Idiosyncratic Volatility and the Intertemporal Capital Asset Pricing Model*

Gang Li†

January 28, 2020

Abstract

When the true asset pricing model cannot be identified, the idiosyncratic volatility obtained from a misspecified model contains information of the hedge portfolio in Merton's (1973) ICAPM. Empirically, I find that from 1815 to 2018, more than two centuries, neither equal-weighted idiosyncratic volatility (EWIV) nor value-weighted idiosyncratic volatility (VWIV) can forecast stock market returns. However, EWIV and VWIV when applied together are strong predictors of stock market returns over short- and long-term horizons. The explanatory power is economically significant with an out-of-sample forecasting R^2 around 1% for one month and 12% for one year. This finding suggests that EWIV and VWIV together are linked to state variables that capture time-varying investment opportunities. Furthermore, EWIV and VWIV jointly can explain the cross-section of average stock returns with a beta quintile spread of 7.88% per year. I argue that the combination of EWIV and VWIV is a proxy for the conditional covariance risk in the ICAPM. I revisit the debate between Goyal and Santa-Clara (2003) and Bali et al. (2005) and reconcile their mixed findings between the idiosyncratic volatility and future stock market returns. Finally, this paper also gives new insights for the tail risk measure proposed by Kelly and Jiang (2014).

JEL Classification: G12, G13, G14, G17

Keywords: idiosyncratic volatility, stock market variance, conditional covariance, time-series stock return predictability, cross-section of stock returns, expected stock returns, intertemporal capital asset pricing model, economic state variable, risk-return tradeoff

---

*I am grateful to my thesis committee: Peter Christoffersen, Bing Han, Raymond Kan, and Chayawat Ornthanalai for many valuable discussions. I appreciate valuable comments from Susan Christoffersen, Alexandre Corhay, David Goldreich, Andrey Golubov, Yoontae Jeon, Charles Martineau, Tom McCurdy, Marco Salerno, Mike Simutin, Jincheng Tong, Karen Williams, as well as seminar participants at the University of Toronto. All errors are my own. Comments are welcome.

†Rotman School of Management, University of Toronto, `utgang.li@rotman.utoronto.ca`

1 Introduction

Goyal and Santa-Clara (2003) found that equal-weighted idiosyncratic volatility (EWIV) can positively forecast future stock market returns, suggesting that idiosyncratic risk matters for asset pricing. The finding, however, is criticized by several studies. For example, Bali et al. (2005) and Wei and Zhang (2005) argue that this empirical finding is attributed to liquidity premium, small-stock bias, and sample-specific period. That is, when looking at a more recent period, EWIV is unable to predict aggregate stock returns. However, the value-weighted idiosyncratic volatility (VWIV) is negatively related to future stock market returns, although the significance is not as strong as other existing predictors.[1] The difference in forecasting performance between EWIV and VWIV is confusing and ambiguous in the literature. To date, existing studies do not find consistent return predictability by aggregate idiosyncratic volatility. As a result, how aggregate idiosyncratic volatility affects asset prices remains unclear, especially for the time-series return predictability. I revisit this topic and provide an alternative explanation as to why existing studies obtain mixed conclusions. I show that aggregate idiosyncratic volatility, in fact, can be used to forecast stock market returns at both high and low frequencies.

Theoretically, if the CAPM holds, idiosyncratic volatility is related to neither time series nor cross-section of stock returns. Levy (1978), Campbell et al. (2001), and Malkiel and Xu (2002) claim that if investors hold undiversified portfolios, the expected return of any asset will be positively linked to idiosyncratic volatility. Their models imply that when investors' portfolios are less diversified, aggregate idiosyncratic volatility should be positively related to future stock market returns in time series. However, recent empirical evidence has not consistently supported this prediction. While most of the previous studies focus on explaining

[1]For example, the relevant papers include Guo and Savickas (2008), Chen and Petkova (2012), and Pollet and Wilson (2010). However, no papers document that VWIV has return predictive power longer than one quarter.

the cross-sectional effect of idiosyncratic volatility with much success,[2] the time-series effect of aggregate idiosyncratic volatility is still unclear and less explored. My research aims to fill this gap.

In this paper, I first document a novel finding that when running a bivariate regression of aggregate stock returns on EWIV and VWIV, both variables express strong and significant forecasting capacities for stock market returns. The explanatory power is economically significant with an out-of-sample R^2 around 1% for one month and 12% for one year. EWIV is positively related to future stock market returns (consistent with Goyal and Santa-Clara (2003)), while the coefficient of VWIV remains negative (consistent with Guo and Savickas (2008) and Pollet and Wilson (2010)). The empirical findings are robust for various concerns of the predictive power of aggregate idiosyncratic volatility in the literature such as small-stock effect, sample-specific periods, market liquidity premium, business-cycle conditions, and predictor persistency.[3] The significance of both coefficients cannot be explained by multicollinearity because, otherwise both coefficients would be insignificant (smaller t-statistic) with a high R^2 (larger F-statistic) regression. Statistically, the combined effect is more consistent with the omitted variable issue, which leads to estimation biases of univariate OLS regression.

Theoretically, the combined effect of EWIV and VWIV can be understood through Merton's (1973) ICAPM. In his analysis, when the investor dynamically decides her optimal portfolio position, and if there exist certain state variables which can forecast future investment opportunities, the expected return of any asset (including the market portfolio) is determined by both its covariance with the market portfolio and with the innovation of the state variables (Hedegaard and Hodrick (2014)). While the market portfolio is easy to observe, the state variables are difficult to detect and thus to value (Rossi and Timmer-

[2]The typical papers include: Ang et al. (2006), Brandt et al. (2010), Campbell and Taksler (2003), Fu (2009), Chen and Petkova (2012), Cao and Han (2016), Herskovic et al. (2016).

[3]More information can be found in Bali et al. (2005), Wei and Zhang (2007), Amihud,Hurvich, and Wang (2008), and Bekaert, Hodrick, and Zhang (2012).

mann (2015)). Given that not all state variables are identifiable and that econometricians use a misspecified asset pricing model, the residual terms will consist of multiple systematic components, some of which are driven by the missed state variables and thus related to equity risk premium. I show in theory that although a single aggregate idiosyncratic variance could not capture the covariance risk in the ICAPM, a combination between two types of weighted average idiosyncratic variance is able to pin down the conditional covariance risk, and therefore is linked to stock market returns. Econometrically, the choice of using EWIV and VWIV is also close to the optimal combination to extract the conditional covariance risk among individual idiosyncratic variance.[4]

The return predictability based on multiple predictors is not uncommon in the literature. Lettau and Ludvigson (2001a) use the combination of consumption in excess of labor income and asset wealth (CAY) to predict stock returns. Baker, Greenwood, and Wurgler (2003) study a two-predictor model of excess bond returns as a function of the lagged short- and long-term debt new issues. Menzly, Santos, and Veronesi (2004) propose a general equilibrium model to prove that the predictive regressions for returns should include both dividend yield and consumption-to-price ratio to disentangle the effect that changes in risk preference and expected dividend growth have on prices and returns. Santos and Veronesi (2006) develop a model where stock returns are predicted by labor income to consumption and by dividend yield. Ang and Bekaert (2007) show that the predictive ability of dividend yield is best seen in a bivariate regression with short rates at short horizons. Guo and Savickas (2008) run a bivariate regression of stock market returns on both market and idiosyncratic volatility. Using a latent variable approach within a present-value model, Van Binsbergen and Koijen (2010) estimate the expected stock market return as a function of both lagged dividend-price ratio and lagged dividend growth rate.[5]

[4]The detailed explanation is provided in Section 4.1.

[5]More advanced techniques are also applied to deal with omitted and latent variable problems in the literature, such as Kelly and Pruitt (2013), Kelly and Pruitt (2015), Light, Maslov, and Rytchkov (2017) and Giglio and Xiu (2018).

Guo and Savickas (2008) found that when combined with stock market volatility (SMV), value-weighted idiosyncratic volatility (VWIV) is negatively related to aggregate stock returns, although SMV and EWIV jointly do not work well.[6] My explanation is different from theirs. While they argue VWIV is a proxy for the volatility of the missing factor, I contend that VWIV is a proxy for a combination of the variance of the hedge portfolio and the covariance of the hedge portfolio with the market portfolio. Because both terms are time-varying, we need to construct another variable with the same components to figure out the conditional covariance risk. Theoretically, both EWIV and VWIV are not redundant when SMV is included in the model. Empirically, the predictive power of both EWIV and VWIV is robust when SMV is controlled in the regression.

In addition to time-series stock return predictability, I also find cross-sectional evidence to support the hypothesis that the combined effect of EWIV and VWIV is consistent with Merton's ICAPM. The analysis in Section 2 conjectures that under certain conditions, the combination of EWIV and VWIV can help pinpoint conditional risk exposures to the hedge portfolio, without identifying state variables. Consistent with the theoretical prediction, by running the time-series regression of individual stock returns on both EWIV and VWIV and sorting firms into quintile based on the firm loadings, I find that firms in the highest quintile earn average returns 7.88% per year higher than those in the lowest quintile. The cross-sectional evidence is robust to various econometric tests and unable to be explained by the Fama-French five-factor model and other relevant anomalies.

My study contributes to estimating the conditional covariance risk as well. A large literature has proposed different ways to estimate the conditional market variance,[7] but far less work has been undertaken on estimating the conditional covariance in the ICAPM. A key hurdle in measuring the conditional covariance is the difficulty in identifying the hedge

[6] See Table 4 in Guo and Savicaks (2008).

[7] For example, French, Schwert, and Stambaugh (1987), Bali and Peng (2006), Ghysels, Santa-Clara, and Valkanov (2005), Andersen et al. (2003), Glosten, Jagannathan, and Runkle (1993), Nelson (1991), Engle and Ng (1993), Hansen and Richard (1987), Harvey (2001), Ludvigson and Ng (2007).

portfolio. Scruggs (1998) uses a two-factor GARCH-in-mean model to estimate the covariance risk where the nominal risk-free rate drives movements in the conditional covariance. Guo and Whitelaw (2006) assume that the conditional covariance is a linear function of a vector of observable state variables. Lo and Wang (2006) creatively pinpoint the hedge portfolio using weekly volume and return data. Bali and Engle (2010) extend the estimation to a model with dynamic conditional correlations. Rossi and Timmermann (2015) propose a method for constructing the conditional covariance risk using a daily summary measure of economic activity to track time-varying investment opportunities. Differing from theirs, my study provides an alternative approach to estimate the conditional covariance risk through the combination of two types of aggregate idiosyncratic variance. Compared with other proposals, my approach has several advantages: it is model-free and the estimation is feasible at high frequency such as daily and weekly.

This paper also contributes to how to measure tail risk. Motivated by the power-law distribution, Kelly and Jiang (2014) propose a measure of time-varying tail risk through the cross-section of stock returns. They find that the proposed tail index has strong predictive power for aggregate stock returns and can also explain the cross-sectional variation of stock returns. Chapman, Gallmeyer, and Martin (2018) raise some empirical concerns regarding the explanation and robustness of this variable. One of their main arguments is that the tail index proposed by Kelly and Jiang (2014) explains the cross-section of the discount rate component rather than the cash-flow component, thus contradicting most of the theoretical motivations for tail risk such as the modified long-run risk model and the disaster risk model.[8] In contrast, Chapman, Gallmeyer, and Martin (2018) found that the tail index is closely related to the level and slope of the term structure. An extension of my theoretical analysis in Section 2 provides an alternative explanation of the debate regarding the tail risk measure. I conjecture that the asset pricing implications of the variable proposed by

[8] Most of those models are motivated through the cash-flow channel such as Barro (2006), Gabaix (2012), Wachter (2013), Bansal and Yaron (2004), Bansal and Shaliastovich (2011), and Drechsler and Yaron (2011).

Kelly and Jiang (2014) can also be understood under the framework of Merton's ICAPM.[9] The alternative hypothesis can justify both the time-series and cross-sectional implications of the tail risk variable, and at the same time justify why Chapman, Gallmeyer, and Martin (2018) find that the predictive power of the tail risk measure comes through the discount rate channel rather than the cash-flow channel.

Last but not least, my paper contributes to the time-series return prediction literature as well. Recent papers find more evidence that stock market returns are predictable in both short- and long-term horizons, although most of the predictors either lose the significance in recent years (Goyal and Welch (2008)) or are only available for short-sample periods.[10] By collecting data from both the Center for Research in Security Prices (CRSP) and the Global Financial Data (GFD), I provide new and strong empirical evidence that stock market returns are predictable with consistent performance and sufficient long-sample periods from 1815 to 2018, over two centuries. The return prediction by the combination of EWIV and VWIV exists at daily and weekly frequency as well. Previous studies find mixed evidence regarding stock return predictability before 1925. Goetzmann, Ibbotson, and Peng (2001) estimate an alternative stock index for the New York Stock Exchange between 1815 and 1925 and find little evidence of return prediction. Chen (2009) studies the U.S. stock market return prediction by dividend growth between 1872 and 2005, and documents that returns are largely unpredictable before 1926. Using comprehensive databases, Goleza and Koudijs (2018) found strong evidence of return prediction by dividend yield for the past four centuries. Different from the previous studies focus on dividend yield, my paper examines the conditional covariance risk, which complements their studies and provides strong support for the hypothesis of time-varying risk premium. The variation of equity risk premium is consistent with Merton's ICAPM that investors are seeking to hedge against the shortfalls in consumption or against changes in the future investment opportunity sets.

