolios. In particular, the strategies using the Q-CFER expected returns generally perform the best, followed by the strategies using the CA-CFER expected returns. This result has an important implication for the mean-variance portfolio strategies, that is, the choice of the expected return specification is of first-order importance while that of the covariance matrix is of second-order importance, once the constraints on the weight are imposed. Contrary to the unconstrained case, using the ZC specification does not necessarily reduce the OOS portfolio volatility nor the turnover rate.

[Table 3.7 about here.]

3.5.3 Mechanism behind the empirical findings: Discussion

3.5.3.1 Shrinkage-like effect of constraints on portfolio weights

Our empirical results so far show that imposing constraints on portfolio weights drastically improves the performance of mean-variance portfolios. We argue that this finding complements a theoretical result in Jagannathan and Ma (2002, 2003) regarding the mean-variance portfolio problem.

A well-known result of the seminal work of Jagannathan and Ma (2003) is that the upper bound and the lower bound constraints on portfolio weights of the form of equation (3.26) can be interpreted to have a *shrinkage-like* effect on the *covariance matrix* in the *minimum* variance portfolio problem. They show that the lower (upper) bound on the portfolio decreases (increases) possibly overestimated (underestimated) covariance matrix elements to mitigate measurement errors in the covariance matrix input. Albeit less famous compared to their result on the minimum variance portfolio, Jagannathan and Ma (2002), which is the working paper version of Jagannathan and Ma (2003), provides a corresponding theoretical result on the *mean*-variance portfolio problem. They show that portfolio weight constraints on the mean-variance portfolio problem can be interpreted to have a shrinkage-like effect on the *expected stock returns*; Jagannathan and Ma (2003, p.1660) show that

> [T]he effect of the lower bound on portfolio weights is to adjust the mean returns upward by an amount proportional to the Lagrange multiplier [of the weight constraints], and the effect of the upper bound on portfolio weights is a similar downward adjustment.

However, unlike their main result on the minimum variance problem, Jagannathan and Ma (2002, 2003) fail to find the empirical usefulness of the shrinkage-like effect in the mean-variance problem; they find that the constrained mean-variance portfolio based on the historical sample moments performs worse than the global minimum variance portfolio, even though the constrained portfolio construction is benefited from the shrinkage-like effect. Consequently, they conclude that *"the estimates of the mean returns are so noisy that simply imposing the portfolio weight constraint is not enough, even though the constraints still have a shrinkage effect"* (Jagannathan and Ma, 2003, p.1654) and they urge us to *"bring additional information about the population mean instead of relying on sample mean as an estimator of population mean"* (Jagannathan and Ma, 2003, p.1660).

Our result shows for the first time that this shrinkage-like effect of the weight constraints on the mean-variance portfolio problem is useful once the option-implied expected return is employed. This result also implies that our option-based forward-looking estimate of the expected returns bring additional information about the population mean as requested by Jagannathan and Ma.

Our empirical results indicate that constraints on portfolio weights are not only useful but also indispensable to obtain practically reasonable portfolio allocations. This finding is closely related to the *error-maximizing* nature of the mean-variance optimization discussed in Michaud (1989), Best and Grauer (1991) and Chopra and Ziemba (1993) among others. These studies document that the unconstrained mean-variance optimization amplifies even tiny estimation errors in the expected stock return input, leading to unstable and poorly performing portfolio allocations. Therefore, even though our option-implied estimate of the expected stock returns are less susceptible to measurement errors compared to the historical sample means, a mechanism for mitigating the adverse effect of measurement errors is necessary. Moreover, Chopra and Ziemba (1993) documents that the mean-variance weight is generally more sensitive to measurement errors in the expected returns than those in the covariance matrix, implying that the estimation quality of the expected return is more important than that of the covariance matrix. In line with their argument, we find that the choice of the expected return specification is more pertinent to the OOS constrained mean-variance portfolio performance than the specification of the covariance matrix.

3.5.3.2 Trade-off between better forecast and measurement errors

The error maximization property of the mean-variance portfolio gives rise to the trade-off between better forecast and measurement errors; it is often the case that using more informative estimates of future returns and covariance do not necessarily improve or even do deteriorate the performance of mean-variance portfolios when these estimates are contaminated by large measurement errors. This is because the gain from using more informative estimate is overwhelmed by measurement errors which are amplified in the mean-variance optimization process. The most well-known example of this trade-off is the poor performance of the mean-variance portfolio based on historical sample means; historical sample means are so noisy that the minimum variance portfolio performs better, even though the minimum variance optimization does not exploit the information on the expected returns at all.

Another example in our results is that the ZC covariance matrix specification results in lower volatility and turnover compared to the HC specification in the unconstrained mean-variance construction. Best and Grauer (1991) show that the sensitivity of the optimal mean-variance portfolio weight to small changes in the expected return (i.e., the degree of the error amplification) is determined by the maximum to minimum eigenvalue ratio of the covariance matrix; when this max-to-min eigenvalue ratio is large, the mean-variance portfolio allocation is more sensitive and susceptible to measurement errors in the expected returns. We find that the max-to-min eigenvalue ratio of the HV-ZC covariance matrix is in the order of $10^{1.5}$, whereas that of the HV-HC matrix is in the order of $10^{4.5}$, suggesting that the sensitivity of the portfolio weight to small changes in the expected returns is about 1,000 times larger for the HC specification than the ZC one. The finding that the HV-ZC covariance matrix has a lower max-to-min eigenvalue ratio is sensible, because it equals the ratio of the maximum and minimum variance of individual stocks. On the other hand, the minimum eigenvalue of the historical covariance matrix is much smaller given empirical findings that the covariance matrix exhibits strong factor structure (e.g., Chan et al., 1999), implying that few eigenvectors can explain most of the covariance structure. In effect, using ZC helps to improve some characteristics of unconstrained portfolios by reducing the measurement error issue, even though zero correlations are less informative than the historical correlations.

The discussion here is also related to our results on the option-implied risk-adjusted variance (RIV). For general mean-variance portfolio construction, we find that RIV is not useful to improve the portfolio performance, even though DPUV document that RIV contains richer, forward-looking information on the future variance compared to the sample variance. Indeed, DPUV acknowledge that using RIV improves the portfolio performance in limited situations such as the GMVP with the ZC covariance matrix. They argue that *"implied volatility is a highly unstable estimator of future volatility; this instability increases the error in portfolio weights and reduces the gains from having a better predictor of RV [realized volatility]"* (p.1833, footnote 20). Their argument suggests that the gain from the better variance prediction by RIV is visible only when the degree of measurement error amplification is low (e.g., the GMVP with the ZC covariance matrix). Therefore their discussion is in line with our finding that using RIV does not improve the mean-variance portfolio construction problem, where measurement errors have relatively larger impact on the resulted optimal weight compared to the minimum variance portfolio construction.

Finally, the discussion here may be another reason (in addition to our discussion in Section 3.5.1.3) why using the option-implied risk premium component does not improve the portfolio performance. While Martin and Wagner (2018) document that the option-implied risk premium provides rich information on the risk premium of stocks, its formula contains the unobservable residual term (see equation (3.13)). Even though Martin and Wagner (2018) document that this residual term does not cause problems for their empirical analysis, it may have large impact on the mean-variance portfolio construction given the amplification mechanism in the mean-variance optimization.

3.5.4 Mean-variance portfolios: Alternative universe

Next, we report the mean-variance portfolio performance under the alternative 500 randomly selected optionable stocks universe. This universe contains smaller stocks compared to the S&P 500 universe; about three quarters of the stocks do not belong to the S&P 500 index.

Table 3.8 reports the result. Panel A shows the performance of the benchmark EW and VW portfolios as well as those of the rank-based strategies and short-sale constrained parametric portfolio strategies. The performance of these strategies are largely unchanged compared to those obtained from the S&P 500 universe case except that some short-sale constrained parametric portfolio strategies improve the Sharpe ratio. Panel B shows the performance of the mean-variance portfolios using the historical covariance matrix and various expected return specifications with the NS10 constraint. We can see that two quantitative option-implied expected return specifications, Q-CFER and Q-FULL, attain even better performance compared to the S&P 500 universe. This result is plausible since CFER and (shifted) risk premium exhibit larger variations in smaller stocks (see Section 3.3.3) and hence there are more opportunities to improve the portfolio performance by utilizing large dispersion in the expected stock returns.

[Table 3.8 about here.]

3.5.5 Transaction costs and portfolio performance

Our results above suggest that using the Q-CFER expected returns vector drastically improves the mean-variance portfolio performance. However, this strategy (and other strategies using the option-implied expected returns measure in general) exhibits a high turnover rate, reflecting large variations in the option-implied expected stock returns across the cross-section of stocks and over the time-series. Therefore, it is crucial to take transaction costs into account when it comes to evaluating the practical usefulness of the option-based mean-variance portfolio strategies.

Before examining the transaction costs issues, we briefly overview typical transaction cost values considered in the stock portfolio selection literature. Based on Hasbrouck's (2009) statistical method, Novy-Marx and Velikov (2016) calculate the effective transaction costs for each stock in yearly basis. According to Figure 1 of their article, which displays the average transaction costs in each decade as a function of the firm size rank, the effective *round-trip* transaction costs for the top 500 largest stocks are around 40 to 60 bps (hence the one-way transaction costs are about 20 to 30 bps). Brandt et al. (2009) employ a simplified transaction costs specification which incorporates the decline in transaction costs in recent years and the fact that larger stocks have lower transaction costs. Specifically, they assume that transaction costs after 2003 are 35 bps for the largest stock and 60 bps for the smallest stock. This setting is adopted in the following studies such as DeMiguel et al. (2019). Given these conventional values in the literature, we take 30 bps for large S&P 500 constituent stocks and 60 bps for smaller stocks as indicative transaction costs values. These values are indicative because transaction costs depend on various factors such as the size of trade and the degree of the sophistication of traders.[13]

We begin with reporting the equivalent transaction costs between pairs of two strategies. The equivalent transaction costs of the strategy i with respect to the competing strategy j, $TC_{i,j}$, is the constant transaction costs that would equate the performance measure of the two strategies. We calculate the equivalent transaction costs regarding the OOS mean excess returns and the Sharpe ratio, $TC_{i,j}^{mean}$ and $TC_{i,j}^{SR}$, defined as follows:

$$\mu_i - TC_{i,j}^{mean} TO_i = \mu_j - TC_{i,j}^{mean} TO_i, \quad \frac{\mu_i - TC_{i,j}^{SR} TO_i}{\sigma_i} = \frac{\mu_j - TC_{i,j}^{SR} TO_j}{\sigma_j},$$

where μ_i, σ_i, and TO_i are the OOS mean excess return, volatility and turnover rate of the strategy i.

Table 3.9 reports the equivalent transaction costs of the Q-CFER strategy with the historical covariance matrix and the NS10 constraint against various competing strategies. The symbol "D" indicates that the Q-CFER strategy has a lower performance measure than the strategy in comparison *before* transaction costs are taken into account. A positive equivalent transaction cost means that the Q-CFER strategy's after-transaction cost performance measure is better than that of the competing strategy as long as the transaction cost is lower than the displayed value.

In the S&P 500 universe, the toughest rivals are EW and GMVP; the Q-CFER strategy beats the EW (GMVP) strategy in terms of the Sharpe ratio as long as the transaction costs are smaller than 35 bps (32 bps). The Q-CFER strategy may beat even these two strategies given the indicative transaction costs of 30 bps. Q-CFER comfortably outperforms other competing strategies given a moderate size of transaction costs for trading large stocks, despite of its high turnover.

[Table 3.9 about here.]

Over the randomly selected 500 stocks universe, the equivalent transaction costs of the Q-CFER strategy becomes larger in general. The equivalent transaction costs

[13]For example, based on proprietary data, Frazzini et al. (2014) document that actual stock trading costs for stock funds are much smaller than previous studies suggest.

about the Sharpe ratio regarding EW and GMVP are approximately doubled to about 70 bps, making the Q-CFER strategies even easier to beat the competing strategies. Note that the transaction costs of the Q-CFER strategy against the Q-FULL strategy is about 20 bps. This does not change the implication of our findings that the option-implied expected returns improves the mean-variance portfolio performance, because Q-FULL is an alternative specification of the option-based expected return.

When the Q-CFER strategy is compared with other option-based strategies examined in DPUV (i.e., short-sale constrained parametric portfolios and characteristic-adjusted mean strategies), the equivalent transaction costs in the S&P 500 universe range between about 50 to 100 bps and those in the randomly selected 500 stocks universe range between about 90 to 180 bps. This corroborates our finding that the mean-variance portfolio based on our option-implied estimate of the expected return utilizes richer informational contents of option prices compared to the parametric portfolio approach and the characteristic-adjusted return approach.

Novy-Marx and Velikov (2016) document that simple transaction costs reduction techniques are helpful for performance enhancement, especially for high turnover long-short anomaly portfolio strategies. Among the techniques they examined, we apply the *staggered partial rebalancing* technique to our mean-variance portfolio construction. Under this technique, only a part of the portfolio is rebalanced (e.g., only one third of the portfolio is rebalanced in each month under the quarterly staggered rebalancing scheme) and hence the overall turnover becomes lower.[14]

Table 3.10 reports the performance of staggered partial rebalancing Q-CFER portfolio with the historical covariance matrix and NS10 constraint. We consider three alternative frequencies of rebalancing: quarterly (3m), biannual (6m), and annual (12m) rebalancing. For the ease of comparison, we also include the performance metrics of selected non-staggered strategies. We can see that the turnover rate of staggered portfolios decreases roughly proportionally to the rebalance frequency. For example, the turnover rate of the quarterly rebalancing portfolios are about one third of the corresponding non-staggered strategy. The Sharpe ratio of the staggered portfolios

[14] Novy-Marx and Velikov (2016) examine two alternative transaction costs reduction strategies, limiting the investment universe to low transaction costs stocks and the so-called buy/hold strategy. We do not consider the first alternative technique because our universes are already confined to large stocks. We do not consider the second alternative technique neither because it is not directly applicable to the mean-variance portfolio construction problem.

also decreases, yet it is still higher than the benchmark strategies including EW, VW, and GMVP. As a result, the equivalent transaction costs of the Q-CFER strategy about the Sharpe ratio increase compared to the non-staggered case. For instance, with the biannual rebalancing, the Q-CFER strategy earns a higher after transaction costs Sharpe ratio than the three benchmark strategies as long as transaction costs are below about 100 bps. This result suggests that simple transaction costs reduction techniques further enhance the performance of the mean-variance portfolio strategies, in line with Novy-Marx and Velikov (2016).

[Table 3.10 about here.]

3.5.6 Why does Q-CFER strategy improve the portfolio performance?

Conceptually, CFER measures the degree of limits-of-arbitrage each stock faces. This implies that CFER-based portfolios load on stocks which are more difficult to trade. Therefore, one may expect that it is difficult to profit from CFER-based strategies. On the contrary, our empirical results so far suggest that the Q-CFER mean-variance portfolio strategy can beat other portfolio strategies even after considering transaction costs. We provide two caveats on these arguably puzzling empirical results.

First, we consider only transaction costs as market frictions for the evaluation of the performance of portfolio strategies, although there may be other types of frictions that are relevant to portfolio investing. While it is standard in the portfolio literature to consider only transaction costs for frictions to execute portfolio strategies, our approach may underestimate the difficulty in trading portfolio strategies (in our case, Q-CFER strategy). Second, our empirical evaluation is critically affected by the choice of the size of transaction costs. For investors who face transaction costs greater than the "indicative" transaction costs we used, it is difficult for them to profit from employing the Q-CFER strategy. Our discussion at the beginning of Section 3.5.5 suggests that the "indicative" transaction costs we used are standard in the literature. However, given the fact that transaction costs vary across investors and also given that they are difficult to estimate with precision, we do not conclude that the Q-CFER strategy is always profitable for any investors.

On the other hand, there is a possibility that the Q-CFER strategy improves the portfolio performance via the risk premium component. If CFER is positively correlated with the risk premium component, titling portfolios toward stocks with higher CFER can earn profit through a higher risk premium component. This mechanism might be present in the randomly selected 500 stock universe. The variance decomposition result in Table 3.1, Panel C, shows that CFER and the shifted risk premium component is slightly positively correlated over the all optionable stock universe. This may contribute to the good performance of Q-CFER strategy in the randomly selected 500 stock universe. On the other hand, this mechanism seems to be absent in the S&P 500 universe as Table 3.1, Panel C, shows almost zero correlation between the estimated CFER and the shifted risk premium.

3.6 Conclusions

In this Chapter, we document that the option-implied measures of expected stock returns are useful to construct the Markowitz (1952) mean-variance portfolios.

To construct estimates of the expected stock returns under a realistic market model with frictions, we generalize Martin and Wagner's (2018) expected return formula to allow the presence of market frictions. We show that the expected stock return can be estimated as the sum of the contribution of frictions to expected return (CFER) analyzed in Hiraki and Skiadopoulos (2019) and the scaled risk-neutral simple variance, apart from the cross-sectionally constant term which does not affect the mean-variance portfolio optimization outcome. Since CFER and the risk-neutral simple variance can be estimated from option price data based on the result in Hiraki and Skiadopoulos (2019) and Martin (2017), respectively, this allows us to construct a real-time, forward-looking measure of expected stock returns vector solely from option price information.

Our empirical results yields three implications worth summarizing here. First, we document that the Q-CFER mean-variance strategy, for which we use the estimated CFER as the estimate of the expected stock returns, outperforms other portfolio strategies which have been documented to have good performance including the equally-weighted portfolio, minimum variance portfolios, and previously proposed option-based portfolio strategies, both before and after considering a moderate size of transaction costs. This may change the common perception in the literature that it is notoriously difficult to construct a mean-variance portfolio with good empirical performance.

Second, the outperformance of the Q-CFER strategy suggests that incorporating the effect of market frictions into the estimation of the expected return is of importance. Tilting portfolio weights toward stocks with a high CFER component is effective to improve the portfolio expected return *without* much increasing the portfolio variance, because CFER is not a compensation for risk exposures. On the other hand, we find that utilizing the option-implied estimate of the risk premium component does not improve the portfolio performance in general.

Finally, we find that imposing constraints on portfolio weights dramatically improves the performance of mean-variance portfolios. This result complements the empirically unexploited theoretical result in Jagannathan and Ma (2002, 2003) on the *mean*-variance portfolio construction; they theoretically show that portfolio weight constraints on the mean-variance portfolio problem can be interpreted to have a *shrinkage-like* effect to mitigate measurement errors in the expected returns. Nevertheless, they document that this shrinkage-like effect is not useful empirically when the noisy historical sample mean is employed. Our result shows for the first time that this shrinkage-like effect of the weight constraints on the mean-variance portfolio problem is useful once the option-implied expected return is employed.

3.A Proofs

3.A.1 Proof of Proposition 3.2.2

By taking the value-weighted average of equation (3.12), we obtain

$$\mathbb{E}_t^{\mathbb{P}}[R_{M,t,t+1}] - R_{f,t,t+1} = \sum_{i=1}^{N} \omega_{i,t}^{M} \left(CFER_{i,t,t+1} + V_{i,t,t+1}\right) + \frac{Var_t^{\mathbb{Q}^*}(\xi_{t,t+1}) - \overline{u}_t + 2\overline{v}_{i,t}}{2R_{f,t,t+1}}.$$

Since the last term in the right-hand side of this equation equals c_t, it follows that

$$c_t = (\mathbb{E}_t^{\mathbb{P}}[R_{M,t,t+1}] - R_{f,t,t+1}) - \sum_{i=1}^{N} \omega_{i,t}^{M} CFER_{i,t,t+1} - \sum_{i=1}^{N} \omega_{i,t}^{M} V_{i,t,t+1}.$$

Therefore, it suffices to show that the expected excess market return satisfies

$$\mathbb{E}_t^{\mathbb{P}}[R_{M,t,t+1}] - R_{f,t,t+1} = \frac{Var_t^{\mathbb{Q}^*}(R_{M,t,t+1})}{R_{f,t,t+1} + CFER_{M,t,t+1}} + CFER_{M,t,t+1}. \tag{3.A.1}$$

To show equation (3.A.1), first we apply the same trick in Martin (2017) to obtain

$$\begin{aligned} \frac{\mathbb{E}_t^{\mathbb{Q}^*}[R_{M,t,t+1}^2]}{R_{f,t,t+1}} &= \mathbb{E}_t^{\mathbb{P}}[m_{t,t+1}^* R_{M,t,t+1}^2] \\ &= \mathbb{E}_t^{\mathbb{P}}[m_{t,t+1}^* R_{M,t,t+1}]\mathbb{E}_t^{\mathbb{P}}[R_{M,t,t+1}] + Cov_t^{\mathbb{P}}(m_{t,t+1}^* R_{M,t,t+1}, R_{M,t,t+1}). \end{aligned}$$

The last covariance term vanishes under the assumption that the inverse of the IMRS coincides with the market portfolio return. In this case, we obtain

$$\mathbb{E}_t^{\mathbb{P}}[R_{M,t,t+1}] - R_{f,t,t+1} = \frac{1}{R_{f,t,t+1}} \frac{\mathbb{E}_t^{\mathbb{Q}^*}[R_{M,t,t+1}^2]}{\mathbb{E}_t^{\mathbb{P}}[m_{t,t+1}^* R_{M,t,t+1}]} - R_{f,t,t+1} \tag{3.A.2}$$

The numerator of the first term in the right-hand side of equation (3.A.2) can be calculated as

$$\begin{aligned} &\mathbb{E}_t^{\mathbb{Q}^*}[R_{M,t,t+1}^2] = Var_t^{\mathbb{Q}^*}(R_{M,t,t+1}) + \mathbb{E}_t^{\mathbb{Q}^*}[R_{M,t,t+1}]^2 \\ &= Var_t^{\mathbb{Q}^*}(R_{M,t,t+1}) + (R_{f,t,t+1}^2 + 2R_{f,t,t+1}CFER_{M,t,t+1} + CFER_{M,t,t+1}^2) \end{aligned} \tag{3.A.3}$$

due to equation (3.5). On the other hand, the denominator of the same term can be calculated as

$$\frac{1}{R_{f,t,t+1}\mathbb{E}_t^{\mathbb{P}}[m_{t,t+1}^* R_{M,t,t+1}]} = \frac{1}{\mathbb{E}_t^{\mathbb{Q}^*}[R_{M,t,t+1}]} = \frac{1}{R_{f,t,t+1} + CFER_{M,t,t+1}} \tag{3.A.4}$$

again due to equation (3.5). Substituting equation (3.A.3) and (3.A.4) into (3.A.2) and performing algebra prove equation (3.A.1).

Note that when the effect of market frictions on the market index is absent (i.e., $CFER_{M,t,t+1} = 0$), equation (3.A.1) boils down to $\mathbb{E}_t^{\mathbb{P}}[R_{M,t,t+1}] - R_{f,t,t+1} = Var_t^{\mathbb{Q}^*}(R_{M,t,t+1})/R_{f,t,t+1}$, which coincides with Martin's (2017) frictionless result with the assumption that the inverse of the IMRS equals the market return. Therefore, equation (3.A.1) can be viewed as the generalization of his result to the case where the market frictions may affect the expected market return. □

3.A.2 Proof of Proposition 3.2.3

For the ease of notation, we suppress the stock index i and the double time subscript $t, t+1$. The stock return satisfies $R = (S_{t+1} + \widetilde{D})/S_t = (S_{t+1} - F + R_f S_t)/S_t$. Then, the following equation holds:

$$\begin{aligned} Var_t^{\mathbb{Q}^*}(R) &= \mathbb{E}_t^{\mathbb{Q}^*}[R^2] - (\mathbb{E}_t^{\mathbb{Q}^*}[R])^2 = \frac{\mathbb{E}_t^{\mathbb{Q}^*}[(S_{t+1} - F + R_f S_t)^2]}{S_t^2} - (\mathbb{E}_t^{\mathbb{Q}^*}[R])^2 \\ &= \underbrace{\frac{\mathbb{E}_t^{\mathbb{Q}^*}[S_{t+1}^2]}{S_t^2} - \frac{2F\mathbb{E}_t^{\mathbb{Q}^*}[S_{t+1}]}{S_t^2} + \frac{F^2}{S_t^2}}_{Y_1} + \underbrace{\frac{2R_f}{S_t}(\mathbb{E}_t^{\mathbb{Q}^*}[S_{t+1}] - F) + R_f^2 - (\mathbb{E}_t^{\mathbb{Q}^*}[R])^2}_{Y_2}. \end{aligned} \tag{3.A.5}$$