[9]The detailed investigation is shown in Section 5.
[10]For example, most of the option predictors are only available since 1990.

The remainder of the paper is organized as follows. Section 2 derives the theoretical framework of the relationship between the combination of EWIV and VWIV and the conditional covariance risk in the ICAPM. Section 3 provides both in-sample and out-of-sample time-series empirical evidence. Section 4 conducts various robustness checks, revisits previous studies in the literature, and discusses asset pricing implications at cross-section. Section 5 examines the relationship between the conditional covariance risk and the tail risk proposed by Kelly and Jiang (2014). Section 6 provides a simulation study based on the theoretical framework in Section 2. Section 7 concludes the paper.

2 Theoretical Motivation

The derivation begins with the discrete-time version of Merton's ICAPM. By analyzing the dynamic optimal portfolio choice of the representative agent, Merton (1973) suggests the expression of the conditional expected return on asset i as:

$$E_t(R_{i,t+1}) = \gamma_M Cov_t(R_{i,t+1}, R_{M,t+1}) + \sum_{k=1}^{K} \gamma_k Cov_t(R_{i,t+1}, \Delta Z_{k,t+1}), \quad i = 1, ..., N, \tag{1}$$

where

$R_{i,t+1}$ is the excess return (rates of return minus a risk-free rate) for asset i;

$R_{M,t+1} = \sum_{i=1}^{N} w_{i,t} R_{i,t+1}$ is the market excess return, where $w_{i,t}$ is the market portfolio's weight for asset i;

$Z_{k,t+1}$ stands for certain state variable k, which contains information about future investment opportunities,

γ_M is the relative risk aversion of the representative agent, assumed to be constant;

γ_k is the weighted average across investors of their state-variable aversions, assumed to be constant as well;

E_t is the conditional expectation operator on information at time t;

Cov_t is the conditional covariance operator on information at time t.

Equation (1) has been deduced in many papers such as Merton (1973), Long Jr (1974), and Cox, Ingersoll Jr, and Ross (1985). The ICAPM predicts that the expected return on asset i is determined by not only its covariance with the market portfolio but also its covariance with the innovation of certain state variables, which are linked to future investment opportunities (e.g., aggregate stock returns or volatilities). Equation (1) can be used to derive the conditional equity risk premium:

$$E_t(R_{M,t+1}) = \gamma_M Var_t(R_{M,t+1}) + \sum_{k=1}^{K} \gamma_k Cov_t(R_{M,t+1}, \Delta Z_{k,t+1}) \tag{2}$$

where Var_t stands for the conditional variance operator on information at time t. For simplicity, I will use μ to denote the expected return and σ^2 to denote the (co)variance term for the rest of the derivations. $\mu_{i,t}$, $\mu_{M,t}$, $\mu_{H,t}$ ($\sigma_{i,t}$, $\sigma_{M,t}$, $\sigma_{H,t}$) are the conditional expected excess return (volatility) of asset i, the market portfolio, and the hedge portfolio at time t. $\sigma_{iM,t}$, $\sigma_{iH,t}$, $\sigma_{MH,t}$ stand for the covarianec between any two of them.

Aside from the traditional risk-return tradeoff based on the CAPM, the ICAPM states that the conditional equity risk premium is determined not only by the conditional variance of the market portfolio but also by the conditional covariance of the market portfolio with the innovation of the state variables. Although $\sigma^2_{M,t}$ and $\sigma_{M\Delta Z_k,t}$ are time-varying, since both variables are very persistent and stationary over time, one would expect empirically that either the variance term or the covariance terms can predict future aggregate realized stock returns in time series. Equation (2) can be simplified to a two-factor model by creating a new giant covariance term including all state variables:

$$\mu_{M,t} = \gamma_M \sigma^2_{M,t} + Cov_t(R_{M,t+1}, \sum_{k=1}^{K} \gamma_k \Delta Z_{k,t+1}). \tag{3}$$

Without loss of generality, in the following analysis, I consider the case of $K = 1$. In practice, since the state variable (Z_K) is not tradeable, people often project the innovation of the state variable to a factor mimicking portfolio, namely a hedge portfolio, which is correlated with future investment opportunities. Based on the two traded portfolios, the ICAPM can be expressed as:

$$\mu_{i,t} = \gamma_M \sigma_{iM,t} + \gamma_H \sigma_{iH,t}. \tag{4}$$

Note that the hedge portfolio has the maximum absolute correlation with the state variables, but not perfectly correlated with them. Nevertheless, previous studies show that $\sigma_{i\Delta Z,t}$ and $\sigma_{iH,t}$ are linearly correlated (Ingersoll (1987), p.218), thus not impacting my main conclusion. Based on equation (4), the corresponding conditional equity risk premium can be written as:

$$\mu_{M,t} = \gamma_M \sigma_{M,t}^2 + \gamma_H \sigma_{MH,t}. \tag{5}$$

When multiple state variables are involved in Merton's ICAPM, $\sigma_{MH,t}$ can be treated as a giant hedge portfolio which hedges the combination of all state variables ($\sigma_{MH,t} \propto Cov_t(R_{M,t+1}, \sum_{k=1}^{K} \gamma_k \Delta Z_{k,t+1})$). Equation (4) can be expressed as the beta representation:

$$\mu_{i,t} = \beta_{iM,t} \mu_{M,t} + \beta_{iH,t} \mu_{H,t}. \tag{6}$$

Assuming rational expectations, the realized return generating process of (6) is given by:

$$R_{i,t+1} = \beta_{iM,t}(\mu_{M,t} + \varepsilon_{M,t+1}) + \beta_{iH,t}(\mu_{H,t} + \varepsilon_{H,t+1}) + \varepsilon_{i,t+1}, \tag{7}$$

where $\varepsilon_{M,t+1}$, $\varepsilon_{H,t+1}$, and $\varepsilon_{i,t+1}$ are exogenous shocks for the market portfolio, the hedge portfolio, and asset i. The cross-sectional distribution of the idiosyncratic shock is assumed to be normal and stable over time (Ferson, Kandel, and Stambaugh (1987)). A sufficient condition for the assumption is that the joint distribution of the returns and the relevant information set is a stationary normal distribution. It is worth noting that the risk exposure of asset i to the market portfolio is affected by the extra covariance term in the ICAPM. Based on Merton and Samuelson (1990, P.390), the risk exposures under the ICAPM are: $\beta_{iM,t} = \frac{\sigma_{iM,t}\sigma^2_{H,t} - \sigma_{iH,t}\sigma_{MH,t}}{\sigma^2_{H,t}\sigma^2_{M,t} - \sigma^2_{MH,t}}$, $\beta_{iH,t} = \frac{\sigma_{iH,t}\sigma^2_{M,t} - \sigma_{iM,t}\sigma_{MH,t}}{\sigma^2_{H,t}\sigma^2_{M,t} - \sigma^2_{MH,t}}$, where the market portfolio and the hedge portfolio are not necessarily orthogonal to each other. In the case of $\sigma_{MH,t} = 0$, both the equity risk premium and expected stock return will not be affected by the covariance term. In the following derivations, I therefore assume that $\sigma_{MH,t} \neq 0$. Note that since $\sigma_{MH,t}$, $\sigma^2_{H,t}$, and $\sigma^2_{M,t}$ are time-varing, the risk exposures change over time as well, although previous studies find that the risk exposures are relatively stable and that most of the predictable variation in returns are attributed to predictable variation in risk premia instead of time-varying risk exposures (Ferson and Harvey (1991); Evans (1994)). Their empirical studies are also important motivations to decompose idiosyncratic variance into risk exposures (i.e., persistent parts) and other components (i.e., volatile parts).

While the market portfolio can easily be observed, the hedge portfolio is difficult to estimate. Assume that econometricians only use the CAPM as the asset pricing model,

$$R_{i,t+1} = b_{iM,t}(\mu_{M,t} + \varepsilon_{M,t+1}) + \eta_{i,t+1}, \tag{8}$$

where $b_{iM,t} = \frac{\sigma_{iM,t}}{\sigma^2_{M,t}}$ is the misspecified risk exposures; $\eta_{i,t+1}$ is the misspecified idiosyncratic shock for asset i under the CAPM. I first show that both the first and the second conditional moments of the misspecified firm idiosyncratic shock can be expressed as a combination of conditional variance of the hedge portfolio and the conditional covariance between the market portfolio and the hedge portfolio.

Proposition 1. *Suppose that the true conditional asset pricing model follows Merton's ICAPM defined in (7), but econometricians use the conditional CAPM as the asset pricing model defined in (8). The first and second conditional moment of the misspecified firm* *idiosyncratic shock* $\eta_{i,t+1}$ *are given by:*

$$E_t(\eta_{i,t+1}) = \gamma_H b_{iH,t}\sigma^2_{H,t} + \gamma_H b_{iM,t}\sigma_{MH,t}, \tag{9}$$

$$Var_t(\eta_{i,t+1}) = \beta_{iH,t}b_{iH,t}\sigma^2_{H,t} + \beta_{iH,t}b_{iM,t}\sigma_{MH,t} + \sigma^2_{\varepsilon_i,t}, \tag{10}$$

$$\textit{where} \begin{cases} \beta_{iM,t} = \frac{\sigma_{iM,t}\sigma^2_{H,t} - \sigma_{iH,t}\sigma_{MH,t}}{\sigma^2_{H,t}\sigma^2_{M,t} - \sigma^2_{MH,t}} \\ \beta_{iH,t} = \frac{\sigma_{iH,t}\sigma^2_{M,t} - \sigma_{iM,t}\sigma_{MH,t}}{\sigma^2_{H,t}\sigma^2_{M,t} - \sigma^2_{MH,t}} \end{cases}, \begin{cases} b_{iM,t} = \frac{\sigma_{iM,t}}{\sigma^2_{M,t}} \\ b_{iH,t} = \frac{\sigma_{iH,t}}{\sigma^2_{H,t}} \end{cases}.$$

Proofs are provided in the Appendix. Proposition 1 shows that the misspecified idiosyncratic variance contains at least two common components, but only one of them (i.e., $\sigma_{MH,t}$) is related to the conditional equity risk premium. Corollary 1.1 below proves how two types of different weighted-average idiosyncratic variance can capture the conditional covariance in Merton's ICAPM, thus are linked to stock market returns. In my study, I use equal-weighted average idiosyncratic variance (EWIV) and value-weighted average idiosyncratic variance (VWIV) as an example. Theoretically, the weighting schemes are not necessarily restricted to this choice. The robustness checks of using alternative weighting schemes are provided in Section 4, where I also show econometrically that EWIV and VWIV are close to an ex-post optimal combination to estimate the conditional covariance.