We separately calculate Y_1 and Y_2 in equation (3.A.5). First, to further transform Y_1, note that the following equation holds:

$$S_{t+1}^2 = 2\int_0^{S_{t+1}} (S_{t+1} - K)dK = 2\int_0^{\infty} (S_{t+1} - K)^+ dK, \tag{3.A.6}$$

where $(x)^+ = \max(x, 0)$. Therefore, taking $\mathbb{Q}^*$-expectation yields

$$\mathbb{E}_t^{\mathbb{Q}^*}[S_{t+1}^2] = 2R_f \int_0^{\infty} C_t(K)dK \quad \Leftrightarrow \quad \frac{\mathbb{E}_t^{\mathbb{Q}^*}[S_{t+1}^2]}{S_t^2} = \frac{2R_f}{S_t^2}\int_0^{\infty} C_t(K)dK. \tag{3.A.7}$$

Due to the risk-neutral valuation formula, it follows that $C_t(K) - P_t(K) = (\mathbb{E}_t^{\mathbb{Q}^*}[S_{t+1}] - K)/R_f$. This further yields

$$\begin{aligned} \int_0^F C_t(K)dK &= \int_0^F P_t(K)dK + \frac{\mathbb{E}_t^{\mathbb{Q}^*}[S_{t+1}]}{R_f}\int_0^F dK - \frac{1}{R_f}\int_0^F KdK \\ &= \int_0^F P_t(K)dK + \frac{\mathbb{E}_t^{\mathbb{Q}^*}[S_{t+1}]}{R_f}F - \frac{F^2}{2R_f}. \end{aligned} \tag{3.A.8}$$

Therefore, combining equations (3.A.7) and (3.A.8) yield

$$\frac{\mathbb{E}_t^{\mathbb{Q}^*}[S_{t+1}^2]}{S_t^2} = \frac{2R_f}{S_t^2}\left(\int_F^{\infty} C_t(K)dK + \int_0^F P_t(K)dK\right) + \frac{2F\mathbb{E}_t^{\mathbb{Q}^*}[S_{t+1}]}{S_t^2} - \frac{F^2}{S_t^2}.$$

This shows that

$$Y_1 = \frac{2R_f}{S_t^2}\left(\int_F^{\infty} C_t(K)dK + \int_0^F P_t(K)dK\right). \tag{3.A.9}$$

Next, we transform Y_2. To this end, note that $\mathbb{E}_t^{\mathbb{Q}^*}[S_{t+1} - F]/S_t = \mathbb{E}_t^{\mathbb{Q}}[R] - R_f$

holds. Therefore,

$$Y_2 = 2R_f(\mathbb{E}_t^{\mathbb{Q}^*}[R] - R_f) + R_f^2 - (\mathbb{E}_t^{\mathbb{Q}^*}[R])^2 = -(\mathbb{E}_t^{\mathbb{Q}^*}[R] - R_f)^2. \tag{3.A.10}$$

Equation (3.A.10) equals $-CFER^2$ because of equation (3.5). Finally, dividing equation (3.A.5) by R_f and using equations (3.A.9) and (3.A.10) prove equation (3.19). □

3.A.3 Proof of Lemma 3.2.1

Under the constraint $\omega' e = 1$, the optimization problem given Σ and $\widetilde{\mu} = \mu + ke$ is

$$\min_{\omega} \frac{\gamma}{2}\omega'\Sigma\omega - \omega'\mu - k \quad \Leftrightarrow \quad \min_{\omega} \frac{\gamma}{2}\omega'\Sigma\omega - \omega'\mu, \tag{3.A.11}$$

subject to the constraint $\omega' e = 1$. Therefore, the optimal solution ω^* is the same for the pair (μ, Σ) and the pair $(\widetilde{\mu}, \Sigma)$. □

3.B Description of dataset and variables

In this Appendix, we provide a detailed explanation for the variable construction procedures and the estimation procedures of option-implied measures.

3.B.1 Construction of the firms' and stocks' characteristics

We construct the firms' size characteristics using the CRSP database. First, we calculate the stock market capitalization as the product of the close price (CRSP item prc) and the shares outstanding (shrout). Then, we take the natural logarithm of the market capitalization. For the momentum (MOM), we follow DPUV and calculate the momentum on day t as the cumulative daily return from day $t - 251 - 21$ to day $t - 21$.

We construct the log book-to-market value logBM by following Davis et al. (2000). Specifically, we measure book equity as stockholders' book equity, plus balance sheet deferred taxes and investment tax credit (Compustat annual item TXDITC) if available, minus the book value of preferred stock. Stockholders' equity is the value reported by Compustat (SEQ), if it is available. If not, we measure stockholders' equity as the book value of common equity (CEQ) plus the par value of preferred stock (PSTK), or the book value of assets (AT) minus total liabilities (LT). Depending on availability, we use redemption (PSTKRV), liquidating (PSTKL), or par value (PSTK) for the book value of preferred stock. From June of each year t to May of $t+1$, the book-to-market equity is calculated as the ratio of the book equity for the fiscal year ending in calendar year $t-1$ to the market equity at the end of December of year $t-1$. We treat non-positive book-to-market equity data as missing. Then, we take the natural logarithm to calculate logBM.

3.B.2 Linking the OM, CRSP, and Compustat databases

In our analysis, we match three databases, the OM, CRSP, and Compustat. To link the CRSP and Compustat, we use the SAS linking macro provided by the WRDS and match PERMNO (the CRSP identifier) and GVKEY (the Compustat identifier). To link the CRSP and OM, we link SECID (the OM identifier) and PERMNO by the SAS linking macro provided by the WRDS. This WRDS macro links SECID and PERMNO based on the CUSIP code, the similarity in the issuer's (company) name recorded in the OM and CRSP, and partly the security ticker. The macro also provides the score of the strength of the link based on a 0–6 integer scale (0 stands for the strongest link, where the eight-digit CUSIP exactly matches as well as the issuer's name matches). We keep SECID-PERMNO link with the score less than or equal to three. After linking SECID, GVKEY and PERMNO, we keep only U.S. common stocks (SHRCD 10 or 11) traded in NYSE/Amex/NASDAQ (EXCHCD 1, 2, or 3).

3.B.3 Estimation of CFER

HS show that the scaled deviations from put-call parity, equation (3.21), reliably estimates the underlying stocks' CFER. In the actual estimation procedure, we use an alternative equivalent expression for the calculation of the deviations from put-call parity $\widetilde{S}_t(K) - S_t$ in equation (3.21):

$$\begin{aligned}\widetilde{S}_t(K) - S_t = BS_{call}(S_t, K, T-t, r_f, \widetilde{D}_{t,T}, IV^c_{t,T}(K)) \\ - BS_{call}(S_t, K, T-t, r_f, \widetilde{D}_{t,T}, IV^p_{t,T}(K)),\end{aligned} \tag{3.B.1}$$

where $BS_{call}(S, K, \tau, r, D, \sigma)$ is the Black-Scholes European call option function with the deterministic dividend payment D, r_f is the continuously compounded risk-free rate, and $IV^c_{t,T}(K)$ $(IV^p_{t,T}(K))$ is the IV of the call (put) option. See the proof of Proposition 2.1 of HS for the proof of equation (3.B.1).

To calculate equation (3.B.1), we use the data from the *OM Standardized Option Price file*, which provides the call and put option prices as well as their respective IVs at the forward at-the-money (ATM), $K = F_{t,T} = R_{f,t,T}S_t - \widetilde{D}_{t,T}$, for standardized maturities such as 30, 60, 91, 182, and 365 calender days. The stock price is obtained from the OM underlying price file. Regarding the risk-free rate, we obtain the data from the OM zero yield file and linearly interpolate it to obtain the risk-free rate for the option maturity. We back out the projected dividend payments $\widetilde{D}_{t,T}$ from the forward prices in the file as $\widetilde{D}_{t,T} = F_{t,T} - R_{f,t,T}S_t$ and we use 30-day-to-maturity data.

3.B.4 Estimation of the risk-neutral simple variance and risk-neutral moments

We estimate the risk-neutral simple variance (equation (3.19)) and the risk-neutral model-free implied volatility (MFIV) and the risk-neutral skewness (RNS) (equations (3.22) and (3.23)) relying on the *OM Volatility Surface File*. This file contains IVs and associated option prices and strike prices at standardized delta equals 0.2, 0.25, ..., 0.8 for call options and -0.8, -0.75, ..., -0.2 for put options for the standardized maturities including 30, 60, 91, 182, and 365 calendar days.[15] Therefore, this file allows us to use call and put option data for 13 different strikes for standardized time-to-maturities. We discard few day-stock IV surface data if either the strike prices provided by the OM of calls or puts are not monotonic in deltas. Such a situation may occur when the interpolated volatility surface exhibits an extremely steep skew. We use 30-day-to-maturity data for the purpose of estimating the risk-neutral simple variance and risk-neutral moments.

Before calculating the integrals, we interpolate and extrapolate the IV curves to obtain option prices over a fine grid. To this end, we follow recent studies (e.g., Stilger

[15] The OM estimate the smooth volatility surface based on a kernel density method for each trading day and for each stock if there are sufficient option data. See the OM documentation for the detail.

et al., 2017; Borochin and Zhao, 2018) and interpolate and extrapolate IVs as follows. We interpolate call IVs and put IVs separately in the strike-IV metric based on the cubic piecewise Hermite polynomial interpolation. Then, we horizontally extrapolate the IV beyond the highest and the lowest strikes. Once the continuous IV curve is obtained, we calculate option prices at 1,001 equally-spaced strike points over the moneyness from 1/3 to 3. Finally, we numerically calculate the integrals in equations (3.19) and (3.24) with these splined option prices. The boundary value between the put and call integrals in equation (3.19) is $F_{i,t,t+1}$ and that in equation (3.24) is $S_t - \widetilde{D}_{t,T}$. We use the forward price recorded in the OM Standardized Option Price file for equation (3.19). For equation (3.24), we back out $\widetilde{D}_{t,T}$ from the same forward price as $\widetilde{D}_{t,T} = R_{f,t,T} S_{i,t} - F_{i,t,T}$. See Appendix 3.B.4 for more detail.

Using the OM volatility surface file has several advantageous points. First, it provides the IVs at a standardized maturity. This helps us to avoid any issues arising from mixing estimated RNMs from different maturities. Second, this dataset provides IVs as a function of delta in the unified range (0.2 to 0.8 in the absolute value) for all stocks. This makes it more reliable to compare the estimated RNMs among different stocks and time periods. Third, this dataset has been used in the recent studies for the estimation of the risk-neutral simple variance (Martin and Wagner, 2018) and the estimation of risk-neutral moments (e.g., DeMiguel et al., 2013; Borochin and Zhao, 2018), thus making possible the comparison of our results to theirs.

Table 3.1. Summary statistics of the estimated CFER and the shifted risk premium term

Panels A and B report the summary statistics of the estimated CFER and the estimated shifted risk-premium term $V_{i,t,t+1}$ (equation (3.14)), respectively. At the end of each month, we estimate these two option-implied measures using options with 30 day-to-maturity. The sample period ranges from January 1996 to December 2017. The unit of CFER and the shifted risk-premium is percent per year. Px denotes the x-th percentile point value and IQR stands for the inter-quartile range (the difference between 75th and 25th percentile point values). Panel C reports the variance decomposition of the cross-section of the shifted expected returns, $CFER_{i,t,t+1} + V_{i,t,t+1}$, into the variance of CFER, that of the shifted risk premium term, and their cross variation term. At the end of each month, we conduct the variance decomposition and then we calculate the time-series average of each component. The unit is the percentage (par year) squared. The values in the square brackets denotes the proportion of the variation attributable to respective components. The results in the rows labeled "All" are calculated based on all optionable stocks observations, while those in the rows labeled "S&P 500" are calculated based on the observations of the constituent stocks of the S&P 500 index.

Panel A: Summary statistics of the estimated CFER										
	Mean	Std	Skew	P1	P10	Median	P90	P99	IQR	Obs
All	-0.82	15.90	-1.05	-46.15	-9.72	-0.37	7.63	40.02	5.58	575,580
S&P 500	-0.44	5.57	-2.62	-13.43	-3.74	-0.27	2.65	11.28	2.64	124,777
Panel B: Summary statistics of the shifted risk premium term										
All	16.93	21.96	5.06	1.32	3.06	10.07	37.20	104.93	14.82	575,580
S&P 500	7.18	9.48	7.08	0.99	1.83	4.54	14.20	45.75	5.19	124,777
Panel C: Variance decomposition of the variations in the expected returns										
	Total		CFER		Shifted RP		Cross Var			
All	684.5	[100.0]	235.5	[34.4]	422.1	[61.7]	26.9	[3.9]		
S&P 500	94.9	[100.0]	29.9	[31.5]	65.4	[68.9]	-0.4	[-0.4]		

Table 3.2. List of portfolio strategies

Entries in Panel A show the descriptions and the symbols of portfolio strategies which are not based on the mean-variance portfolio optimization, equation (3.20). Entries in Panel B show the three elements of the mean-variance portfolio strategies, the expected return specification, the covariance matrix specification, and the constraints on the stock weights. A mean-variance portfolio strategy is specified as a combination of these three elements.

Panel A: Non mean-variance portfolio strategies			
Naive		No short-sale Parametric portfolio	
EW	equally-weighted port	PPNS-FFM	Size, logBM, MOM
VW	value-weighted port	PPNS-MFIV	FFM and MFIV
		PPNS-RNS	FFM and RNS
rank-based weights		PPNS-VRP	FFM and VRP
RK-CFER	based on CFER	PPNS-CFER	FFM and CFER
RK-RP	shifted risk premium	PPNS-FULL	FFM and all option var.
RK-FULL	CFER plus RP		

Panel B: Mean-variance portfolio strategies	
Expected returns	
GMVP	global minimum variance portfolio ($\mu_{i,t} = 0$)
Hist	historical sample mean
Q-CFER	quantitative: CFER
Q-RP	Shifted risk premium $V_{i,t}$
Q-FULL	CFER + shifted risk premium
CA-MFIV	characteristic adjusted return: MFIV
CA-RNS	RNS
CA-VRP	VRP
CA-CFER	CFER
Covariance matrix	
HV-HC	Historical covariance matrix
RIV-HC	Risk adjusted implied variance & historical correlations
HV-ZC	Historical variance & zero correlations
RIV-ZC	Risk adjusted implied variance & zero correlations
LWI	Historical cov. matrix with Lediot-Wolf shrinkage toward identity matrix
Constraints	
NC	No constraints
NS	No short-sale
NS10	No short-sale and 10% bound
ABS10	10% Absolute weight bound

Table 3.3. Mean-variance portfolio performance: Unconstrained strategies

Entries in Panel A report the performance measures of the benchmark portfolio strategies which are not based on the mean-variance optimization. EW and VW stand for the equally- and value-weighted portfolio strategy, respectively. RK-X stands for the rank-based strategy in which the portfolio weight is determined by the cross-sectional ranking of the variable X, where X is either CFER, the shifted risk premium (RP), or the sum of these two (FULL). Panel B reports the performance measures of the unconstrained mean-variance portfolios with nine alternative expected returns vector specifications. We consider the minimum-variance portfolio (GMVP), historical mean returns (Hist), the scaled risk premium (Q-RP), CFER (Q-CFER), the sum of the scaled risk premium and CFER (Q-FULL) and four characteristic-adjusted (CA-) returns based on MFIV, VRP, RNS and CFER. We use the historical covariance matrix. The universe of stocks is the S&P 500 member stocks and the out-of-sample (OOS) portfolio return period is from January 1997 to December 2017. The unit of the OOS excess return (ExRet), the volatility (Vol), the Fama and French (1993) three-factor alpha (α_{FF3}) and the Fama and French (2018) five-factor alpha (α_{FF5}) is percent per year. The t-statistics of alphas are adjusted for heteroskedasticity and serial correlations and reported in the parentheses. The certainty equivalent (CE) is calculated based on the power utility with the relative risk-aversion parameter of five. SR and TO stand for the Sharpe ratio and turnover rate, respectively.

Model	ExRet	Vol	SR	CE	TO	α_{FF3}		α_{FF5}	
Panel A: Benchmark strategies									
EW	9.7	16.9	0.58	4.4	0.08	1.10	(0.96)	-0.17	(-0.16)
VW	7.2	14.8	0.48	3.7	0.02	0.25	(1.09)	-0.23	(-0.97)
RK-CFER	9.5	17.4	0.54	3.7	0.95	0.63	(0.52)	-0.50	(-0.45)
RK-RP	9.7	16.8	0.57	4.4	1.01	1.16	(0.92)	-0.12	(-0.10)
RK-FULL	9.6	17.5	0.55	3.7	0.98	0.81	(0.66)	-0.38	(-0.31)
Panel B: Unconstrained mean-variance strategies									
GMVP	4.4	16.3	0.27	-0.8	5.92	2.63	(0.71)	0.62	(0.17)
Hist	-124.9	980.8	-0.13	-1168.2	1546.03	8.50	(0.04)	62.97	(0.28)
Q-RP	120.7	604.4	0.20	-1136.1	640.23	46.63	(0.25)	108.95	(0.50)
Q-CFER	720.6	861.0	0.84	-1171.2	922.44	762.67	(2.84)	832.93	(2.87)
Q-FULL	850.0	1151.7	0.74	-1064.7	891.21	815.69	(2.30)	948.95	(2.27)
CA-MFIV	10.8	60.8	0.18	-150.0	67.74	8.01	(0.61)	10.57	(0.75)
CA-VRP	-1.7	67.0	-0.03	-148.9	68.60	-0.68	(-0.04)	-2.97	(-0.16)
CA-RNS	51.1	91.0	0.56	-166.7	107.12	46.22	(2.42)	47.79	(2.36)
CA-CFER	58.6	79.3	0.74	-233.5	97.73	55.95	(3.42)	55.82	(3.12)

Table 3.4. Mean-variances portfolio performance: Constrained strategies

Entries in Panel A reports the performance measures of the benchmark portfolios strategies which are not based on the mean-variance optimization. EW and VW stand for the equally- and value-weighted portfolio strategy, respectively. PPNS-X stands for the parametric portfolio with no short-sale constraints. See Table 3.2 for the set of characteristic variables for each PPNS strategy. Panels B to D report the performance measures of the mean-variance portfolios with nine alternative expected return vector specifications under three types of constraints on the portfolio weight: no short-sale constraints for Panel B, no short-sale constraints and the 10% limit for Panel C and the 10% limit on the absolute weight for Panel D. We consider the minimum-variance portfolio (GMVP), historical mean returns (Hist), the scaled risk premium (Q-RP), CFER (Q-CFER), the sum of the scaled risk premium and CFER (Q-FULL) and four characteristic-adjusted (CA-) returns based on MFIV, VRP, RNS and CFER. We use the historical covariance matrix. The universe of stocks is the S&P 500 member stocks and the out-of-sample (OOS) portfolio return period is from January 1997 to December 2017. The unit of the OOS excess return (ExRet), the volatility (Vol), the Fama and French (1993) three-factor alpha (α_{FF3}) and the Fama and French (2018) five-factor alpha (α_{FF5}) is percent per year. The *t*-statistics of alphas are adjusted for heteroskedasticity and serial correlations and reported in the parentheses. The certainty equivalent (CE) is calculated based on the power utility with the relative risk-aversion parameter of five. SR and TO stand for the Sharpe ratio and turnover rate, respectively.

Model	ExRet	Vol	SR	CE	TO	α_{FF3}		α_{FF5}	
Panel A: Benchmark strategies									
EW	9.7	16.9	0.58	4.4	0.08	1.10	(0.96)	-0.17	(-0.16)
VW	7.2	14.8	0.48	3.7	0.02	0.25	(1.09)	-0.23	(-0.97)
PPNS									
-FFM	10.3	16.7	0.62	5.1	0.50	2.65	(1.71)	1.07	(0.65)
-MFIV	9.2	15.0	0.61	5.6	0.67	2.43	(1.74)	1.64	(1.13)
-RNS	10.6	16.4	0.65	5.7	0.95	2.84	(1.87)	1.40	(0.83)
-VRP	10.2	16.1	0.63	5.6	0.70	2.72	(1.89)	1.40	(0.95)
-CFER	10.6	17.6	0.60	4.8	1.23	2.43	(1.58)	1.49	(0.84)
-ALL	10.6	16.3	0.65	5.9	1.18	3.06	(2.33)	1.91	(1.39)
Panel B: No short-sale ($0 \leq \omega_i$)									
GMVP	7.2	10.8	0.66	6.4	0.26	3.40	(1.84)	1.03	(0.57)
Hist	11.6	19.7	0.59	4.2	0.52	6.21	(2.11)	7.03	(2.07)
Q-RP	-0.1	37.7	0.00	-676.5	1.24	-9.71	(-1.20)	-10.3	(-1.01)
Q-CFER	14.0	19.2	0.73	4.8	1.66	8.60	(2.13)	8.63	(1.97)
Q-FULL	6.3	37.0	0.17	-676.3	1.57	-3.18	(-0.41)	-1.82	(-0.19)
CA-MFIV	8.0	11.5	0.70	6.8	0.63	3.76	(2.01)	1.23	(0.69)
CA-VRP	6.9	10.6	0.65	6.2	0.51	3.20	(1.78)	0.94	(0.54)
CA-RNS	8.3	11.0	0.75	7.3	0.85	4.40	(2.36)	2.23	(1.23)
CA-CFER	8.6	11.1	0.78	7.6	0.84	4.66	(2.50)	2.26	(1.26)

Continued on next page

Model	ExRet	Vol	SR	CE	TO	α_{FF3}		α_{FF5}	
Panel C: No short-sale and 10% upper bound ($0 \leq \omega_i \leq 0.1$)									
GMVP	7.4	10.8	0.68	6.6	0.25	3.55	(1.96)	1.21	(0.68)
Hist	9.6	17.6	0.54	3.8	0.50	4.30	(1.87)	4.95	(1.82)
Q-RP	9.4	24.9	0.38	-4.3	1.00	0.56	(0.13)	1.41	(0.30)
Q-CFER	13.9	14.2	0.98	11.0	1.43	8.89	(3.93)	7.62	(3.25)
Q-FULL	13.9	24.6	0.56	0.6	1.41	4.82	(1.24)	5.85	(1.40)
CA-MFIV	8.1	11.4	0.71	6.9	0.60	3.80	(2.06)	1.28	(0.72)
CA-VRP	7.1	10.6	0.67	6.4	0.49	3.33	(1.88)	1.09	(0.64)
CA-RNS	8.4	11.0	0.76	7.4	0.80	4.42	(2.42)	2.29	(1.30)
CA-CFER	8.8	11.0	0.80	7.9	0.79	4.87	(2.65)	2.57	(1.46)
Panel D: Constraints on the absolute size 10% ($\lvert\omega_i\rvert \leq 0.1$)									
GMVP	5.2	15.8	0.33	0.5	5.49	3.73	(1.01)	1.98	(0.54)
Hist	-25.2	115.4	-0.22	-1100.3	33.4	0.18	(0.01)	-1.07	(-0.04)
Q-RP	32.2	62.2	0.52	-58.2	22.5	16.0	(1.40)	24.4	(2.00)
Q-CFER	50.1	40.9	1.22	16.4	34.9	49.4	(5.15)	46.4	(4.67)
Q-FULL	61.0	63.3	0.96	-49.2	33.0	46.7	(3.54)	51.7	(3.54)
CA-MFIV	11.3	24.8	0.46	-3.2	17.5	7.42	(1.40)	7.52	(1.32)
CA-VRP	-0.4	27.0	-0.02	-17.5	20.4	-0.44	(-0.06)	-2.76	(-0.39)
CA-RNS	22.4	27.9	0.80	5.9	23.1	19.7	(3.10)	20.0	(2.99)
CA-CFER	21.2	29.1	0.73	1.7	25.0	18.9	(2.78)	15.3	(2.27)

Table 3.5. Pairwise performance comparison: Bootstrapped p-values

Entries report the bootstrapped pairwise comparison of the three performance measures, the out-of-sample (OOS) average excess return, Sharpe ratio (SR) and certainty equivalent (CE), for the three mean-variance strategies based on the option-based quantitative expected return. For each pair of two strategies, we generate 10,000 bootstrapped time-series of the OOS returns. Values in the columns labeled "Q-CFER" report the p-value of the null hypothesis that the Q-CFER strategy (with the historical covariance and NS10 constraint) performs no better than the competing strategy specified in each row. Values in the columns labeled "Q-RP" and "Q-FULL" are defined similarly. The universe of stocks is the S&P 500 member stocks and the OOS portfolio return period is from January 1997 to December 2017.