Corollary 1.1. *Suppose that the true conditional asset pricing model follows Merton's ICAPM defined in (7), but econometricians use the conditional CAPM as the asset pricing model defined in (8). The conditional variance of the hedge portfolio and the conditional covariance between the market portfolio and the hedge portfolio are given by*

$$\sigma^2_{H,t} = \frac{B^{EW}_t}{A^{VW}_t B^{EW}_t - A^{EW}_t B^{VW}_t}\widetilde{VWIV}_t - \frac{B^{VW}_t}{A^{VW}_t B^{EW}_t - A^{EW}_t B^{VW}_t}\widetilde{EWIV}_t, \tag{11}$$

$$\sigma_{MH,t} = \frac{A^{VW}_t}{A^{VW}_t B^{EW}_t - A^{EW}_t B^{VW}_t}\widetilde{EWIV}_t - \frac{A^{EW}_t}{A^{VW}_t B^{EW}_t - A^{EW}_t B^{VW}_t}\widetilde{VWIV}_t, \tag{12}$$

$$
where \begin{cases} A_t^{EW} = \sum_{i=1}^{N_t} \frac{1}{N_t}\beta_{iH,t}b_{iH,t} \\ A_t^{VW} = \sum_{i=1}^{N_t} w_{i,t}\beta_{iH,t}b_{iH,t} \end{cases}, \begin{cases} B_t^{EW} = \sum_{i=1}^{N_t} \frac{1}{N_t}\beta_{iH,t}b_{iM,t} \\ B_t^{VW} = \sum_{i=1}^{N_t} w_{i,t}\beta_{iH,t}b_{iM,t} \end{cases}, \begin{cases} EWIV_t = \sum_{i=1}^{N_t} \frac{1}{N_t}Var_t(\eta_{i,t+1}) \\ VWIV_t = \sum_{i=1}^{N_t} w_{i,t}Var_t(\eta_{i,t+1}) \end{cases}
$$

$$
\begin{cases} \Omega_t^{EW} = \sum_{i=1}^{N_t} \frac{1}{N_t}\sigma^2_{\varepsilon_i,t} \\ \Omega_t^{VW} = \sum_{i=1}^{N_t} w_{i,t}\sigma^2_{\varepsilon_i,t} \end{cases}, \begin{cases} \widetilde{EWIV}_t = EWIV_t - \Omega_t^{EW} \\ \widetilde{VWIV}_t = VWIV_t - \Omega_t^{VW} \end{cases}.
$$

Proofs and corresponding assumptions are provided in the Appendix. Since the weighting schemes are not perfectly linearly correlated (i.e., $A_t^{VW}B_t^{EW} - A_t^{EW}B_t^{VW} \neq 0$). The solution exists and is unique. Both the conditional variance of the hedge portfolio and the conditional covariance of the market portfolio with the hedge portfolio are a linear combination of EWIV and VWIV. Empirically, the second conditional moment of $\eta_{i,t+1}$ can be estimated by the standard deviation of the realized idiosyncratic shocks from the CAMP using the daily returns within each month:

$$
\widehat{Var_t}(\eta_{i,t+1}) = \frac{1}{21}\sum_{s=0}^{21}(R_{i,t-s} - \hat{b}_{iM,t}R_{M,t-s})^2. \tag{13}
$$

Because both $\sigma^2_{H,t}$ and $\sigma_{MH,t}$ are time-varying and because only $\sigma_{MH,t}$ matters for the conditional equity risk premium, a single aggregate idiosyncratic variance, either EWIV or VWIV, is a noisy proxy for $\sigma_{MH,t}$, and therefore is inconsistent in its performance forecasting future stock market returns. This is why previous research obtains mixed empirical findings of the predictive power of aggregate idiosyncratic variance. In Section 6, I conduct a simulation study to illustrate the estimation bias of the omitted-variable issue.

Based on Proposition 1, besides the conditional misspecified idiosyncratic variance, the first conditional moment of the misspecified idiosyncratic shock can also be used to estimate $\sigma_{MH,t}$. Therefore, it is feasible to estimate the common component using the cross-section of individual residual returns, instead of idiosyncratic variance.

Corollary 1.2. *Suppose that the true conditional asset pricing model follows Merton's*

ICAPM defined in (7), but econometricians use the conditional CAPM as the asset pricing model defined in (8). The conditional variance of the hedge portfolio and the conditional covariance between the market portfolio and the hedge portfolio are given by

$$\sigma^2_{H,t} = \frac{G_t^{EW}}{F_t^{VW}G_t^{EW} - F_t^{EW}G_t^{VW}} VWAP_t - \frac{G_t^{VW}}{F_t^{VW}G_t^{EW} - F_t^{EW}G_t^{VW}} EWAP_t, \tag{14}$$

$$\sigma_{MH,t} = \frac{F_t^{VW}}{F_t^{VW}G_t^{EW} - F_t^{EW}G_t^{VW}} EWAP_t - \frac{F_t^{EW}}{F_t^{VW}G_t^{EW} - F_t^{EW}G_t^{VW}} VWAP_t, \tag{15}$$

where $\begin{cases} EWAP_t = \sum_{i=1}^{N_t} \frac{1}{N_t} E_t(\eta_{i,t+1}) \\ VWAP_t = \sum_{i=1}^{N_t} w_{i,t} E_t(\eta_{i,t+1}) \end{cases}, \begin{cases} F_t^{EW} = \sum_{i=1}^{N_t} \frac{1}{N_t} \gamma_H b_{iH,t} \\ F_t^{VW} = \sum_{i=1}^{N_t} w_{i,t} \gamma_H b_{iH,t} \end{cases}, \begin{cases} G_t^{EW} = \sum_{i=1}^{N_t} \frac{1}{N_t} \gamma_H b_{iM,t} \\ G_t^{VW} = \sum_{i=1}^{N_t} w_{i,t} \gamma_H b_{iM,t} \end{cases}.$

Proofs and corresponding assumptions are provided in the Appendix. The result is similar to Corollary 1.1, except that I use the first conditional moment to estimate the conditional covariance. Empirically, the first conditional moment of the misspecified idiosyncratic shock is similar to the intercept (α) from the Market model using daily returns within each month:

$$\widehat{E_t}(\eta_{i,t+1}) = \hat{\alpha}_{i,t} = \frac{1}{21}\sum_{s=0}^{21}(R_{i,t-s} - \hat{b}_{iM,t}R_{M,t-s}). \tag{16}$$

While the focus of my paper is idiosyncratic variance, I will discuss the implication of Corollary 1.2 in Section 5, when examing the tail risk measure. Thereafter, I will use EWIV and VWIV as the primary setup for the derivations in this section. The conclusion is similar if I replace them with the first-moment case.

Proposition 2. *Suppose that the true conditional asset pricing model follows Merton's ICAPM defined in (7), but econometricians use the conditional CAPM as the asset pricing model defined in (8). The conditional equity risk premium can be expressed as:*

$$\mu_{M,t} = \gamma_M \sigma^2_{M,t} + \gamma_H \sigma_{MH,t} = \gamma_M \times \sigma^2_{M,t} + C_t \times \widetilde{EWIV}_t - D_t \times \widetilde{VWIV}_t, \tag{17}$$

where $C_t = \frac{\gamma_H A_t^{VW}}{A_t^{VW} B_t^{EW} - A_t^{EW} B_t^{VW}}$, $D_t = \frac{\gamma_H A_t^{EW}}{A_t^{VW} B_t^{EW} - A_t^{EW} B_t^{VW}}$.

Similarly, the portfolio/firm conditional expected excess return is given by:

$$\mu_{i,t} = \beta_{iM,t}\mu_{M,t} + C_{i,t} \times \widetilde{EWIV}_t - D_{i,t} \times \widetilde{VWIV}_t, \tag{18}$$

where $C_{i,t} = \beta_{iH,t}\frac{\gamma_M A_t^{VW} - \gamma_H B_t^{VW}}{A_t^{VW} B_t^{EW} - A_t^{EW} B_t^{VW}}$, $D_{i,t} = \beta_{iH,t}\frac{\gamma_M A_t^{EW} - \gamma_H B_t^{EW}}{A_t^{VW} B_t^{EW} - A_t^{EW} B_t^{VW}}$.

Proofs and corresponding assumptions are given in the Appendix. Note that $\sigma_{MH,t}$ is replaced by the combination of $\widetilde{EWIV}$ and $\widetilde{VWIV}$, since the latter are identifiable. As to the coefficients, if A_t^{EW} and A_t^{VW} have the same sign, the coefficients of EWIV and VWIV are supposed to be opposite each other.

It is worth noting that Proposition 2 has important implications on estimating risk exposures of the missing factors for the conditional ICAPM. Since the hedge portfolio is generally unobservable, how to estimate the risk premium of the hedge portfolio and its corresponding risk exposure is challenging. Previous studies focus on identifying state variables and estimate the risk exposures based on the covariance term with the innovation of certain state variables. Instead of looking for state variables, Proposition 2 shows that one can estimate conditional risk exposures of the hedge portfolio based on the conditional misspecified idiosyncratic variance.

Corollary 2.1. *Suppose that the true conditional asset pricing model follows Merton's ICAPM defined in (7), but econometricians use the conditional CAPM as the asset pricing model defined in (8). The conditional risk exposure of firm/portfolio i to the missed* *hedge portfolio ($\beta_{iH,t}$) is given by one of the following formulas:*

$$\begin{cases} \beta_{iH,t} = C_{i,t} \times \frac{A_t^{VW} B_t^{EW} - A_t^{EW} B_t^{VW}}{\gamma_M A_t^{VW} - \gamma_H B_t^{VW}} \propto C_{i,t} \text{ at cross-section} \\ \beta_{iH,t} = D_{i,t} \times \frac{A_t^{VW} B_t^{EW} - A_t^{EW} B_t^{VW}}{\gamma_M A_t^{EW} - \gamma_H B_t^{EW}} \propto D_{i,t} \text{ at cross-section} \\ \beta_{iH,t} = \frac{A_t^{VW} B_t^{EW} - A_t^{EW} B_t^{VW}}{\gamma_H^2} (D_t \times C_{i,t} - C_t \times D_{i,t}) \propto (D_t \times C_{i,t} - C_t \times D_{i,t}) \text{ at cross-section} \end{cases} \tag{19}$$

where C_t, $C_{i,t}$, D_t, and $D_{i,t}$ are based on Proposition 2.

Proofs and corresponding assumptions are given in the Appendix. Corollary 2.1 predicts that by running the regressions of (17) and (18), we are able to obtain a proxy for the risk exposure to the hedge portfolio from one of the coefficients in (19), which I will provide the empirical supports in Section 4.5. It is worth noting that since $(A_t^{VW} B_t^{EW} - A_t^{EW} B_t^{VW})$ are not identifiable, Corollary 2.1 does not give the explicit formula for $\beta_{iH,t}$. Instead, one can only infer the relative rank of $\beta_{iH,t}$ at cross-section by assuming that the sign of $(A_t^{VW} B_t^{EW} - A_t^{EW} B_t^{VW})$ does not change over time. In addition to this, it should be noted that Corollary 2.1 only applies to the conditional model. As to the unconditional case, since the time-varying risk exposure and the risk premium are correlated, there are extra factors appearing in the unconditional version (Jagannathan and Wang (1996)). Nevertheless, the empirical results in Section 4.5 strongly support Corollary 2.1 that EWIV and VWIV have important cross-sectional implications on stock returns.

3 Time-Series Empirical Evidence

3.1 Data, Variable, and Summary Statistics

Motivated by diverse goals, researchers define idiosyncratic volatilities in different ways. To make my empirical results neat and mostly comparable to other papers in the literature, when conducting the empirical tests, I use idiosyncratic volatility instead of idiosyncratic variance and calculate it based on the CAPM model.[11] The empirical results are robust if I use idiosyncratic variance or log-variance of stock returns.

From 1926 to 2018, for each stock at the end of each month, I use the past 30-day return observations to fit the CAPM model. The data of stock returns, market capitalization, share code, and exchange code are obtained from the CRSP. The stock market excess return (MKTRF) is downloaded from Kenneth French's website. The number of return observations within each month should be greater than 20. The idiosyncratic volatility of each stock is then defined as the standard deviation of the residuals from the benchmark model.[12] The value-weighted idiosyncratic volatility (VWIV) and equal-weighted idiosyncratic volatility (EWIV) are simply obtained by taking the cross-sectional average based on either market-capitalization weights at the end of the previous month or equal weights. The daily EWIV and VWIV are constructed in a similar way at the daily frequency. When constructing aggregate idiosyncratic volatility, I include all available securities traded in the U.S. stock exchanges, so that it is closer to the market portfolio.[13]

In order to construct the sample before 1926, I collect monthly observations including stock price, stock return, and market capitalization from the Global Financial Data (GFD), which provides U.S. stock data beginning from 1790. The monthly firm idiosyncratic volatility before 1926 is estimated using a 48-month rolling window of return observations to fit the CAPM. The market excess returns between 1871 to 1925 are obtained from Amit Goyal's website. From 1815 to 1870, there were no liquid monthly short-term government securities (Golez and Koudijs (2018)). Therefore, I use the raw stock market returns between 1815 and 1870. The data is obtained from G.William Schwert's website (Schwert (1990)). The value-weighted idiosyncratic volatility (VWIV) and equal-weighted idiosyncratic volatility

[11]In this paper, the "volatility" represents the standard deviation of stock returns.

[12]The empirical results are robust if I use other types of definitions such as total volatility proposed by French et al (1987), idiosyncratic volatility based on different benchmarks (e.g., Fama-French three/five-factor model) and frequencies (e.g., daily or monthly), and total variance used in Chen and Petkova (2012).

[13]The empirical results are robust if I only include the common stocks.

(EWIV) are simply derived by taking the cross-sectional average based on either market-cap weights at the end of the previous month or equal weights. When constructing aggregate idiosyncratic volatility, I include all available securities traded in the U.S. stock exchanges with a stock price greater than $1.

The followings are the figures of monthly and daily EWIV and VWIV, and the corresponding summary statistics. As a comparison, I also estimate the stock market volatility (SMV) based on the past 30-day daily return observations of MKTRF.

Insert Figure 1 and Figure 2

Since the data of CRSP before 1963 is subject to quality issues (Shumway (1997)), the main sample period is from 1963 to 2018. The empirical results with longer periods from 1815 to 1962 are provided as robustness checks.

Insert Table 1

Figures 1 and 2 show that EWIV is always higher than VWIV, which is consistent with previous findings in the literature that small firms' volatilities are on average higher than big firms' volatilities. In the meantime, SMV shares a lower correlation with either EWIV or VWIV. All the volatility measures shoot up during the recession periods and revert back to the normal level thereafter. As expected, EWIV and VWIV are highly correlated with each other (correlation equals to 0.893). However, one can see that EWIV and VWIV also perform differently sometimes. For example, EWIV shoots up much higher than VWIV during recessions. The divergence between EWIV and VWIV supports the assumption that the variance and covariance of the hedge portfolio are time-varying.

One important issue for aggregate idiosyncratic volatility is whether there exists a trend of aggregate idiosyncratic volatility. The trend of aggregate idiosyncratic volatility is first documented by Campbell et al. (2001). By extending the sample period to 2008 and to 23 developed equity markets, Bekaert, Hodrick, and Zhang (2012) find no evidence of upward trends in recent periods. I extend both studies to examine whether the trend exists for aggregate idiosyncratic volatility. I follow Bekaert, Hodrick, and Zhang (2012) and conduct the following test:

$$y_t = b_0 + b_1 t + u_t, \tag{20}$$

where y_t is the variable of interest (i.e, EWIV or VWIV), and t is a linear time trend. The null hypothesis test is $b_1 = 0$. Similar to Bekaert, Hodrick, and Zhang (2012), I find no evidence that there is a trend of either EWIV or VWIV in recent periods. The t-stat of the coefficient of b_1 is 1.05 and 0.19 for EWIV and VWIV respectively.

3.2 In-Sample Empirical Evidence

The most commonly used multi-period predictive regression follows Fama and French (1988, 1989):

$$\sum_{k=1}^{K} \frac{r_{t+k}}{K} \equiv r_{t,t+K} = a + b \times X_t + \epsilon_{t,t+K} \tag{21}$$

Similarly, the bivariate regression can be written as:

$$r_{t,t+K} = a + b_1 \times X_{1,t} + b_2 \times X_{2,t} + \epsilon_{t,t+K}, \tag{22}$$

where r_{t+k} is the value-weighted market excess return in logarithm (MKTRF) at time $t+k$ obtained from Kenneth French's Website; X_t is the predictor variable of interest; K stands for the forecast horizon (number of months ahead). I run monthly predictive regressions with K equal to 1, 3, 6, 12, and 24 months. When $K > 1$, I correct the serial correlation and conditional heteroscedasticity using the Newey-West correction with $K - 1$ lags (Newey and West (1987)). The hypothesis testing is: $H_0 : b = 0$ vs. $H_1 : b \neq 0$. When running regressions, to make the coefficients comparable, I scale all independent variables to have zero mean and one standard deviation. Three points are worth mentioning here:

(1) In Proposition 2, the true predictors are $\widetilde{EWIV}_t$ and $\widetilde{VWIV}_t$. Since the aggregate

true idiosyncratic volatility is unable to be observed, I use $EWIV_t$ and $VWIV_t$ as predictors in the regression. In Appendix B, I show that under certain conditions the effect caused by true idiosyncratic volatility is moderate.

(2) In Proposition 2, the variables on the right-hand side should be estimated based on the conditional volatility of the misspecified idiosyncratic shocks. In the empirical tests, I use the lagged volatility as a proxy for the conditional variable, which can be justified by the high persistence of the time series of aggregate idiosyncratic volatility. The empirical results are robust if I use parametric models, such as ARCH or GARCH, to construct the conditional idiosyncratic volatility.

(3) In Proposition 2, since the coefficients (e.g., C_t and D_t) are also time-varying, the predictive regression of (21) does not exactly capture the true time-varying coefficients in equation (17). The empirical results can only be interpreted as evidence that the variation of EWIV and VWIV is more important than the variation of the average risk exposures (i.e., C_t and D_t). One can follow Lettau and Ludvigson (2001b) to model C_t and D_t as a linear function of certain conditional variable Z_t, in order to capture the effects from time-varying C_t and D_t, which is beyond the scope of my study. For simplicity and tractability, C_t and D_t are assumed to be relatively stable over time.

Since there are multiple predictors in the regressions and both of them are highly correlated and persistent, the hypothesis test from OLS might be subject to estimation biases. Amihud, Hurvich, and Wang (2008) propose an alternative hypothesis-testing method for multiple-predictor regressions in small samples. Their approach, the multi-predictor augmented regression method (mARM), shows better performance to reject the null hypothesis of no predictive power than those of OLS and bootstrapping method. Therefore, I also conduct their test and provide the corresponding mARM statistics as a robustness check for the conclusion. The corresponding mARM statistics are provided in Table 2.

Insert Table 2

Table 2 provides convincing evidence that although neither EWIV nor VWIV is able to predict stock market returns, EWIV and VWIV jointly are strong predictors of aggregate stock returns. The in-sample one-month and one-year R^2 are around 1% and 14% respectively. The predictive power of the combination cannot be explained by multicollinearity, because if so one should observe that both coefficients are insignificant with a high R^2 regression. Instead, the lack of predictive power of EWIV (or VWIV) is more likely driven by the omitted-variable bias, which can be understood by the following example. Suppose that the true model is

$$R_{M,t+1} = a + \gamma_M \times \sigma^2_{M,t} + C_t \times EWIV_t - D_t \times VWIV_t + e_{M,t+1}. \tag{23}$$

When running the predictive regression, suppose that we only include EWIV:

$$R_{M,t+1} = a + b \times EWIV_t + \nu_{M,t+1}, \tag{24}$$

the estimation of b is given by:

$$b = \frac{Cov(R_{M,t+1}, EWIV_t)}{Var(EWIV_t)} = C_t + \gamma_M \times \frac{Cov(\sigma^2_{M,t}, EWIV_t)}{Var(EWIV_t)} - D_t \times \frac{Cov(VWIV_t, EWIV_t)}{Var(EWIV_t)}. \tag{25}$$

The estimation bias is driven by two components: $\sigma^2_{M,t}$ and $VWIV_t$. Note that the omitted-variable bias does not always lead to insignificant coefficient tests of existing variables unless the omitted variable is highly correlated with existing variables with opposite values. Since the correlation between $\sigma^2_{M,t}$ and $EWIV_t$ is moderate (44%) and C_t is empirically positive while γ_M is on-average positive as well, the bias caused by $\sigma^2_{M,t}$ is moderate and acceptable. On the contrary, $VWIV_t$ and $EWIV_t$ are highly correlated (89%) and the coefficient of $VWIV_t$ is very close to that of $EWIV_t$ with an opposite sign. Therefore, one would expect that missing VWIV in the regression will lead to a significant bias to the coefficient estimation for EWIV. Empirically, I show that the coefficient of EWIV is highly underestimated if VWIV is missing in the regression and vice versa. For example, in Table 2 the univariate regression, the coefficient of EWIV is around 0.001. However, the coefficient of EWIV increases to greater than 0.01 in the bivariate regression.

In Table 3, I show that the joint predictive power of EWIV and VWIV holds at high frequencies such as daily and weekly as well. When running the daily and weekly regressions in Table 3, I add the one-day (one-week) lagged stock market return as an extra independent variable to control for positive autocorrelation of stock market returns documented by Lo and MacKinlay (1988). Similar to the monthly regression, I correct the serial correlation and conditional heteroscedasticity using the Newey-West correction with $D-1$ lags, where D stands for the daily forecast horizon.

Insert Table 3

Consistent with Table 2, while neither EWIV nor VWIV is able to predict stock market returns at the daily frequency, EWIV and VWIV together create strong signals for future aggregate stock returns. Overall, the empirical finding provides persuasive evidence of stock return predictability both with sufficient long-sample periods and at various forecast horizons. Garci, Mantilla-Garcia, and Martellini (2014) suggest computing aggregate idiosyncratic volatility by calculating the cross-sectional standard deviation of individual stock returns. They show that the new predictor can forecast equal-weighted aggregate stock returns at the daily frequency, although long-term return prediction has not been found. My paper's motivation and findings are different from theirs. They still treat EWIV and VWIV separately in their tests, but I consider EWIV and VWIV together as a proxy for the covariance term in Merton's ICAPM. Empirically, the predictive power is robust if I construct EWIV and VWIV using the method proposed by Garcia, Mantilla-Garcia, and Martellini (2014).

To investigate whether the combination of EWIV and VWIV is only a proxy for any existing predictor in the literature, I run the multiple regressions with the combination of EWIV and VWIV and each of the classic predictors found by other papers. The controlling variables include: 22 predictors based on Goyal and Welch (2008) from Amit Goyal's website, stock market volatility (SMV; Guo and Savickas (2008)), variance risk premium (VRP; Bollerslev, Tauchen, and Zhou (2009)), lower-bound equity premium (SVIX; Martin (2017)), average correlation (Pollet and Wilson (2010)), asset wealth (CAY; Lettau and Ludvigson (2001a)), aggregate stock illiquidity (ILIQ; Chen, Eaton, and Paye (2018)), investor sentiment (Baker and Wurgler (2007)), aggregate short interest (SII; Rapach, Ringgenberg, and Zhou (2016)), and aggregate implied volatility spread (IVS; Han and Li (2017)). The regression specification is given by:

$$r_{t,t+K} = a + b_1 \times EWIV_t + b_2 \times VWIV_t + b_3 \times X_{i,t} + \epsilon_{t,t+K}. \tag{26}$$

Insert Table 4

Table 4 shows that after controlling for those selected predictors, EWIV and VWIV retain the significance at all forecast horizons. The coefficients of EWIV and VWIV in Table 4 are about the same as the corresponding ones in the univariate regressions in Table 2. Furthermore, when stock market volatility is included in the regression, it shows significant positive relation with stock market returns, thus supporting the positive risk-return tradeoff.

Insert Table 5

For return prediction at high frequency, because the number of such predictors is small, I include all available predictors and run one multiple regression together at the daily frequency. For example, I include VRP by Bollerslev, Tauchen, and Zhou (2009) to differentiate from the diffusion effect. Han and Li (2017) found that aggregate implied volatility spread (IVS) between at-the-money call and put equity options can strongly predict stock market returns at daily and weekly frequency. Since both variables are constructed using volatility, I also control for their variable in my test. The multiple regression is displayed in Table 5.

3.3 Out-of-Sample Performance

To alleviate concerns about overfitting and finite sample biases, I conduct out-of-sample tests by separating the data sample (1963 to 2018) into two parts: 1963 to 1980 as the in-sample estimation period and 1981 to 2018 as the out-of-sample performance evaluation period. Starting in January 1981, I run various predictive regressions each month using historical data from January 1963 and then compare the out-of-sample forecast errors (i.e., differences between the realized market returns and the predicted returns) with those from the benchmark model (i.e., historical average).[14] The statistical test of equal predictive accuracy in nested models is based on Clark and West (2007). The regression details are given by:

$$\begin{cases} r_{t,t+K} = \alpha + \beta \times x_t + \epsilon_{t,t+K}, & t = 1, ..., T_0 - K \\ \hat{r}_{t,t+K} = \hat{\alpha} + \hat{\beta} \times x_t, & t = T_0, ..., T \end{cases} \tag{27}$$

$$\text{Benchmark: } r^B_{t,t+K} = \frac{1}{t-K} \sum_{s=1}^{t-K} r_{s,s+K}, \quad t = T_0, ..., T, \tag{28}$$

where K is the forecast horizon, $r_{t,t+K}$ is the market excess return from time t to $t+K$, x_t is the value of the predictor at time t; $\hat{r}_{t,t+K}$ is the forecasted return based on x_t from the recursive regression. The out-of-sample R^2 statistic is defined as 1 minus the ratio of mean squared forecast error of the larger model to that of the benchmark model:

$$MSFE_1 = \frac{1}{T-T_0} \sum_{t=T_0}^{T} (r_{t,t+K} - \hat{r}_{t,t+K})^2; MSFE_0 = \frac{1}{T-T_0} \sum_{t=T_0}^{T} (r_{t,t+K} - r^B_{t,t+K})^2, \tag{29}$$

[14] Similar out-of-sample tests are used by Campbell and Thompson (2008), Goyal and Welch (2008), and Rapach et al. (2010).

$$R^2_{OS} = 1 - \frac{MSFE_1}{MSFE_0}, \tag{30}$$

where $T - T_0$ is the number of out-of-sample evaluation periods. I test the hypothesis $H_0 : MSFE_0 \leq MSFE_1$ vs. $H_1 : MSFE_0 > MSFE_1$, or equivalently $H_0 : R^2_{OS} \leq 0$ vs. $H_1 : R^2_{OS} > 0$. Following the Clark and West (2007) test for nested models. I adjust the point estimate of the difference between two MSFEs for the noise associated with the larger model's forecast and define

$$\hat{f}_{t,t+K} = (r_{t,t+K} - r^B_{t,t+K})^2 - [(r_{t,t+K} - \hat{r}_{t,t+K})^2 - (r^B_{t,t+K} - \hat{r}_{t,t+K})^2]. \tag{31}$$

The test of equal predictive accuracy is conducted by regressing $\hat{f}_{t,t+K}$ on a constant and using the resulting z-statistic for a zero coefficient. The null hypothesis is rejected (equivalent to R^2_{OS} as statistically significant) if this statistic is greater than 1.282 (for a one-sided test at 10% confidence), 1.645 (for a one-sided test at 5% confidence), or 2.334 (for a one-sided test at 1% confidence). When forecast horizon K is greater than one, I adjust for serial correlation and conditional heteroskedasticity using the Newey-West correction with $K - 1$ lags.