	Q-CFER			Q-RP			Q-FULL		
	ExRet	SR	CE	ExRet	SR	CE	ExRet	SR	CE
EW	0.06	0.00	0.00	0.55	0.95	0.99	0.13	0.66	0.86
VW	0.01	0.00	0.00	0.31	0.84	0.97	0.05	0.48	0.77
RK-CFER	0.06	0.00	0.00	0.52	0.92	0.98	0.12	0.57	0.80
RK-RP	0.06	0.00	0.00	0.55	0.95	0.99	0.13	0.66	0.86
RK-FULL	0.06	0.00	0.00	0.54	0.93	0.98	0.13	0.59	0.81
PPNS-FFM	0.13	0.02	0.02	0.61	0.96	0.99	0.21	0.75	0.89
PPNS-MFIV	0.00	0.00	0.00	0.33	0.92	0.98	0.05	0.63	0.85
PPNS-RNS	0.27	0.01	0.03	0.78	0.99	0.99	0.32	0.83	0.92
PPNS-VRP	0.05	0.00	0.00	0.53	0.94	0.99	0.11	0.64	0.85
PPNS-CFER	0.18	0.00	0.01	0.70	0.97	0.99	0.24	0.71	0.88
PPNS-ALL	0.06	0.01	0.01	0.61	0.98	1.00	0.16	0.82	0.93
GMVP	0.00	0.05	0.01	0.32	0.99	0.99	0.06	0.90	0.90
Hist	0.16	0.03	0.03	0.56	0.84	0.93	0.23	0.58	0.74
Q-CFER	—	—	—	0.89	1.00	1.00	0.52	1.00	1.00
Q-RP	0.11	0.00	0.00	—	—	—	0.01	0.00	0.00
Q-FULL	0.48	0.00	0.00	0.99	1.00	1.00	—	—	—
CA-MFIV	0.00	0.06	0.01	0.38	0.99	0.99	0.08	0.91	0.92
CA-VRP	0.00	0.04	0.01	0.30	0.99	0.99	0.06	0.89	0.89
CA-RNS	0.00	0.11	0.02	0.41	1.00	0.99	0.09	0.96	0.93
CA-CFER	0.00	0.16	0.04	0.45	1.00	1.00	0.11	0.98	0.95

Table 3.6. Choice of the covariance matrix: Unconstrained strategies

Entries report the performance measures of unconstrained mean-variance portfolio strategies. In each Panel, the performance of the mean-variance portfolios are calculated using the expected returns vector specified in the title of Panel and the covariance matrix specified in each row. For the first four models of the covariance matrix, the former part of the model name represents the specification for the variance, either historical volatility (HV), risk-adjusted option-implied volatility (RIV). The latter part represents the specification for the correlation, either historical correlation (HC) or zero correlation (ZC). LWI stands for the historical covariance matrix shrunk by Ledoit and Wolf (2004) method toward the scaled identity matrix. The universe of stocks is the S&P 500 member stocks and the out-of-sample (OOS) portfolio return period is from January 1997 to December 2017. The unit of the OOS excess return (ExRet), the volatility (Vol), the Fama and French (1993) three-factor alpha (α_{FF3}) and the Fama and French (2018) five-factor alpha (α_{FF5}) is percent per year. The t-statistics of alphas are adjusted for heteroskedasticity and serial correlations and reported in the parentheses. The certainty equivalent (CE) is calculated based on the power utility with the relative risk-aversion parameter of five. SR and TO stand for the Sharpe ratio and turnover rate, respectively.

Model	ExRet	Vol	SR	CE	TO	α_{FF3}		α_{FF5}	
Panel A: Global minimum Variance portfolio (GMVP)									
HV-HC	4.4	16.3	0.27	-0.8	5.92	2.63	(0.71)	0.62	(0.17)
RIV-HC	8.0	9.1	0.89	8.1	3.67	5.46	(3.09)	3.81	(2.12)
HV-ZC	9.5	13.6	0.69	6.8	0.07	2.38	(2.01)	-0.21	(-0.21)
RIV-ZC	8.3	10.2	0.82	8.0	0.53	3.79	(3.05)	1.37	(1.18)
LWI	5.8	12.0	0.48	4.1	2.38	4.06	(1.59)	1.78	(0.69)
Panel B: CFER-based quantitative expected returns (Q-CFER)									
HV-HC	720.6	861.0	0.84	-1171.2	922.4	762.7	(2.84)	832.9	(2.87)
RIV-HC	1958.3	28067.7	0.07	-730.2	20263.0	851.9	(0.15)	-506.0	(-0.09)
HV-ZC	69.8	65.2	1.07	-25.1	17.7	68.3	(3.78)	69.7	(3.83)
RIV-ZC	-275.5	4490.6	-0.06	-1148.0	1023.2	-152.2	(-0.15)	-149.4	(-0.14)
LWI	255.5	279.1	0.92	-678.6	174.7	263.4	(3.35)	288.3	(3.27)
Panel C: Fully estimated quantitative expected returns (Q-FULL)									
HV-HC	850.0	1151.7	0.74	-1064.7	891.2	815.7	(2.30)	949.0	(2.27)
RIV-HC	3797.7	28826.6	0.13	-730.4	22047.0	2343.1	(0.39)	1056.2	(0.17)
HV-ZC	97.5	201.0	0.49	-1044.4	30.0	29.1	(0.90)	65.6	(1.81)
RIV-ZC	-21.8	4737.4	0.00	-1003.5	995.3	-533.3	(-0.53)	-800.7	(-0.74)
LWI	304.6	384.3	0.79	-875.1	221.5	277.4	(2.49)	320.2	(2.41)
Panel D: CFER-based characteristic-adjusted returns (CA-CFER)									
HV-HC	58.6	79.3	0.74	-233.5	97.7	56.0	(3.42)	55.8	(3.12)
RIV-HC	1383.3	6173.5	0.22	-868.4	4687.2	1306.3	(1.25)	1350.6	(1.19)
HV-ZC	16.1	16.1	1.00	11.6	3.4	8.7	(4.61)	6.5	(3.26)
RIV-ZC	227.4	883.9	0.26	-1104.8	430.3	258.1	(1.24)	290.4	(1.20)
LWI	29.3	29.9	0.98	10.5	29.7	27.3	(4.41)	24.8	(3.80)

Table 3.7. Choice of the covariance matrix: NS10 constrained strategies

Entries report the performance measures of mean-variance portfolio strategies with the NS10 weight constraint ($0 \leq \omega_{i,t} \leq 0.1$). In each Panel, the performance of the mean-variance portfolios are calculated using the expected returns vector specified in the title of Panel and the covariance matrix specified in each row. For the first four models of the covariance matrix, the former part of the model name represents the specification for the variance, either historical volatility (HV), risk-adjusted option-implied volatility (RIV). The latter part represents the specification for the correlation, either historical correlation (HC) or zero correlation (ZC). LWI stands for the historical covariance matrix shrunk by Ledoit and Wolf (2004) method toward the scaled identity matrix. The universe of stocks is the S&P 500 member stocks and the out-of-sample (OOS) portfolio return period is from January 1997 to December 2017. The unit of the OOS excess return (ExRet), the volatility (Vol), the Fama and French (1993) three-factor alpha (α_{FF3}) and the Fama and French (2018) five-factor alpha (α_{FF5}) is percent per year. The t-statistics of alphas are adjusted for heteroskedasticity and serial correlations and reported in the parentheses. The certainty equivalent (CE) is calculated based on the power utility with the relative risk-aversion parameter of five. SR and TO stand for the Sharpe ratio and turnover rate, respectively.

Model	ExRet	Vol	SR	CE	TO	α_{FF3}		α_{FF5}	
Panel A: Global minimum Variance portfolio (GMVP)									
HV-HC	7.4	10.8	0.68	6.6	0.25	3.55	(1.96)	1.21	(0.68)
RIV-HC	6.1	9.7	0.63	5.8	0.98	3.05	(1.76)	0.67	(0.41)
HV-ZC	9.3	13.8	0.67	6.4	0.07	2.46	(2.03)	-0.05	(-0.05)
RIV-ZC	8.3	11.4	0.72	7.1	0.39	2.95	(2.32)	0.15	(0.14)
LWI	7.4	10.8	0.69	6.6	0.24	3.56	(1.96)	1.17	(0.67)
Panel B: CFER-based quantitative expected returns (Q-CFER)									
HV-HC	13.9	14.2	0.98	11.0	1.43	8.89	(3.93)	7.62	(3.25)
RIV-HC	13.0	15.4	0.85	9.4	1.78	7.01	(3.07)	5.08	(2.25)
HV-ZC	18.5	19.2	0.96	11.7	1.74	10.43	(3.81)	9.63	(3.29)
RIV-ZC	15.6	18.9	0.83	9.1	1.78	7.79	(2.86)	6.02	(2.23)
LWI	14.0	14.2	0.98	11.0	1.43	8.90	(3.89)	7.69	(3.27)
Panel C: Fully estimated quantitative expected returns (Q-FULL)									
HV-HC	13.9	24.6	0.56	0.6	1.41	4.82	(1.24)	5.85	(1.40)
RIV-HC	15.7	22.3	0.70	5.4	1.57	6.89	(2.00)	6.60	(1.77)
HV-ZC	15.3	38.7	0.40	-20.3	1.31	0.41	(0.07)	4.99	(0.88)
RIV-ZC	16.2	34.5	0.47	-11.4	1.33	1.78	(0.37)	6.12	(1.26)
LWI	14.0	24.9	0.56	0.4	1.41	4.82	(1.23)	5.93	(1.41)
Panel D: CFER-based characteristic-adjusted returns (CA-CFER)									
HV-HC	8.8	11.0	0.80	7.9	0.79	4.87	(2.65)	2.57	(1.46)
RIV-HC	10.3	12.3	0.84	8.6	1.70	5.78	(2.86)	3.24	(1.75)
HV-ZC	12.5	14.7	0.85	9.0	1.28	5.45	(3.85)	3.34	(2.28)
RIV-ZC	12.0	13.3	0.90	9.5	1.70	6.20	(3.40)	3.64	(2.10)
LWI	9.0	11.0	0.82	8.0	0.79	4.99	(2.73)	2.68	(1.54)

Table 3.8. Mean-variance portfolios: Randomly selected 500 optionable stocks

Entries in Panel A reports the performance measures of the benchmark portfolios strategies which are not based on the mean-variance optimization problem. EW and VW stand for the equally- and value-weighted portfolio strategy, respectively. RK-X stands for the rank-based strategy in which the portfolio weight is determined by the cross-sectional ranking of the variable X, where X is either CFER, the shifted risk premium (RP), or the sum of these two (FULL). PPNS-X stands for the parametric portfolio with no short-sale constraints. See Table 3.2 for the set of characteristic variables for each PPNS strategy. Panel B reports the performance measures of the mean-variance portfolios with nine alternative expected return vector specifications under the NS10 constraint, $0 \leq \omega_{i,t} \leq 0.1$. We consider the minimum-variance portfolio (GMVP), historical mean returns (Hist), the scaled risk premium (Q-RP), CFER (Q-CFER), the sum of the scaled risk premium and CFER (Q-FULL) and four characteristic-adjusted (CA-) returns based on MFIV, VRP, RNS and CFER. We use the historical covariance matrix. The universe of stocks is the randomly selected 500 optionable stocks. In each year y, we keep stocks with non-missing option-implied data and non-missing 630 daily historical returns on each end of month days in y. Then, we randomly choose 500 stocks from these stocks. We update the set of stocks in every January. The out-of-sample (OOS) portfolio return period is from January 1997 to December 2017. The unit of the OOS excess return (ExRet), the volatility (Vol), the Fama and French (1993) three-factor alpha (α_{FF3}) and the Fama and French (2018) five-factor alpha (α_{FF5}) is percent per year. The t-statistics of alphas are adjusted for heteroskedasticity and serial correlations and reported in the parentheses. The certainty equivalent (CE) is calculated based on the power utility with the relative risk-aversion parameter of five. SR and TO stand for the Sharpe ratio and turnover rate, respectively.

Model	ExRet	Vol	SR	CE	TO	α_{FF3}		α_{FF5}	
Panel A: Benchmark strategies									
EW	11.1	20.4	0.55	2.2	0.20	0.93	(0.65)	0.54	(0.41)
VW	6.6	15.9	0.42	2.2	0.13	-0.79	(-1.22)	-1.46	(-2.60)
RK-CFER	11.2	21.6	0.52	1.0	0.98	0.59	(0.40)	0.64	(0.46)
RK-RP	11.5	20.3	0.56	2.7	1.05	1.39	(0.89)	0.84	(0.59)
RK-FULL	11.5	21.3	0.54	1.6	1.02	1.12	(0.77)	1.14	(0.82)
PPNS-FFM	11.2	20.7	0.54	2.3	0.59	1.95	(1.09)	2.27	(1.18)
PPNS-MFIV	10.9	20.1	0.54	3.2	0.69	2.37	(1.25)	2.30	(1.09)
PPNS-RNS	15.1	21.7	0.70	5.3	1.11	4.75	(2.94)	4.16	(2.31)
PPNS-VRP	13.4	21.4	0.62	4.0	0.76	3.84	(2.01)	3.77	(1.83)
PPNS-CFER	17.0	24.1	0.71	4.5	1.28	5.97	(2.61)	6.57	(2.66)
PPNS-ALL	15.1	22.6	0.67	4.0	1.25	4.96	(2.32)	4.74	(2.16)
Panel B: Mean-variance strategies under the NS10 constraint									
GMVP	7.7	11.5	0.67	6.2	0.31	3.07	(1.72)	0.73	(0.43)
Hist	4.0	26.2	0.15	-10.7	0.54	-3.36	(-1.15)	-1.40	(-0.42)
Q-RP	26.0	33.5	0.78	2.3	1.18	14.22	(2.88)	16.33	(3.20)
Q-CFER	21.1	18.9	1.12	14.5	1.56	14.15	(4.37)	11.39	(4.19)
Q-FULL	33.9	33.6	1.01	10.6	1.45	21.54	(4.16)	23.15	(4.52)
CA-MFIV	7.3	12.4	0.59	5.3	0.58	2.19	(1.16)	-0.06	(-0.03)
CA-VRP	7.9	11.4	0.69	6.4	0.46	3.23	(1.86)	0.93	(0.55)
CA-RNS	9.5	11.9	0.80	7.7	0.80	4.57	(2.66)	2.24	(1.37)
CA-CFER	9.7	11.7	0.83	8.1	0.65	4.87	(2.66)	2.80	(1.69)

Table 3.9. Equivalent transaction costs of the Q-CFER strategy

Entries report the equivalent transaction costs of the mean-variance portfolio strategy based on the Q-CFER expected returns and the historical covariance matrix with the NS10 weight constraint. A positive number indicate that the Q-CFER strategy has superior after-transaction costs performance measure, either the out-of-sample (OOS) mean excess return or the Sharpe ratio, than the competing strategy specified in each row as long as the transaction costs are lower than the displayed value. The unit is basis points per transaction. The entries with "D" indicate that the performance statistics of the Q-CFER strategy is lower than the competing strategy before transaction costs are taken into account. The universe of stocks is the S&P 500 member stocks and the OOS portfolio return period is from January 1997 to December 2017.

	S&P 500 universe		Random 500 universe	
	TC^{mean}	TC^{SR}	TC^{mean}	TC^{SR}
EW	26.1	35.4	61.5	66.0
VW	40.0	41.8	84.8	78.9
RK-RP	78.1	79.4	145.2	136.3
RK-CFER	84.2	83.1	160.1	151.1
RK-FULL	79.3	80.1	149.3	139.5
PPNS-FFM	32.4	43.0	86.1	89.8
PPNS-MFIV	52.3	55.3	98.1	99.7
PPNS-RNS	56.9	64.6	111.6	112.2
PPNS-VRP	42.5	50.8	80.7	87.5
PPNS-CFER	139.5	102.8	123.2	117.1
PPNS-ALL	114.6	99.1	166.1	139.2
GMVP (NS10)	46.5	32.3	89.9	67.8
Hist (NS10)	39.1	50.6	140.5	130.3
Q-RP (NS10)	88.6	83.3	D	60.4
Q-FULL (NS10)	17.2	80.2	D	23.4
CA-MFIV (NS10)	59.0	47.8	117.9	123.6
CA-VRP (NS10)	60.8	48.2	101.2	85.8
CA-RNS (NS10)	73.1	64.8	129.2	182.9
CA-CFER (NS10)	66.6	51.8	105.5	90.9

Table 3.10. Performance of staggered mean-variance portfolios

Entries report the performance measures and the equivalent transaction costs of the four selected non-staggered strategies (EW, VW, GMVP, and Q-CFER) and the staggered rebalancing Q-CFER strategies with three alternative rebalancing frequencies: quarterly (3m), biannual (6m), and annual (12m) rebalancing. The GMVP and Q-CFER strategies are calculated using the historical covariance matrix with the NS10 weight constraint, $0 \leq \omega_{i,t} \leq 0.1$. The universe of stocks for Panel A (Panel B) is the S&P 500 member stocks (randomly selected 500 optionable stocks) and the out-of-sample (OOS) portfolio return period is from January 1997 to December 2017. The unit of the OOS mean excess return and the volatility is percent per year. The certainty equivalent is calculated based on the power utility with the relative risk-aversion parameter of five. The equivalent transaction costs $TC^{SR}_{\text{Q-CFER,X}}$ report the transaction costs which equate the after-transaction costs Sharpe ratio of the Q-CFER strategies and the competing strategy X (either EW, VW, or GMVP). A positive equivalent transaction cost indicates that the Q-CFER strategy has a higher after-transaction costs Sharpe ratio compared to the competing strategy as long as the transaction costs are lower than the displayed value. The entries with "S" indicate that the Q-CFER strategy always has a higher after-transaction costs Sharpe ratio for any positive transaction cost value.

	Non-staggered strategies				Staggered Q-CFER		
	EW	VW	GMVP	Q-CFER	3m	6m	12m
Panel A: S&P 500 universe							
Excess return	9.69	7.17	7.36	13.92	9.55	9.20	8.69
Volatility	16.86	14.80	10.78	14.17	12.30	11.57	11.20
Sharpe ratio	0.58	0.48	0.68	0.98	0.78	0.80	0.78
Certainty Equiv.	4.40	3.67	6.55	11.01	7.79	7.91	7.57
Turnover	0.08	0.02	0.25	1.43	0.50	0.27	0.15
$TC^{SR}_{\text{Q-CFER, EW}}$	—	—	—	35.4	46.8	100.1	186.4
$TC^{SR}_{\text{Q-CFER, VW}}$	—	—	—	41.8	62.2	120.1	199.2
$TC^{SR}_{\text{Q-CFER, GMVP}}$	—	—	—	32.3	45.0	S	S
Panel B: Randomly selected 500 stocks universe							
Excess return	11.13	6.63	7.71	21.12	13.29	9.52	6.52
Volatility	20.42	15.93	11.51	18.87	15.29	13.72	10.98
Sharpe ratio	0.55	0.42	0.67	1.12	0.87	0.69	0.59
Certainty Equiv.	2.21	2.22	6.22	14.48	9.19	6.55	5.52
Turnover	0.20	0.13	0.31	1.56	0.56	0.31	0.18
$TC^{SR}_{\text{Q-CFER, EW}}$	—	—	—	66.0	101.2	97.2	61.4
$TC^{SR}_{\text{Q-CFER, VW}}$	—	—	—	78.9	132.9	160.0	177.6
$TC^{SR}_{\text{Q-CFER, GMVP}}$	—	—	—	67.8	175.5	S	59.7

Conclusions

In this thesis, I have conducted theoretical and empirical studies on the *contribution of frictions to expected returns* (CFER) of individual equities. CFER represents the effect of market frictions on the expected asset returns not attributable to the conventional covariance risk premium term, that is, the covariance between the asset returns and the stochastic discount factor (agents' marginal utility).

In Chapter One, I have developed the formal theoretical framework for my analysis on CFER. Then, I have theoretically and empirically documented that the CFER term can be reliably estimated by properly scaled deviations from put-call parity, that is, the difference between the underlying stock price and the synthetic stock price calculated from option prices. My empirical findings on U.S. common stocks show that the estimated CFER has strong return predictive power. Moreover, I have confirmed that this strong return predictive ability of CFER is robust to the recent data snooping concerns and does not stem from omitted risk factors. I have also documented that the size of market frictions is a more pertinent explanation to the return predictability of CFER than option trading activity. The results in Chapter One suggest that market frictions, especially transaction costs, have a non-negligible effect even on optionable stocks, which tend to be large and liquidly traded stocks. In particular, I have theoretically shown that the upper bound of the alpha of CFER-sorted spread portfolios is at least twice the round-trip transaction costs. This result accommodates the sizable alpha of the CFER-sorted spread portfolio I have found.

In Chapter Two, I have examined the implication of CFER on the risk-neutral expected asset return and the return predictive ability of the risk-neutral skewness (RNS). To this end, I have shown that a non-zero CFER is equivalent to the violation of the *martingale restriction* (MR), that is, the deviation of the expected risk-neutral return from the risk-free rate caused by the presence of market frictions. Then, I have shown that the Bakshi et al. (2003) (BKM) formulae for risk-neutral moments (RNMs) do not correctly estimate RNMs if MR is violated, due to their implicit assumption that the underlying asset satisfies MR. To remedy this drawback of the original BKM formulae, I propose the *generalized* BKM formulae, which account for the possible violation of MR, that is, a non-zero CFER. My empirical analysis in Chapter Two have revealed that the violation of MR has an important implication on the return predictive power of RNS; the estimated RNS based on the original BKM formula (O-RNS) strongly predicts future stock returns, whereas the estimated RNS based on my generalized BKM formula does not. Furthermore, I have provided evidence that the predictive power of O-RNS stems from its estimation bias component caused by the violation of MR, which is highly correlated with CFER. O-RNS predicts future returns because it signals the effect of frictions on the expected stock returns via its correlation with CFER.

In Chapter Three, I have applied my CFER framework to the Markowitz (1952) mean-variance portfolio construction. To this end, I have generalized Martin and Wagner's (2018) formula for the expected stock return by allowing the existence of the CFER-type market frictions, where the expected stock return is expressed by the sum of CFER and the scaled risk-neutral simple stock variance of Martin (2017). Relying on this theoretical result, I have examined mean-variance portfolio strategies which use the option-based expected stock returns as an input to the mean-variance optimization. My empirical analysis has shown that the mean-variance portfolio strategy, for which I use the estimated CFER as the estimate of the expected stock returns (the Q-CFER strategy), outperforms other portfolio strategies which have been documented to have good performance, including minimum variance portfolios and option-based portfolio strategies proposed by the existing literature. The outperformance of the Q-CFER strategy suggests that incorporating the effect of market frictions into the estimation of the expected return is of importance, because it allows us to improve the portfolio expected return without much increasing the portfolio variance due to the fact that CFER is not a compensation for risk exposures. Moreover, I have found that imposing constraints on portfolio weights dramatically improves the performance of mean-variance portfolios. This is the first empirical evidence that the *shrinkage-like* effect of the weight constraints on the *mean*-variance portfolio problem, which is theoretically documented by Jagannathan and Ma (2002, 2003), is useful for portfolio selection once the option-implied expected return is employed.

The three studies in my thesis contribute to the literature on asset pricing and the informational content of options in four important ways. First, my empirical findings suggest that the estimated CFER provides useful information on the expected stock returns. Specifically, Chapter One has documented that the estimated CFER has strong return predictive power. Moreover, my analysis in Chapter Three has shown that the estimated CFER provides practically useful information for the portfolio selection problem, as my CFER-based mean-variance portfolio outperforms various other portfolio strategies.

Second, the estimated CFER provides a forward-looking, model-free estimate of the effect of market frictions on the expected stock returns. This contributes to the large body of the theoretical and empirical literature on the asset pricing and market frictions, by helping us to better understand how market frictions affect the time-series and cross-sectional variations in the expected stock returns. For instance, my analysis in Chapter One suggests that market frictions affect the expected return of even large optionable stocks, in contrast to the findings in Hou et al. (2018) that almost all previously documented friction-related anomalies vanish once micro-cap stocks are weighted less in the analysis. Utilizing the theoretically-founded, quantitative nature of CFER, I have documented that the sizable return predictive power of CFER is in line with the degree of market frictions, especially that of transaction costs, for trading large stocks.