Insert Table 6

Table 6 Panel A reports the R^2_{OS} statistics for various predictors and forecast horizons. The out-of-sample R^2_{OS} for the combination of EWIV and VWIV is as high as 0.64% for one-month ahead, 5.46% for six-month ahead, and 11.90% for a one-year ahead forecast horizon. All of them are statistically significant at the 1% level.

Time-series predictability of stock market returns has important implications for market timing by guiding investors to optimally allocate wealth between stock investments and a risk-free asset.[15] I consider a mean-variance-utility investor who allocates wealth between the market portfolio and T-bill. Given an investment horizon of K periods, her optimal weight on the market portfolio is:

$$w_{t,t+K} = \frac{1}{\gamma}\frac{\hat{r}_{t,t+K}}{\hat{\sigma}^2_{t,t+K}}, \tag{32}$$

where $\hat{r}_{t,t+K}$ is conditional expected market excess return (i.e., forecast based on a predictor) given by equation (27). The $\hat{\sigma}^2_{t,t+K}$ is estimated using the variance of the past five-year historical returns, and the relative risk aversion γ is set to be 3. The portfolio is rebalanced every month. The corresponding Sharpe ratio of the investor's optimal portfolio is given by:

$$SR = \frac{R_P}{\sigma_P}, \tag{33}$$

where R_P and σ_P are the mean and the standard deviation of the portfolio return. The average utility gain or the certainty equivalent return (CER) is computed as:

$$CER = R_P - 0.5\gamma\sigma_P^2. \tag{34}$$

To gauge the economic benefit of a predictor to the mean-variance investor, I compare the CER above associated with the optimal portfolio based on the forecasts provided by the predictor to $\overline{CER}$, the certainty equivalent return of a benchmark portfolio formed based on the average return and standard deviation estimated from historical returns. The difference is defined as the CER gain:

$$CER\ Gain = CER - \overline{CER}. \tag{35}$$

[15]The implication of out-of-sample return prediction is documented in for example Kandel and Stambaugh (1996), Campbell and Thompson (2008), Rapach, Strauss, and Zhou (2010), and Ferreira and Santa-Clara (2011).

Table 6 Panel B compares the economic value of using out-of-sample forecasts provided by each predictor to form the optimal portfolio. Consistent with the results in Table 6 Panel A, EWIV and VWIV together outperform other predictors at horizons from one month to two years.

4 Robustness Checks and Discussions

4.1 Alternative Weighting Schemes and Optimal Combinations

In the analysis above, I choose the aggregate idiosyncratic volatilities using EWIV and VWIV, which are commonly used in the literature.[16] However, it is not necessary to be restricted to this choice. Based on Section 2, as long as the two aggregate idiosyncratic volatilities are not perfectly correlated, we should be able to obtain similar empirical results. Therefore, there is flexibility in deciding how to aggregate idiosyncratic volatilities. In this section, I select three alternative weighting schemes to check the robustness of the empirical results. I then econometrically show that among those choices, EWIV and VWIV are closest to the optimal weighting scheme.

The first choice is a combination of EWIV and price-weighted idiosyncratic volatility (PWIV):

$$PWIV_t = \sum_{i=1}^{N_t} p_{i,t} Var_t(\eta_{i,t+1}) = \sum_{i=1}^{N_t} p_{i,t}\beta_{iH,t}b_{iH,t}\sigma^2_{H,t} + \sum_{i=1}^{N_t} p_{i,t}\beta_{iH,t}b_{iM,t}\sigma_{MH,t} + \sum_{i=1}^{N_t} p_{i,t}\sigma^2_{\varepsilon_i,t}, \tag{36}$$

where $p_{i,t}$ is the market price weight and sums up to one. The second choice is a combination of EWIV and aggregate idiosyncratic volatility weighted by the level of idiosyncratic

[16] Most of the previous studies use EWIV and VWIV, including: Guo and Savickas (2008), Bekaert, Hodrick, and Zhang (2012), Bartram, Brown, and Stulz (2018).

volatility (IWIV):

$$IWIV_t = \sum_{i=1}^{N_t} IV_{i,t} Var_t(\eta_{i,t+1}) = \sum_{i=1}^{N_t} IV_{i,t}\beta_{iH,t} b_{iH,t}\sigma^2_{H,t} + \sum_{i=1}^{N_t} IV_{i,t}\beta_{iH,t} b_{iM,t}\sigma_{MH,t} + \sum_{i=1}^{N_t} IV_{i,t}\sigma^2_{\varepsilon_i,t}, \tag{37}$$

where $IV_{i,t}$ is the idiosyncratic volatility weight and sums up to one. The last choice is a combination of two different aggregate idiosyncratic volatility at the portfolio level. Based on Section 2, it is not necessary to aggregate all individuals to obtain the proxies for the conditional covariance term. Instead, as long as I choose two types of aggregate idiosyncratic volatility such that the weights are not perfectly linearly related and the aggregate true idiosyncratic volatility is stable and diversified, one should be able to use the combination to forecast stock market returns. Motivated by this logic, I choose the aggregate idiosyncratic volatilities of two portfolios to obtain the conditional covariance term. In each month I first sort the stocks into five portfolios based on the market capitalization and then compute the portfolio aggregate idiosyncratic volatility as the equal-weighted average of the individual stock idiosyncratic volatility across stocks belonging to the corresponding portfolio. I then choose the combination of aggregate idiosyncratic volatility from the bottom portfolio (SWIV) and from the top portfolio (BWIV) to predict stock market returns:

$$\begin{cases} SWIV_t = \sum_{i=1}^{N_t^S} \frac{1}{N_t^S} Var_t(\eta_{i,t+1}) = \sum_{i=1}^{N_t^S} \frac{1}{N_t^S}\beta_{iH,t} b_{iH,t}\sigma^2_{H,t} + \sum_{i=1}^{N_t^S} \frac{1}{N_t^S}\beta_{iH,t} b_{iM,t}\sigma_{MH,t} + \sum_{i=1}^{N_t^S} \frac{1}{N_t^S}\sigma^2_{\varepsilon_i,t} \\ BWIV_t = \sum_{i=1}^{N_t^B} \frac{1}{N_t^B} Var_t(\eta_{i,t+1}) = \sum_{i=1}^{N_t^B} \frac{1}{N_t^B}\beta_{iH,t} b_{iH,t}\sigma^2_{H,t} + \sum_{i=1}^{N_t^B} \frac{1}{N_t^B}\beta_{iH,t} b_{iM,t}\sigma_{MH,t} + \sum_{i=1}^{N_t^B} \frac{1}{N_t^B}\sigma^2_{\varepsilon_i,t} \end{cases} \tag{38}$$

Insert Table 7

Table 7 shows that all of those alternative weighting schemes are able to predict stock market returns. Note that similar to the case of the combination of EWIV and VWIV, in both situations above, the true idiosyncratic volatility has a limited effect on the empirical results. One should be noted that although the choice of weighting schemes is flexible, it does not mean the approach can work for any combinations. The failure of certain cases may be due to reasons such as unstable aggregate risk exposures (e.g., C_t and D_t) and volatile aggregate true idiosyncratic variance (e.g., Ω_t^{EW} and Ω_t^{VW}).

In addition to this, Corollary 1.2 shows that instead of using conditional variance, one can forecast stock market returns by calculating the first conditional moment across individual firms. As a robustness check, I replace EWIV and VWIV with the aggregate alpha defined in (16) to examine and validate this hypothesis. The aggregate first conditional moments are defined as:

$$\begin{cases} EWAP_t = \sum_{i=1}^{N_t} \frac{1}{N_t} E_t(\eta_{i,t+1}) = \sum_{i=1}^{N_t} \frac{1}{N_t} \gamma_H b_{iH,t} \sigma_{H,t}^2 + \sum_{i=1}^{N_t} \frac{1}{N_t} \gamma_H b_{iM,t} \sigma_{MH,t} \\ VWAP_t = \sum_{i=1}^{N_t} w_{i,t} E_t(\eta_{i,t+1}) = \sum_{i=1}^{N_t} w_{i,t} \gamma_H b_{iH,t} \sigma_{H,t}^2 + \sum_{i=1}^{N_t} w_{i,t} \gamma_H b_{iM,t} \sigma_{MH,t} \end{cases} \tag{39}$$

Table 7 provides consistent evidence that one can achieve the same conclusion of stock market return predictability by using aggregate first conditional moments instead of second moments. The empirical results provide further support of my theoretical explanation using the ICAPM, which can justify the predictive power of both the first and second conditional moments of the misspecified idiosyncratic shocks.

A natural question regarding weighting schemes is: **what is the optimal combination to extract the conditional covariance risk**, or in other words, to give the best return prediction on stock market returns? To answer this question, I apply the three-pass regression filter methodology proposed by Kelly and Pruitt (2015). Kelly and Pruitt (2015) derive a closed-form formula to forecast a single time series using many predictors, which are driven by both infeasible relevant factors and irrelevant factors. Econometrically, the method is able to identify the best subsets of the common components which are related to the forecast target. Following their procedures and adjusting to fit my problem, I implement the three-pass regression filter through the following steps:

(1) Run separate time-series regressions for each firm. The independent variables are (au-

tomatic) proxies constructed based on Kelly and Pruitt (2015, Section 2.5.1). The dependent variables are firm idiosyncratic volatilities.

(2) Run separate cross-sectional regressions at each point of time to extract latent factors that are related to the forecast target. The independent variables are the coefficients of idiosyncratic volatilities of each firm from step (1). The dependent variables are again firm idiosyncratic volatilities. The slopes of the coefficients from the cross-sectional regressions are treated as the estimations of latent factors.

(3) Run a single time-series regression of stock market returns on the predictive factors.

Alternatively, Kelly and Pruitt (2015) show that the whole process can be simplified to one closed-form formula to forecast the target variable:

$$y = \iota\bar{y} + J_T X\hat{\alpha}, \tag{40}$$

$$\hat{\alpha} = W_{XZ}(W_{XZ}^{'} S_{XX} W_{XZ}^{'})^{-1} W_{XZ}^{\prime} s_{Xy}, \tag{41}$$

where $J_t \equiv I_t - \frac{1}{T}\iota_t\iota_t^{\prime}$; I_T is the T-dimensional identity matrix and ι_T is the T-vector of ones (J_N is analogous); $\bar{y} = \iota_T^{\prime} y/T$, $W_{XZ} \equiv J_N X^{\prime} J_T Z$, $S_{XX} \equiv X^{\prime} J_T X$, and $s_{Xy} \equiv X^{\prime} J_T y$. The vector y and matrix X are the target variable (i.e., stock market returns) and observable predictors (i.e., firm idiosyncratic volatility). To implement the three-pass regression filter, during the sample period from 1963 to 2018, in each month I first sort all available stocks into thirty portfolios based on market capitalizations and then compute the portfolio aggregate idiosyncratic volatility as the equal-weighted average among individual idiosyncratic volatilities within the corresponding portfolio.[17] The portfolio aggregate idiosyncratic volatilities are then treated as the predictors in the three-pass regression filter. The target variable is stock market returns one month ahead. The optimal combination of predictors can be

[17]The implementation by portfolios can make sure I have consecutive observations for each range of firm size in the whole sample period. The result and conclusion are robust and consistent if I run the tests at the firm level over time by rolling windows.

obtained by computing $\hat{\alpha}$ in (40). The coefficient $\hat{\alpha}$ can be used as the optimal weights to aggregate idiosyncratic volatilities among the thirty portfolios, and thus construct the best estimation of the conditional covariance risk. As an ex-post analysis, Figure 3 plots the bar chart for the optimal weights of aggregating idiosyncratic volatilities through the thirty portfolios sorted by market capitalization to compute the conditional covariance risk.

Insert Figure 3

The pattern of the bar charts in Figure 3 is consistent with the signs and performance of using EWIV (or EWAP) and VWIV (or VWAP) as an approximate optimal combination to measure the conditional covariance risk. The optimal aggregation gives the most positive weights to small-stock portfolios and most negative weights to big-stock portfolios. Since EWIV (or EWAP) assigns the same weights to all stocks, its dynamic is more close to small-stock idiosyncratic volatilities. Similarly, as VWIV (or VWAP) gives more weights to big stocks, its dynamic is more close to big-stock idiosyncratic volatilities.