Third, my findings in this thesis shed light on the predictive mechanism of various option-implied return predicting variables under a unified framework. In Chapter One, I have shown that my estimate of CFER has superior properties compared to previously proposed measures based on deviations from put-call parity such as the implied volatility spread (e.g., Bali and Hovakimian, 2009; Cremers and Weinbaum, 2010) and the DOTS measure (Goncalves-Pinto et al., 2019). Chapter Two has revealed that the return predictive ability of O-RNS (Rehman and Vilkov, 2012; Stilger et al., 2017; Gkionis et al., 2018; Bali et al., 2018; Borochin and Zhao, 2018; Chordia et al., 2019) and that of the Xing et al. (2010) implied volatility slope measure stem from their mechanical correlation with CFER which arises in the presence of market frictions (i.e., the violation of MR). My analysis in this thesis collectively suggests that these option-implied variables can predict future stock returns because they signal the effect of market frictions on the expected stock returns via their correlation with CFER.

Finally, my analysis in Chapter Two would call for a revisit of the voluminous theoretical and empirical literature which is based on the martingale assumption on asset prices. While I have focused on the estimation of RNS in Chapter Two, my analysis there is relevant to any studies on the informational content of option prices which relies on the martingale assumption, including the estimation of risk-neutral distributions, pricing kernels, risk-aversion parameters, various portfolio allocation applications among others. These topics are best left for future research.

Bibliography

Acharya, V. V. and L. H. Pedersen (2005): "Asset Pricing with Liquidity Risk," *Journal of Financial Economics*, 77, 375–410.

Adrian, T., E. Etula, and T. Muir (2014): "Financial Intermediaries and the Cross-Section of Asset Returns," *Journal of Finance*, 69, 2557–2596.

Aït-Sahalia, Y. and M. W. Brandt (2008): "Consumption and Portfolio Choice with Option-Implied State Prices," NBER Working Paper, No. 13854.

Aït-Sahalia, Y. and A. W. Lo (1998): "Nonparametric Estimation of State-Price Densities Implicit in Financial Asset Prices," *Journal of Finance*, 53, 499–547.

Amihud, Y. (2002): "Illiquidity and Stock Returns: Cross-Section and Time-Series Effects," *Journal of Financial Markets*, 5, 31–56.

Ammann, M. and A. Feser (2019): "Robust Estimation of Risk-Neutral Moments," *Journal of Futures Markets*, forthcoming.

Ang, A., R. J. Hodrick, Y. Xing, and X. Zhang (2006): "The Cross-Section of Volatility and Expected Returns," *Journal of Finance*, 61, 259–299.

Asness, C. S., T. J. Moskowitz, and L. H. Pedersen (2013): "Value and Momentum Everywhere," *Journal of Finance*, 68, 929–985.

Asquith, P., P. A. Pathak, and J. R. Ritter (2005): "Short Interest, Institutional Ownership, and Stock Returns," *Journal of Financial Economics*, 78, 243–276.

Bakshi, G. S., N. Kapadia, and D. Madan (2003): "Stock Return Characteristics, Skew Laws, and the Differential Pricing of Individual Equity," *Review of Financial Studies*, 16, 101–143.

Bali, T. G., N. Cakici, and R. F. Whitelaw (2011): "Maxing out: Stocks as lotteries and the cross-section of expected returns," *Journal of Financial Economics*, 99, 427–446.

Bali, T. G., R. F. Engle, and S. Murray (2016): *Empirical Asset Pricing: The Cross Section of Stock Returns*, Wiley.

Bali, T. G. and A. Hovakimian (2009): "Volatility Spreads and Expected Stock Returns," *Management Science*, 55, 1797–1812.

Bali, T. G., J. Hu, and S. Murray (2018): "Option Implied Volatility, Skewness, and Kurtosis and the Cross-Section of Expected Stock Returns," SSRN Working Paper.

Barberis, N. and M. Huang (2008): "Stocks as Lotteries: The Implications of Probability Weighting for Security Prices," *American Economic Review*, 98, 2066–2100.

Battalio, R. and P. Schultz (2006): "Options and the Bubble," *Journal of Finance*, 61, 2071–2102.

Best, M. J. and R. R. Grauer (1991): "On the Sensitivity of Mean-Variance-Efficient Portfolios to Changes in Asset Means: Some Analytical and Computational Results," *Review of Financial Studies*, 4, 315–342.

Black, F. and R. B. Litterman (1991): "Asset allocation: Combining investor views with market equilibrium," *Journal of Fixed Income*, 1, 7–18.

Black, F. and M. Scholes (1973): "The Pricing of Options and Corporate Liabilities," *Journal of Political Economy*, 81, 637–654.

Bliss, R. R. and N. Panigirtzoglou (2002): "Testing the stability of implied probability density functions," *Journal of Banking and Finance*, 26, 381–422.

Boehme, R. D., B. R. Danielsen, and S. M. Sorescu (2006): "Short-Sale Constraints, Differences of Opinion, and Overvaluation," *Journal of Financial and Quantitative Analysis*, 41, 455–487.

Bollen, N. P. and R. E. Whaley (2004): "Does Net Buying Pressure Affect the Shape of Implied Volatility Functions?" *Journal of Finance*, 59, 711–753.

Borochin, P. and Y. Zhao (2018): "What Information Does Risk Neutral Skewness Contain? Evidence From Momentum Crashes," SSRN Working Paper.

Boyer, B. H., T. Mitton, and K. Vorkink (2010): "Expected idiosyncratic skewness," *Review of Financial Studies*, 23, 169–202.

Brandt, M. W., P. Santa-Clara, and R. Valkanov (2009): "Parametric portfolio policies: Exploiting characteristics in the cross-section of equity returns," *Review of Financial Studies*, 22, 3411–3447.

Brennan, M. J. and A. W. Wang (2010): "The Mispricing Return Premium," *Review of Financial Studies*, 23, 3437–3468.

Brunnermeier, M. K. and L. H. Pedersen (2009): "Market Liquidity and Funding Liquidity," *Review of Financial Studies*, 22, 2201–2238.

Carhart, M. M. (1997): "On Persistence in Mutual Fund Performance," *Journal of Finance*, 52, 57–82.

Carr, P. P. and D. Madan (2001): "Optimal positioning in derivative securities," *Quantitative Finance*, 1, 19–37.

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

Chabakauri, G. (2013): "Dynamic equilibrium with two stocks, heterogeneous investors, and portfolio constraints," *Review of Financial Studies*, 26, 3104–3141.

Chan, L. K. C., J. Karceski, and J. Lakonishok (1999): "On Portfolio Optimization: Forecasting Covariances and Choosing the Risk Model," *Review of Financial Studies*, 12, 937–974.

Chen, J., H. Hong, and J. C. Stein (2002): "Breadth of Ownership and Stock Returns," *Journal of Financial Economics*, 66, 171–205.

Chopra, V. K. and W. T. Ziemba (1993): "The Effect of Errors in Means, Variances, and Covariances on Optimal Portfolio Choice," *Journal of Portfolio Management*, 19, 6–11.

Chordia, T., T. C. Lin, and V. Xiang (2019): "Risk-neutral Skewness and Informed Trading," SSRN Working Paper.

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

Christoffersen, P., K. Jacobs, and B. Y. Chang (2013): "Forecasting with Option-Implied Information," in *Handbook of Economic Forecasting, Volume 2*, ed. by G. Elliott and A. Timmermann, Elsevier, 581–656.

Cochrane, J. H. (2005): *Asset Pricing, Revised Edition*, Princeton University Press.

Conrad, J., R. F. Dittmar, and E. Ghysels (2013): "Ex Ante Skewness and Expected Stock Returns," *Journal of Finance*, 68, 85–124.

Constantinides, G. M., J. B. Donaldson, and R. Mehra (2002): "Junior Can't Borrow : A New Perspective on the Equity Premium Puzzle," *Quarterly Journal of Economics*, 117, 269–296.

Cox, J. C., S. A. Ross, and M. Rubinstein (1979): "Option Pricing: A Simplified Approach," *Journal of Financial Economics*, 7, 229–263.

Cremers, M. and D. Weinbaum (2010): "Deviations from Put-Call Parity and Stock Return Predictability," *Journal of Financial and Quantitative Analysis*, 45, 335–367.

Davis, J. L., E. F. Fama, and K. R. French (2000): "Characteristics, Covariances, and Average Returns: 1929 to 1997," *Journal of Finance*, 55, 389–406.

D'Avolio, G. (2002): "The market for borrowing stock," *Journal of Financial Economics*, 66, 271–306.

DeMiguel, V., L. Garlappi, F. J. Nogales, and R. Uppal (2009a): "A Generalized Approach to Portfolio Optimization: Improving Performance by Constraining Portfolio Norms," *Management Science*, 55, 798–812.

DeMiguel, V., L. Garlappi, and R. Uppal (2009b): "Optimal versus Naive Diversification: How Inefficient Is the 1/N Portfolio Strategy?" *Review of Financial Studies*, 22, 1915–1953.

DeMiguel, V., A. Martin-Utrera, F. J. Nogales, and R. Uppal (2019): "A Transaction-Cost Perspective on the Multitude of Firm Characteristics," SSRN Working Paper.

DeMiguel, V., Y. Plyakha, R. Uppal, and G. Vilkov (2013): "Improving Portfolio Selection Using Option-Implied Volatility and Skewness," *Journal of Financial and Quantitative Analysis*, 48, 1813–1845.

Dennis, P. and S. Mayhew (2009): "Microstructural biases in empirical tests of option pricing models," *Review of Derivatives Research*, 12, 169–191.

Doran, J. S. and K. Krieger (2010): "Implications for asset returns in the implied volatility skew," *Financial Analysts Journal*, 66, 65–76.

Drechsler, I. and Q. F. Drechsler (2014): "The Shorting Premium and Asset Pricing Anomalies," NBER Working Paper, No. 20282.

Duffie, D., N. Gârleanu, and L. H. Pedersen (2002): "Securities lending, shorting, and pricing," *Journal of Financial Economics*, 66, 307–339.

Easley, D., M. O'Hara, and P. S. Srinivas (1998): "Option Volume and Stock Prices: Evidence on Where Informed Traders Trade," *Journal of Finance*, 53, 431–465.

Engelberg, J. E., A. V. Reed, and M. C. Ringgenberg (2018): "Short Selling Risk," *Journal of Finance*, 73, 755–786.

Engle, R. and B. Neri (2010): "The Impact of Hedging Costs on the Bid and Ask Spread in the Options Market," Working paper, New York University.

Fama, E. F. (1991): "Efficient Capital Markets: II," *Journal of Finance*, 46, 1575–1617.

Fama, E. F. and K. R. French (1993): "Common Risk Factors in the Returns on Stocks and Bonds," *Journal of Financial Economics*, 33, 3–56.

——— (2002): "The Equity Premium," *Journal of Finance*, 47, 427–465.

——— (2015): "A Five-Factor Asset Pricing Model," *Journal of Financial Economics*, 116, 1–22.

——— (2018): "Choosing Factors," *Journal of Financial Economics*, 128, 234–252.

Fama, E. F. and J. D. MacBeth (1973): "Risk, Return, and Equilibrium: Empirical Tests," *Journal of Political Economy*, 81, 607–636.

Fan, J., J. Zhang, and K. Yu (2012): "Vast portfolio selection with gross-exposure constraints," *Journal of the American Statistical Association*, 107, 592–606.

Filippou, I., P. Garcia-Ares, and F. Zapatero (2018): "Demand for Lotteries: The Choice between Stocks and Options," SSRN Working Paper.

Frazzini, A., R. Israel, and T. J. Moskowitz (2014): "Trading Costs of Asset Pricing Anomalies," Chicago Booth Research Paper No. 14-05, University of Chicago.

Frazzini, A. and L. H. Pedersen (2012): "Embedded Leverage," NBER Working Paper No. 18558.

——— (2014): "Betting Against Beta," *Journal of Financial Economics*, 111, 1–25.

French, K. R. (2008): "Presidential address: The cost of active investing," *Journal of Finance*, 63, 1537–1573.

Fu, X., Y. E. Arisoy, M. B. Shackleton, and M. Umutlu (2016): "Option-Implied Volatility Measures and Stock Return Predictability," *Journal of Derivatives*, 24, 58–78.

Gao, X. and J. R. Ritter (2010): "The Marketing of Seasoned Equity Offerings," *Journal of Financial Economics*, 97, 33–52.