To investigate alternative aggregations, I also construct the portfolio idiosyncratic volatilities sorted by various firm characteristics, such as book-to-market ratio, market beta, trading volume, return on equity, and stock illiquidity. I then apply the three-pass regression filter using those portfolios as inputs. Unlike the case of market capitalization, the optimal weights do not show clear and regular patterns among portfolios sorted by other types of firm characteristics. The ex-post optimal weights based on the three-pass regression filter, therefore, justify the choice and the performance of using EWIV and VWIV as an optimal approximation to estimate the conditional covariance risk. Similar results and conclusions can be achieved by using the combination of EWAP and VWAP. The weighting pattern of using first moments is shown in Figure 3 as well.

4.2 Portfolio Return Prediction by EWIV and VWIV

Recall that based on Proposition 2, EWIV and VWIV can be also linked to firm/portfolio return predictability. In order to examine this hypothesis, I investigate time-series return predictability at various portfolio levels. I download different types of (quintile sorted) portfolio returns from Kenneth French's website including market size, book-to-market, operating profitability, and investment portfolios. I run the predictive regression for each portfolio on EWIV, VWIV, and the combination of EWIV and VWIV separately. The results are shown in Table 8. To save space, Table 8 only displays the return prediction for forecast horizons of 6 months and 1 year, though the results are similar for 1 month, 3 months, or 2 years as forecast horizons.

Insert Table 8

Similar to the case of stock market return prediction, Table 8 shows that although neither EWIV nor VWIV is able to forecast portfolio returns independently, EWIV and VWIV jointly are strong predictors for most of the portfolios. The empirical evidence is again consistent with the theoretical prediction in Section 2. It is worth noting that the sign of EWIV (VWIV) is not always positive (negative) at the portfolio level. How the signs of EWIV and VWIV are determined is an interesting question to explore.

4.3 Revisit Goyal and Santa-Clara (2003) and Bali et al. (2005)

Goyal and Santa-Clara (2003) found that EWIV can significantly forecast future stock market returns. Bali et al. (2005) and Wei and Zhang (2005) revisited their study and found that the empirical results are not robust. Bali et al. (2005) extend the sample to recent periods and conclude that idiosyncratic risk does not matter. For example, Bali et al. (2005) found that the predictive power of EWIV does not hold for the extended sample of 1963 to 2001 and for the NYSE stocks. Based on the theoretical framework in Section 2, I show that by considering EWIV and VWIV together, one can obtain consistent predictive power of aggregate idiosyncratic volatility. The results are displayed in Table 9 Panel A and B. The mixed findings between Goyal and Santa-Clara (2003) and Bali et al. (2005) are mainly due to the omitted-variable problem. As stated in Section 2, a single average of idiosyncratic variance includes (at least) both the conditional variance of the hedge portfolio and the conditional covariance between the market portfolio and the hedge portfolio. Since the conditional variance of the hedge portfolio is not related to the equity risk premium, a single aggregate idiosyncratic provides noisy information (i.e., $\sigma_{MH,t}$) about future stock market returns.

Table 9 also documents the robustness of the return predictive power of the combination of EWIV and VWIV. I consider those concerns of aggregate idiosyncratic volatility by Bali et al. (2005) and Wei and Zhang (2005) including sub-sample periods, liquidity issues, and small stock effects.[18] Bali et al. (2005) argue that the predictive power of EWIV (or VWIV) is mainly driven by NASDAQ stocks (i.e., small-stock effect). In order to investigate whether the combined effect of EWIV and VWIV is driven by NASDAQ stocks, I construct an alternative version of EWIV and VWIV by aggregating stocks from the NYSE/AMEX exchange only. Table 9 shows that in all cases one can obtain consistent significant return predictive power through applying the combination of EWIV and VWIV.

While idiosyncratic volatility is used in most of the empirical tests, based on Section 2, the right variable to use is idiosyncratic variance. As a robustness check, I rerun the predictive regression using idiosyncratic variance to construct EWIV and VWIV. Table 9, Panel D confirms that the main conclusion still holds when idiosyncratic variance is used in the regression.

Insert Table 9

Table 9 shows that the predictive power of both EWIV and VWIV holds for different

[18]The market liquidity issues are discussed and examined in the multiple predictive regression in Section 3.

tests and variable specifications. Consequently, the combination largely reconciles the mixed findings between Goyal and Santa-Clara (2003) and Bali et al. (2005). In both sample periods, we can observe consistent forecasting performance of EWIV and VWIV.

4.4 EWIV, VWIV, and Conditional Covariance Risk

Recall that based on Corollary 1.1, the conditional covariance between the market and the hedge portfolio can be expressed as the combination of EWIV and VWIV. In order to explicitly view it, I construct the conditional covariance risk as follows: at the end of each month beginning from 1927 (1816) using the database of CRSP (GFD), I use all available observations of $EWIV_t$, $VWIV_t$, and $\sigma^2_{M,t}$ since 1926 (1815) to fit an expanding-window regression:

$$R_{M,s} = \gamma_{M,t} \times \sigma^2_{M,s-1} + C_t \times EWIV_{s-1} - D_t \times VWIV_{s-1} + \epsilon_{M,s}, \quad s = 1, ..., t \tag{42}$$

Then the estimated conditional covariance is computed as:

$$\hat{\sigma}_{MH,t} \propto (\hat{C}_t \times EWIV_t - \hat{D}_t \times VWIV_t). \tag{43}$$

One should be noted that the estimated variable is not exactly the conditional covariance because γ_H is still unobservable. However, since γ_H is assumed to be constant (or persistent) over time, ($\hat{C}_t \times EWIV_t - \hat{D}_t \times VWIV_t$) can be treated as a one-to-one mapping to the true conditional covariance and therefore has the same dynamic as the true conditional covariance in the ICAPM. Figure 4 plots the conditional covariance estimated from aggregate idiosyncratic variance.

Insert Figure 4

In addition to using aggregate idiosyncratic variance, one can use the first conditional moment of the misspecified idiosyncratic shocks based on Corollary 1.2. Similar to (42) and (43), the conditional covariance can be estimated through:

$$\hat{\sigma}_{MH,t} \propto \left(\frac{\hat{F}_t^{VW}}{\hat{F}_t^{VW}\hat{G}_t^{EW} - \hat{F}_t^{EW}\hat{G}_t^{VW}} EWAP_t - \frac{\hat{F}_t^{EW}}{\hat{F}_t^{VW}\hat{G}_t^{EW} - \hat{F}_t^{EW}\hat{G}_t^{VW}} VWAP_t \right). \tag{44}$$

As a comparison, Figure 5 plots the conditional covariance estimated using either (43) or (44). For convenience, two variables are scaled to have zero mean and one standard deviation. Figure 5 shows that the two time series track closely to each other. The correlation between the two variables is 76%. The empirical evidence further supports the theoretical explanation in Section 2.

Insert Figure 5

Although Merton (1973) acknowledges the additional fact that investors seek to hedge against shortfalls in consumption or against changes in the future investment opportunity set, he does not specify what the state variables are and how such variables map into the conditional covariance. Previous articles try to measure the conditional covariance risk through certain pre-specified models. For example, Merton (1973) suggests that the risk-free interest rate is a proxy to which investors would desire to hedge against its unanticipated adverse changes. Scruggs (1998) was the first to include the return on a long-term Treasury bond as an additional state variable noting that its omission induces a negative bias in the estimate of the price of risk associated with the return on the market. Guo and Whitelaw (2006) assume that the conditional covariance is a linear function of a vector of observable state variables including the relative Treasury bill rate and the CAY variable of Lettau and Ludvigson (2001a). Rossi and Timmermann (2015) link the conditional covariance risk to a broad economic activity index. If my hypothesis is correct that EWIV and VWIV together are a proxy for the conditional covariance, it should be largely related to the time-varying investment opportunities and those state variables.

To investigate the link between the conditional covariance and those state variables, I conduct an empirical test using the framework of the vector autoregression (VAR) with one lag. Based on the previous studies, I consider the following state variables in my test: consumption growth, income growth, CAY, unemployment growth, industrial production growth, term spread, default spread, and dividend-price ratio. The VAR is model specified as:

$$Z_{K,t+1} = \phi_0 + \phi_1 Z_{K,t} + \nu_{K,t+1}, \tag{45}$$

where $Z_{K,t+1} = \{\hat{\sigma}_{MH,t}, \text{consumption growth}, \text{income growth}, \text{CAY}, ...\}$ is a vector of variables of interest; ϕ_0 and ϕ_1 are two vectors of coefficients; $\nu_{K,t+1}$ is the vector of residuals of each state variable. I collect the monthly data of consumption growth, income growth, unemployment growth and industrial production growth from FRED-MD, a monthly frequency macroeconomic database with consistent and comparable adjustment provided by Federal Reserve Economic Data, St. Louis. CAY data is from Sydney Ludvigson's website. I use the latest quarterly CAY as the observation for each month. The rest of the data are obtained from Amit Goyal's website.

Insert Table 10

Table 10 exhibits consistent evidence that the measure $\hat{\sigma}_{MH,t}$ is closely related to those state variables, which are commonly used to model conditional covariance risk. For example, the results in Table 10, Panel A are consistent with both Guo and Whitelaw (2006) that the CAY variable is an important state variable to model conditional covariance, and Rossi and Timmermann (2015) that economic activities such as consumption growth, term spread, and dividend yield can be used to track time-varying investment opportunities.

4.5 Cross-Sectional Implication of EWIV and VWIV

Corollary 2.1 sustains that EWIV and VWIV together can help identify the conditional risk exposures to the missing hedge portfolio in Merton's ICAPM. I conduct the following tests to verify this hypothesis. During the sample period from 1963 to 2018, I first run a time-series regression of market excess returns on SMV, EWIV, and VWIV using a 48-month rolling window:

$$R_{M,s} = \gamma_{M,t} \times \sigma^2_{M,s-1} + C_t \times EWIV_{s-1} - D_t \times VWIV_{s-1} + \epsilon_{M,s}, \quad s = t-47, ..., t. \quad (46)$$

I then run a similar regression at the individual level:

$$R_{i,s} = \beta_{iM,t} \times R_{M,s} + C_{i,t} \times EWIV_{s-1} - D_{i,t} \times VWIV_{s-1} + (\beta_{iH,t} \times \epsilon_{H,s} + \epsilon_{i,s}), \quad s = t-47, ..., t \quad (47)$$

where $R_{i,t}$ is the excess return of firm i at time t. Each firm to be included should have more than 30-month observations and stock prices are great than \$5. Based on Corollary 2.1, the conditional risk exposures to the missing factor (i.e., hedge portfolio) can be obtained from one of the coefficients below:

$$\begin{cases} \hat{\beta}_{iH,t} \propto \hat{C}_{i,t} \text{ at cross-section} \\ \hat{\beta}_{iH,t} \propto \hat{D}_{i,t} \text{ at cross-section} \\ \hat{\beta}_{iH,t} \propto (\hat{D}_t \times \hat{C}_{i,t} - \hat{C}_t \times \hat{D}_{i,t}) \text{ at cross-section} \end{cases} \quad (48)$$

Four conditions are specific to my analysis:

(1) In order to be consistent with the framework in Section 2, when running the time-series regression, I do not include other risk factors. Furthermore, instead of using changes/innovations of EWIV or VWIV to calculate risk loadings, I use the levels of EWIV and VWIV, which are different from some of the previous studies such as Chen and Petkova (2012) and Herskovic

et al. (2016).

(2) Since Corollary 2.1 is based on the conditional ICAPM, both risk exposures ($\beta_{iH,t}$) and risk premium $\mu_{H,t}$ are time-varying. Jagannathan and Wang (1996) show that the conditional CAPM can be unconditioned to a two-factor model due to the covariance term between risk exposures and risk premia. Therefore, a rolling window is more reasonable in my case to capture the time-varying changes of the conditional risk exposures.

(3) Equation (47) is not an unbiased regression, since the residual term $\epsilon_{H,s}$ is correlated with $R_{M,s}$. An alternative regression is to use all conditional expectations on the right-hand side, for example:

$$R_{i,s} = \beta_{iM,t}\gamma_{M,t}\times\sigma^2_{M,s-1}+C_{i,t}\times EWIV_{s-1}-D_{i,t}\times VWIV_{s-1}+(\beta_{iM,t}\epsilon_{M,s}+\beta_{iH,t}\epsilon_{H,s}+\epsilon_{i,s}), \quad s=t-47,...,t. \tag{49}$$

It is worth mentioning here that the mechanism of computing risk loadings using conditional expectations of random variables above is different from that by using realized random variables. The former is based on time-varying variation of random variables, while the latter is based on realized shocks.

(4) Corollary 2.1 only demonstrates that the true conditional risk exposures proportionate to the coefficients from EWIV and VWIV under certain conditions. They are not true risk exposures and only applied to a two-factor model.

After obtaining the estimated conditional risk exposures, at the end of each month, I sort all eligible stocks based on either $\hat{\beta}_{iM,t}$ or one of those three proxies for $\hat{\beta}_{iH,t}$ into quintiles. Within each quintile, I compute both equal-weighted and value-weighted returns based on the market capitalization at the end of the previous month and then construct a long-short portfolio between the top and bottom quintiles. The portfolios are held until the end of next month. Table 11 displays each portfolio's raw returns and the long-short portfolio returns and alphas based on the Fama-French five-factor model. In addition to investigating the effect of $\hat{\beta}_{iM,t}$, I extend the portfolio analysis in Panel C through double sorting the stocks by $\hat{\beta}_{iM,t}$ first and then by $\hat{\beta}_{iH,t}$ within each quintile.

Insert Table 11

Table 11 confirms the theoretical prediction that the combined effect of EWIV and VWIV has an important implication for the cross-sectional variation of stock returns. Firms with higher $\hat{\beta}_{iH,t}$ earn higher returns than firms with lower $\hat{\beta}_{iH,t}$ on average, implying that the risk premium of the hedge portfolio is positive. Furthermore, the sorting performance is consistent and robust among all three proxies specified in Corollary 2.1. The magnitude of the monthly long-short strategy is economically significant and comparable to other factors in the literature. For example, the long-short strategy based on $\hat{\beta}_{iH,t} \propto (\hat{D}_t \times \hat{C}_{i,t} - \hat{C}_t \times \hat{D}_{i,t})$ earns 0.66% in raw returns monthly (around 7.88% annually). The double-sorting test also confirms that the explanatory power of $\hat{\beta}_{iH,t}$ on average returns of stocks cannot be explained by market risk exposures. It is worth mentioning here that when $\hat{\beta}_{iM,t}$ is calculated under the control of EWIV and VWIV, one is able to observe a more monotonic increasing relationship between $\hat{\beta}_{iM,t}$ and future stock returns. The cross-sectional results then highlight the importance of considering EWIV and VWIV together to reconcile the market beta anomaly, which is well documented in the literature.

To further affirm the cross-sectional finding and its difference from existing risk factors, I conduct the Fama-MacBeth regression proposed by Fama and MacBeth (1973). For each month, I run a cross-sectional regression of their returns on $\hat{\beta}_{iH,t}$ and other control variables, which have been linked to stock returns in the literature. The cross-sectional regression above is used to obtain the coefficients for the independent variables. After obtaining the time series of the coefficients for the independent variables, I conduct the t-test for each coefficient with one-lag correction of Newey and West (1987). The control variables include the classic factors such as $\hat{\beta}_{iM,t}$, market size, book-to-market ratio, momentum, CIV beta (Herskovic, et al. 2016), and idiosyncratic volatility. The empirical results are provided in Table 12.

Insert Table 12

Table 12 supports that EWIV and VWIV when used together become an important determinant of cross-sectional variations of stock returns. The cross-sectional evidence is consistent with the theoretical framework in Section 2, and further confirms that the empirical findings can be explained under Merton's ICAPM.

5 Tail Risk and Conditional Covariance Risk

Kelly and Jiang (2014) (hereafter KJ (2014)) propose a new measure of time-varying tail risk through the cross-section of stock returns. They show that tail risk has strong predictive power for aggregate stock market returns and that tail risk also has cross-sectional implementations on individual stock returns. The tail risk measure can negatively forecast real economic activity as well. The asset pricing facts are consistent with the perspective of structural models with heavy-tailed firm-level shock distributions that are preserved under aggregation. They conclude that the power-law aggregation and the real effects of uncertainty shocks represent potential channels through which firm-level tail risk can influence asset prices.

Chapman, Gallmeyer, and Martin (2018) (hereafter CGM (2018)) revisit KJ (2014) and raise some empirical concerns regarding this finding. First, CGM (2018) found that the tail risk proposed by KJ (2014) has a weak correlation with theoretically motivated measures of tail risk.[19] CGM (2018) also found that the variable explains the cross-section of the discount rate component of returns, but not the cash-flow component. The impact of tail risk on real quantities appears to be inconsistent with a structural model, such as a rare disaster model, that generates tail outcomes through large real cash-flow effects. On the contrary, CGM (2018) found that there is a strong and negative correlation between changes

[19] Typical examples include aggregate uncertainty (Jurado, Ludvigson, and Ng (2015); Bloom (2014)) and systemic risk (Allen, Bali, and Tang (2012)).

in the tail measure and subsequent changes in the 10-year Treasury yield. While CGM (2018) questioned the link of the risk measure proposed by KJ (2014) to tail risk, they don't explain why this measure can predict time-series stock market returns and explain the cross-sectional variation of individual stock returns.

In this section, I provide an alternative reconciliation regarding the debate between KJ (2014) and CGM (2018). I first show theoretically that under certain conditions the variable proposed by KJ (2014) may capture the conditional covariance in Merton's ICAPM. Empirically, I find that the tail risk measure proposed by KJ (2014) is highly correlated with the conditional covariance estimated by EWIV and VWIV. Furthermore, when including the conditional covariance in the predictive regression, the tail risk measure loses the significance to forecast stock market returns. My hypothesis can explain why CGM (2018) found that the tail risk measure mainly forecasts stocks' discount rate components instead of cash flow components. The new hypothesis can also interpret the strong correlations between changes in the tail measure and subsequent changes in the level of yield on a 10-year Treasury bond, and changes in the slope of the Treasury yield curve. Both of them are commonly used state variables to model the conditional covariance in the literature (Rossi and Timmermann (2015); Guo and Whitelaw (2006)).

Recall that based on Proposition 1 and Corollary 1.2, besides the conditional misspecified idiosyncratic variance, the first conditional moment of the misspecified idiosyncratic shock can also be used to estimate the conditional covariance between the market portfolio and the hedge portfolio. Therefore, it is possible that the tail risk measure (i.e., λ_{t+1}^{Hill}), which is constructed based on the first conditional moment of stock returns, may be linked to the conditional covariance risk like EWIV and VWIV.

Proposition 3. *Suppose that the stock return follows Merton's ICAPM defined in (7), but econometricians use the CAPM as the asset pricing model defined in (8). Furthermore, the* *empirical* λ_{t+1}^{Hill} *is defined as:*

$$\lambda_{t+1}^{Hill} = \frac{1}{K_{t+1}} \sum_{k=1}^{K_{t+1}} \log\left(\frac{\eta_{k,t+1}}{u_{t+1}}\right). \quad (50)$$

Under moderate assumptions, λ_{t+1}^{Hill} *can be written as a linear function of* $\sigma_{MH,t}$ *in the ICAPM*

$$\lambda_{t+1}^{Hill} \approx \log\left(\frac{J_t^K}{J_t^u}\right) + \frac{1}{x_0}\sigma_{MH,t} + \varepsilon_{t+1}, \quad (51)$$

where

$$\begin{cases} J_t^u = \frac{F_t^u}{F_t^u G_t^K - F_t^K G_t^u} \\ J_t^K = \frac{F_t^K}{F_t^u G_t^K - F_t^K G_t^u}, \end{cases} \quad \begin{cases} F_t^K = \sum_{k=1}^{K_{t+1}} \frac{1}{K_{t+1}} \gamma_H b_{kH,t} \\ G_t^K = \sum_{k=1}^{K_{t+1}} \frac{1}{K_{t+1}} \gamma_H b_{kM,t}, \end{cases} \quad \begin{cases} F_t^u = \gamma_H b_{uH,t} \\ G_t^u = \gamma_H b_{uM,t} \\ x_0 = \frac{1}{2}[E(J_t^u)E(\bar{\eta}_{k,t+1}) + E(J_t^K)E(u_{t+1})]. \end{cases}$$

Proofs and corresponding assumptions are given in the Appendix. Proposition 3 provides an alternative theoretical explanation to why λ_t^{Hill} is able to forecast stock market returns is because it is related to the conditional covariance in the ICAPM. Empirically, I confirm that the variable proposed by KJ (2014) is a robust predictor for aggregate stock market returns from 1963 to 2018. The predictive power is both statistically and economically significant. I first show that besides using λ_t^{Hill}, $\bar{\eta}_{k,t}$ and u_t are also strong predictors of stock market returns.

Insert Table 13

Table 13 shows that similar to the previous empirical results, although separately $\bar{\eta}_{k,t}$ nor u_t is able to forecast stock market returns, $\bar{\eta}_{k,t}$ and u_t together become strong predictors. To compare directly with λ_t^{Hill}, I then construct the conditional covariance risk estimate based on Section 4.4 by running a recursive predictive regression since 1926 to the present date and using the coefficients to compute the conditional covariance risk ($\hat{\sigma}_{MH,t}$). One would expect that the two variables will have a very high correlation with each other. Empirically, I find the same evidence in the data.

Insert Figure 6

Figure 6 shows that the two variables track each other closely. The correlation between $\hat{\sigma}_{MH,t}$ and λ_t^{Hill} is as high as 0.795 from 1926 to 2018.[20] In addition to this, Section 4.1 and 4.3 have shown that equal-weighted alpha (EWAP) and value-weighted alpha (VWAP) also have comparable predictive power for the stock market returns, and that the conditional covariance risk estimated using EWAP and VWAP is highly correlated with that computed from EWIV and VWIV (see Figure 4). Given that $\hat{\sigma}_{MH,t}$ uses all information (either first or second conditional moments) across individual stocks, it is difficult to be treated as tail risk.

Regarding the relationship between λ_t^{Hill} and cross-sectional variations of stock returns, the connection can be justified through the ICAPM as well. Based on Proposition 2 and 3, the conditional expected return of any asset can be written as:

$$\mu_{i,t} = \beta_{iM,t}\mu_{M,t} + \beta_{iH,t}\gamma_H\sigma_{MH,t} + \beta_{iH,t}\gamma_H\sigma_{H,t}^2 \approx \beta_{iM,t}\mu_{M,t} + \beta_{iH,t}\gamma_H\lambda_t^{Hill} + \beta_{iH,t}\gamma_H\sigma_{H,t}^2. \quad (52)$$

Suppose we further assume that γ_H are constant over time and $\beta_{iH,t}$ are stable within certain time periods, one can observe the asset-pricing implication of λ_t^{Hill} on the cross-section of stock returns because the factor loading is linked to the conditional risk exposure to the hedge portfolio ($\beta_{iH,t}$).[21] Equation (52) is also consistent with the cross-sectional test setting in KJ (2014). When calculating the tail risk sensitivities of individual stocks, they use the predictive regression on the lagged level of λ_t^{Hill} instead of its contemporaneous shocks, which are more widely used to calculate factor loadings in the literature.

[20] An alternative simple proxy for the conditional covariance can be calculated by taking the ratio between EWIV and VWIV. The ratio is also highly correlated with λ_t^{Hill} with a correlation equal to 0.761.

[21] The details are provided in the proof of Corollary 2.1 in the appendix. Note that since λ_t^{Hill} doesn't include the effect of $\sigma_{H,t}^2$, the coefficient from the predictive regression of stock excess returns on λ_t^{Hill} is a biased estimator of $\beta_{iH,t}$.

However, one should be aware that the analysis above only provides an alternative explanation to the tail risk measure. Section 5 does not provide any direct evidence to reject the tail-risk explanation proposed by Kelly and Jiang (2014).

6 Simulation Evidence

This section builds on the framework developed in Section 2. In order to understand the mechanism between idiosyncratic variance and the missing covariance term in the ICAPM, I conduct a numerical analysis based on the theoretical framework in Section 2. The simulation is designed in the following way:

Assuming that stock price dynamics are driven by the ICAPM with pre-specified parameters, I first simulate the whole stock market, which includes individual stocks, the market portfolio and the hedge portfolio, based on the true parameters. For instance, the relative risk aversion (γ_M) and weighted average state-variable aversion (γ_H) are set equal to 3 and 2.6 respectively.[22] The conditional variance of the market portfolio ($\sigma^2_{M,t}$) and the hedge portfolio ($\sigma^2_{H,t}$) is simulated based on a GARCH (1, 1) model fitted using the empirical data of the stock market returns for $\sigma^2_{M,t}$ and the HML factor premiums for $\sigma^2_{H,t}$:

$$\sigma^2_{t+1} = \omega + \alpha\varepsilon^2_t + \beta\sigma^2_t, \text{ with } \alpha + \beta < 1, \tag{53}$$

where ε_t is obtained from a normal distribution with zero mean and a standard deviation, which is matched through the market return residuals from regressing stock market returns (HML factor) on SMV, EWIV, and VWIV for $\sigma^2_{M,t}$ ($\sigma^2_{H,t}$). Then the realized returns of the

[22]The level of state-variable aversions is set based on Rossi and Timmermann (2015).

market portfolio and the hedge portfolio are computed as:

$$\begin{cases} R_{M,t} = \gamma_M \sigma_{M,t}^2 + \gamma_H \sigma_{MH,t} + \varepsilon_{M,t} \\ R_{H,t} = \gamma_M \sigma_{MH,t} + \gamma_H \sigma_{H,t}^2 + \varepsilon_{H,t} \end{cases} \tag{54}$$

To simulate individual stocks, I construct the following variables:

(1) **The beta exposures to the market portfolio (hedge portfolio) for each stock**: is simulated based on an AR (1) model:

$$\beta_{i,t+1} = \phi_{i,0} + \phi_{i,1}\beta_{i,t} + \nu_{i,t+1}, \tag{55}$$

where $\phi_{i,0}$ and $\phi_{i,1}$ are calibrated based on the empirical distribution of the market beta and the HML beta among firms. For example, the initial value of the market beta is generated based on a normal distribution of mean equal to 0.5495 and standard deviation equal to 1.661. $\phi_{i,0}$ ($\phi_{i,1}$) is simulated based on a normal distribution with mean equal to 0.474 (0.439) and standard deviation equal to 0.088 (0.214). $\nu_{i,t+1}$ is an $i.i.d$ random risk-exposure shock across all firms with zero mean and standard deviation equal to 0.064.