Garlappi, L., R. Uppal, and T. Wang (2007): "Portfolio Selection with Parameter and Model Uncertainty: A Multi-Prior Approach," *Review of Financial Studies*, 20, 41–81.

Gârleanu, N. and L. H. Pedersen (2011): "Margin-based Asset Pricing and Deviations from the Law of One Price," *Review of Financial Studies*, 24, 1980–2022.

Gârleanu, N., L. H. Pedersen, and A. M. Poteshman (2009): "Demand-Based Option Pricing," *Review of Financial Studies*, 22, 4259–4299.

Giamouridis, D. and G. Skiadopoulos (2011): "The Informational Content of Financial Options for Quantitative Asset Management: A Review," in *The Oxford Handbook of Quantitative Asset Management*, ed. by B. Scherer and K. Winston, Oxford University Press, 243–265.

Gibbons, M. R., S. A. Ross, and J. Shanken (1989): "A Test of the Efficiency of a Given Portfolio," *Econometrica*, 57, 1121–1152.

Gkionis, K., A. Kostakis, G. Skiadopoulos, and P. S. Stilger (2018): "Positive Stock Information in Out-of-the-Money Option Prices," SSRN Working Paper.

Goncalves-Pinto, L., B. D. Grundy, A. Hameed, T. van der Heijden, and Y. Zhu (2017): "Why Do Option Prices Predict Stock Returns? The Role of Price Pressure in the Stock Market," *Management Science*, forthcoming.

Green, J., J. R. M. Hand, and X. F. Zhang (2017): "The Characteristics that Provide Independent Information about Average U.S. Monthly Stock Returns," *Review of Financial Studies*, 30, 4389–4436.

Greene, W. H. (2003): *Econometric Analysis*, Fifth edition, Prentice Hall.

Gromb, D. and D. Vayanos (2010): "Limits of Arbitrage: The State of the Theory," *Annual Review of Financial Economics*, 2, 251–275.

Grundy, B. D., B. Lim, and P. Verwijmeren (2012): "Do option markets undo restrictions on short sales? Evidence from the 2008 short-sale ban," *Journal of Financial Economics*, 106, 331–348.

Guo, B., Q. Han, and D. Ryu (2013): "Is the KOSPI 200 Options Market Efficient? Parametric and Nonparametric Tests of the Martingale Restriction," *Journal of Futures Markets*, 33, 629–652.

Harris, R. D. F. and F. Qiao (2018): "Moment Risk Premia and the Cross-Section of Stock Returns," SSRN Working Paper.

Harrison, J. M. and D. M. Kreps (1979): "Martingales and arbitrage in multiperiod securities markets," *Journal of Economic Theory*, 20, 381–408.

Harrison, J. M. and S. R. Pliska (1981): "Martingales and Stochastic Integrals in the Theory of Continuous Trading," *Stochastic Processes and their Applications*, 11, 215–260.

Harvey, C. R. (2017): "Presidential Address: The Scientific Outlook in Financial Economics," *Journal of Finance*, 72, 1399–1440.

Harvey, C. R., Y. Liu, and H. Zhu (2016): "... and the Cross-Section of Expected Returns," *Review of Financial Studies*, 29, 5–68.

Harvey, C. R. and A. Siddique (2000): "Conditional Skewness in Asset Pricing Tests," *Journal of Finance*, 55, 1263–1295.

Hasbrouck, J. (2009): "Trading Costs and Returns for U.S. Equities: Estimating Effective Costs from Daily Data," *Journal of Finance*, 64, 1445–1477.

He, H. and D. M. Modest (1995): "Market Frictions and Consumption-Based Asset Pricing," *Journal of Political Economy*, 103, 94–117.

He, Z., B. Kelly, and A. Manela (2017): "Intermediary Asset Pricing: New Evidence from Many Asset Classes," *Journal of Financial Economics*, 126, 1–35.

Heaton, J. and D. J. Lucas (1996): "Evaluating the Effects of Incomplete Markets on Risk Sharing and Asset Pricing," *Journal of Political Economy*, 104, 443–487.

Hentchel, L. (2003): "Errors in Implied Volatility Estimation," *Journal of Financial and Quantitative Analysis*, 38, 779–810.

Hiraki, K. and G. Skiadopoulos (2019): "The Contribution of Frictions to Expected Return," SSRN Working Paper.

Hitzemann, S., M. Hofmann, M. Uhrig-Homburg, and C. Wagner (2017): "Margin Requirements and Equity Option Returns," SSRN Working Paper.

Hlouskova, J. and G. S. Lee (2001): "Legal Restrictions on Portfolio Holdings: Some Empirical Results," Economics Series 93, Institute for Advanced Studies.

Hou, K., S. Kim, and I. M. Werner (2016): "(Priced) Frictions," Fisher College of Business Working Paper Series WP 2016-03-019, Ohio State University.

Hou, K. and T. J. Moskowitz (2005): "Market Frictions, Price Delay, and the Cross-Section of Expected Returns," *Review of Financial Studies*, 18, 981–1020.

Hou, K., C. Xue, and L. Zhang (2015): "Digesting Anomalies: An Investment Approach," *Review of Financial Studies*, 28, 650–705.

——— (2018): "Replicating Anomalies," *Review of Financial Studies*, forthcoming.

Huang, H. H., K. Wang, and Z. Wang (2016): "A test of efficiency for the S&P 500 index option market using the generalized spectrum method," *Journal of Banking and Finance*, 64, 52–70.

Huang, R. D. and H. R. Stoll (1997): "The Components of the Bid-Ask Spread: A General Approach," *Review of Financial Studies*, 10, 995–1034.

Jagannathan, R. and T. Ma (2002): "Risk Reduction in Large Portfolios: Why Imposing the Wrong Constraints Helps," NBER Working Paper, No. 8922.

——— (2003): "Risk Reduction in Large Portfolios: Why Imposing the Wrong Constraints Helps," *Journal of Finance*, 58, 1651–1683.

Jarrow, R. (2016): "Bubbles and Multiple-Factor Asset Pricing Models," *International Journal of Theoretical and Applied Finance*, 19, 1–19.

Jegadeesh, N. and S. Titman (1993): "Returns to Buying Winners and Selling Losers: Implications for Stock Market Efficiency," *Journal of Finance*, 48, 65–91.

Jylhä, P. (2018): "Margin Requirements and the Security Market Line," *Journal of Finance*, 73, 1281–1321.

Kadan, O. and X. Tang (2017): "A Bound on Expected Stock Returns," SSRN Working Paper.

Kempf, A., O. Korn, and S. Sassning (2015): "Portfolio Optimization Using Forward-Looking Information," *Review of Finance*, 19, 467–490.

Kogan, L., I. Makarov, and R. Uppal (2007): "The equity risk premium and the riskfree rate in an economy with borrowing constraints," *Mathematics and Financial Economics*, 1, 1–19.

Kostakis, A., N. Panigirtzoglou, and G. Skiadopoulos (2011): "Market Timing with Option-Implied Distributions: A Forward-Looking Approach," *Management Science*, 57, 1231–1249.

Kozhan, R., A. Neuberger, and P. Schneider (2013): "The skew risk premium in the equity index market," *Review of Financial Studies*, 26, 2174–2203.

Kraus, A. and R. H. Litzenberger (1976): "Skewness Preference and the Valuation of Risk Assets," *Journal of Finance*, 31, 1085–1100.

Kumar, A. (2009): "Who Gambles in the Stock Market," *Journal of Finance*, 64, 1889–1933.

Lai, T. L., H. Xing, and Z. Chen (2011): "Mean-variance portfolio optimization when means and covariances are unknown," *Annals of Applied Statistics*, 5, 798–823.

Ledoit, O. and M. Wolf (2003): "Improved estimation of the covariance matrix of stock returns with an application to portfolio selection," *Journal of Empirical Finance*, 10, 603–621.

——— (2004): "A well-conditioned estimator for large-dimensional covariance matrices," *Journal of Multivariate Analysis*, 88, 365–411.

——— (2017): "Nonlinear shrinkage of the covariance matrix for portfolio selection: Markowitz meets goldilocks," *Review of Financial Studies*, 30, 4349–4388.

Leland, H. E. (1985): “Option Pricing and Replication with Transactions Costs,” *Journal of Finance*, 40, 1469–1484.

Lesmond, D. A., J. P. Ogden, and C. A. Trzcinka (1999): “A New Estimate of Transaction Costs,” *Review of Financial Studies*, 12, 1113–1141.

Lin, Y., T. Lehnert, and C. Wolff (2019): “Skewness risk premium: Theory and empirical evidence,” *International Review of Financial Analysis*, 63, 174–185.

Longstaff, F. A. (1995): “Option Pricing and the Martingale Restriction,” *Review of Financial Studies*, 8, 1091–1124.

Luttmer, E. G. J. (1996): “Asset Pricing in Economies with Frictions,” *Econometrica*, 64, 1439–1467.

Markowitz, H. (1952): “Portfolio Selection,” *Journal of Finance*, 7, 77–91.

Martin, I. (2017): “What is the expected return on the market?” *Quarterly Journal of Economics*, 132, 367–433.

Martin, I. and C. Wagner (2018): “What is the Expected Return on a Stock?” *Journal of Finance*, forthcoming.

McLean, R. D. and J. Pontiff (2016): “Does Academic Research Destroy Stock Return Predictability?” *Journal of Finance*, 71, 5–32.

Mehra, R. (2012): “Consumption-Based Asset Pricing Models,” *Annual Review of Financial Economics*, 4, 385–409.

Merton, R. C. (1980): “On estimating the expected return on the market,” *Journal of Financial Economics*, 8, 323–361.

Michaud, R. O. (1989): “The Markowitz Optimization Enigma: Is ’Optimized’ Optimal?” *Financial Analysts Journal*, 45, 31–42.

Mitton, T. and K. Vorkink (2007): “Equilibrium Underdiversification and the Preference for Skewness,” *Review of Financial Studies*, 20, 1255–1288.

Muravyev, D. (2016): “Order Flow and Expected Option Returns,” *Journal of Finance*, 71, 673–708.

Muravyev, D., N. D. Pearson, and J. M. Pollet (2016): “Is There a Risk Premium in the Stock Lending Market? Evidence from Equity Options,” SSRN Working Paper.

Nagel, S. (2012): “Evaporating Liquidity,” *Review of Financial Studies*, 25, 2005–2039.

Neumann, M. and C. Schlag (1996): “Martingale Restrictions and Implied Distributions for German Stock Index Option Prices,” Diskussionspapier Nr. 196, Universität Karlsruhe.

Novy-Marx, R. and M. Velikov (2016): “A Taxonomy of Anomalies and Their Trading Costs,” *Review of Financial Studies*, 29, 104–147.

Ofek, E., M. Richardson, and R. F. Whitelaw (2004): “Limited Arbitrage and Short Sales Restrictions: Evidence from the Options Markets,” *Journal of Financial Economics*, 74, 305–342.

Pasquariello, P. (2014): “Financial market dislocations,” *Review of Financial Studies*, 27, 1868–1914.

Rehman, Z. and G. Vilkov (2012): “Risk-Neutral Skewness: Return Predictability and Its Sources,” SSRN Working Paper.

Ruf, T. (2012): “Limits to Market Making and the Skewness Risk Premium in Options Markets,” SSRN Working Paper.

Santa-Clara, P. and A. Saretto (2009): “Option Strategies: Good Deals and Margin Calls,” *Journal of Financial Markets*, 12, 391–417.

Stambaugh, R. F., J. Yu, and Y. Yuan (2015): “Arbitrage Asymmetry and the Idiosyncratic Volatility Puzzle,” *Journal of Finance*, 70, 1903–1948.

Stambaugh, R. F. and Y. Yuan (2017): “Mispricing Factors,” *Review of Financial Studies*, 30, 1270–1315.

Stilger, P. S., A. Kostakis, and S.-H. Poon (2017): “What Does Risk-Neutral Skewness Tell Us About Future Stock Returns?” *Management Science*, 63, 1814–1834.

Strong, N. and X. Xu (1999): "Do S&P 500 Index Options Violate the Martingale Restriction?" *Journal of Futures Markets*, 19, 499–521.

Xing, Y., X. Zhang, and R. Zhao (2010): "What Does the Individual Option Volatility Smirk Tell Us About Future Equity Returns?" *Journal of Financial and Quantitative Analysis*, 45, 641–662.