(2) **The realized returns for each stock**: are simulated based on the information of the market and hedge portfolio and their corresponding risk exposures, with the model specified in (7). For simplicity and without loss of generality, the firm idiosyncratic shocks are assumed to be $i.i.d.$

(3) **The price dynamics for each stock**: are then obtained by multiplying one plus the raw return by the previous stock price. The cross-sectional distribution of the initial stock price is generated based on a uniform distribution within the range between \$1.00 and \$2000. During the sample periods, if the stock price goes below \$1.00, the firm will be excluded from the sample as it is treated as delisted.

(4) For simplicity and without loss of generality, I assume **the shares outstanding for**

each firm is constant over time. The cross-sectional distribution of the shares outstanding is generated from a uniform distribution within the range from 2,000 to 4,000,000. **The market capitalization for each firm** is then computed as the product of stock price and shares outstanding.

After obtaining all necessary variables, I construct a synthetic stock market with 2000 stocks over 2000 months and simulate the market for 10,000 times. For each simulation, I follow the procedure in Section 3 and 4 to compute the corresponding SMV, EWIV, VWIV, and λ_t^{Hill} proposed by Kelly and Jiang (2014) with monthly observations. The simulated variables are used to run the time-series regressions similar to the previous empirical tests in Section 3.2. I subsequently count the proportion of the simulation studies, which are consistent with my hypothesis and real empirical evidence (i.e., a predictive regression with significant coefficients and right signs). The results are displayed in Table 14.

Insert Table 14

Table 14 shows a similar pattern that when independent variables are highly correlated and one of them is missing in the specification, the existing independent variables are subject to serious omitted variable bias. For example, consistent with the empirical observations, when running univariate regressions of simulated market returns on EWIV (VWIV), only 20% (12.50%) of the total simulations are significant with correct signs. The simulation results, however, are different when omitted variables are included in the regression. For instance, in column IV and V, both EWIV and VWIV become significant in 98.00% and 99.20% of all simulations with correct signs. Consistent with Proposition 3, Table 14, column VI provides simulation evidence that the tail risk measure proposed by Kelly and Jiang (2014) is also able to forecast stock market returns under the framework of Merton's ICAPM.

7 Conclusion

The importance of idiosyncratic volatility is an essential topic in asset pricing. While most of the previous research focuses on cross-sectional studies of idiosyncratic volatility, this paper examines asset pricing implications of aggregate idiosyncratic volatility. I contribute to the literature that aggregate idiosyncratic volatility in fact matters to asset pricing in both time series and cross-section. Empirically, I document that equal-weighted idiosyncratic volatility (EWIV) and value-weighted idiosyncratic volatility (VWIV) jointly are strong predictors of aggregate stock returns in both short- and long-term horizons. The firm loadings obtained from the regressions including both EWIV and VWIV also explain cross-sectional variations of stock returns. I argue that these findings can be understood under the framework of Merton's (1973) ICAPM that EWIV and VWIV when applied together are a proxy for the conditional covariance between the market portfolio and the hedge portfolio. Econometrically, the choice of using EWIV and VWIV is close to an optimal approximation of the conditional covariance risk. Based on this framework, I revisit two debates regarding idiosyncratic volatility and tail risk in the literature and provide new insights and reconciliations pertaining to their mixed findings.

Several relevant questions are still open for discussion. For example, since the ICAPM affirms that stock market return is related to conditional market variance and conditional covariance, it is reasonable to infer that after controlling for conditional covariance, we should observe a significant positive relationship between conditional variance and stock market return (i.e., risk-return tradeoff). I find this is not true in the data. Although a significant positive risk-return tradeoff is observed in the sample period from 1963 to 2018, it is not as stable as the conditional covariance. For example, if I look at longer sample periods (e.g., 1815 to 2018), or recent sample periods (e.g., 1999 to 2018), the risk-return tradeoff becomes insignificant. Future research can explore why we still are unable to find a stable significant positive relationship between stock market returns and conditional market variance, even after controlling for the conditional covariance. One potential explanation is that the variation of the representative agent's risk aversion (i.e.,γ_M) might be very volatile.

Another question, which has not been solved in this article, is regarding the economic reason why EWIV and VWIV are close to the optimal combination to extract the conditional covariance risk. In Section 4.1, by applying the three-pass regression filter, I demonstrate that the combination of EWIV and VWIV is close to the optimal proxies for the conditional covariance risk econometrically. More specifically, a rise of small-firm idiosyncratic volatilities increases the conditional covariance risk, while a rise of big-firm idiosyncratic volatilities decreases the conditional covariance risk. However, the regression methodology is unable to resolve the economic mechanism behind this conclusion. In other words, why do small-firm (big-firm) idiosyncratic volatilities seem to be a bad (good) thing to investors over time? I will leave it to future research.

Appendix

A Proof of Proposition 1

Consider the true return generating process:

$$R_{i,t+1} = \beta_{iM,t}(\mu_{M,t} + \varepsilon_{M,t+1}) + \beta_{iH,t}(\mu_{H,t} + \varepsilon_{H,t+1}) + \varepsilon_{i,t+1}, \tag{A.1}$$

where $\varepsilon_{M,t+1}$, $\varepsilon_{H,t+1}$, $\varepsilon_{i,t+1}$ are the unexpected shocks for the market portfolio, the hedge portfolio, and asset i. $R_{M,t+1}$ is the market excess return, $R_{H,t+1}$ is the return of the hedge portfolio, $\beta_{iM,t}$ and $\beta_{iH,t}$ are the corresponding risk exposures:

$$\beta_{iM,t} = \frac{\sigma_{iM,t}\sigma^2_{H,t} - \sigma_{iH,t}\sigma_{MH,t}}{\sigma^2_{H,t}\sigma^2_{M,t} - \sigma^2_{MH,t}}, \beta_{iH,t} = \frac{\sigma_{iH,t}\sigma^2_{M,t} - \sigma_{iM,t}\sigma_{MH,t}}{\sigma^2_{H,t}\sigma^2_{M,t} - \sigma^2_{MH,t}}, \tag{A.2}$$

Suppose that econometricians only use the CAPM as the asset pricing model:

$$R_{i,t+1} = b_{iM,t}(\mu_{M,t} + \varepsilon_{M,t+1}) + \eta_{i,t+1}, \tag{A.3}$$

where $\eta_{i,t+1}$ is the misspecified idiosyncratic shock for asset i; $b_{iM,t}$ is the risk exposure under the CAPM and can be written as $b_{iM,t} = \frac{\sigma_{MH,t}}{\sigma^2_{M,t}}$. It can be easily shown that:

$$\beta_{iM,t} - b_{iM,t} = \frac{\sigma_{iM,t}\sigma^2_{H,t} - \sigma_{iH,t}\sigma_{MH,t}}{\sigma^2_{H,t}\sigma^2_{M,t} - \sigma^2_{MH,t}} - \frac{\sigma_{iM,t}}{\sigma^2_{M,t}} = -\beta_{iH,t}\frac{\sigma_{MH,t}}{\sigma^2_{M,t}}. \tag{A.4}$$

The conditional expectation of the misspecified firm idiosyncratic shock is given by:

$$\begin{aligned} E_t(\eta_{i,t+1}) &= \mu_{i,t} - b_{iM,t}\mu_{M,t} \\ &= \beta_{iM,t}\mu_{M,t} + \beta_{iH,t}\mu_{H,t} - (\beta_{iM,t} + \beta_{iH,t}\frac{\sigma_{MH,t}}{\sigma^2_{M,t}})\mu_{M,t} \\ &= \beta_{iH,t}\mu_{H,t} - \beta_{iH,t}\frac{\sigma_{MH,t}}{\sigma^2_{M,t}}\mu_{M,t} \end{aligned} \tag{A.5}$$

In addition, based on Merton's ICAPM, the market portfolio and the hedge portfolio have the following relationship:

$$\begin{cases} \mu_{M,t} = \gamma_M\sigma^2_{M,t} + \gamma_H\sigma_{MH,t} \\ \mu_{H,t} = \gamma_M\sigma_{MH,t} + \gamma_H\sigma^2_{H,t} \end{cases} \tag{A.6}$$

Plug back to the equation of the expected misspecified idiosyncratic shocks:

$$
\begin{aligned}
E_t(\eta_{i,t+1}) &= \beta_{iH,t}(\gamma_M \sigma_{MH,t} + \gamma_H \sigma_{H,t}^2) - \beta_{iH,t}\frac{\sigma_{MH,t}}{\sigma_{M,t}^2}(\gamma_M \sigma_{M,t}^2 + \gamma_H \sigma_{MH,t}) \\
&= \gamma_H \beta_{iH,t}\frac{\sigma_{H,t}^2\sigma_{M,t}^2 - \sigma_{MH,t}^2}{\sigma_{M,t}^2}, \qquad \because \beta_{iH,t} = \frac{\sigma_{iH,t}\sigma_{M,t}^2 - \sigma_{iM,t}\sigma_{MH,t}}{\sigma_{H,t}^2\sigma_{M,t}^2 - \sigma_{MH,t}^2} \\
&= \gamma_H \frac{\sigma_{iH,t}}{\sigma_{H,t}^2}\sigma_{H,t}^2 + \gamma_H \frac{\sigma_{iM,t}}{\sigma_{M,t}^2}\sigma_{MH,t} \\
&= \gamma_H b_{iH,t}\sigma_{H,t}^2 + \gamma_H b_{iM,t}\sigma_{MH,t}.
\end{aligned} \tag{A.7}
$$

Similarly, the conditional variance of the misspecified firm idiosyncratic shock is:

$$
\begin{aligned}
Var_t(\eta_{i,t+1}) &= Var_t(R_{i,t+1} - b_{iM,t}R_{M,t+1}) \\
&= \beta_{iH,t}^2\left(\frac{\sigma_{MH,t}}{\sigma_{M,t}^2}\right)^2 \sigma_{M,t}^2 + \beta_{iH,t}^2\sigma_{H,t}^2 - 2\beta_{iH,t}\beta_{iH,t}\frac{\sigma_{MH,t}^2}{\sigma_{M,t}^2} + \sigma_{\varepsilon_i,t}^2 \\
&= \beta_{iH,t}^2\frac{\sigma_{H,t}^2\sigma_{M,t}^2 - \sigma_{MH,t}^2}{\sigma_{M,t}^2} + \sigma_{\varepsilon_i,t}^2, \qquad \because \beta_{iH,t} = \frac{\sigma_{iH,t}\sigma_{M,t}^2 - \sigma_{iM,t}\sigma_{MH,t}}{\sigma_{H,t}^2\sigma_{M,t}^2 - \sigma_{MH,t}^2} \\
&= \beta_{iH,t}\frac{\sigma_{iH,t}}{\sigma_{H,t}^2}\sigma_{H,t}^2 + \beta_{iH,t}\frac{\sigma_{iM,t}}{\sigma_{M,t}^2}\sigma_{MH,t} + \sigma_{\varepsilon_i,t}^2 \\
&= \beta_{iH,t}b_{iH,t}\sigma_{H,t}^2 + \beta_{iH,t}b_{iM,t}\sigma_{MH,t} + \sigma_{\varepsilon_i,t}^2.
\end{aligned} \tag{A.8}
$$

This completes the prrof.

B Proof of Corollary 1.1

Based on Proposition 1 and the CAPM in equation (8), consider the equal-weighted idiosyncratic variance and value-weighted idiosyncratic variance for the CAPM:

$$
EWIV_t = \sum_{i=1}^{N_t}\frac{1}{N_t}Var_t(\eta_{i,t+1}) = A_t^{EW}\sigma_{H,t}^2 + B_t^{EW}\sigma_{MH,t} + \Omega_t^{EW}, \tag{A.9}
$$

where $A_t^{EW} = \sum_{i=1}^{N_t}\frac{1}{N_t}\beta_{iH,t}b_{iH,t}$, $B_t^{EW} = \sum_{i=1}^{N_t}\frac{1}{N_t}\beta_{iH,t}b_{iM,t}$, $\Omega_t^{EW} = \sum_{i=1}^{N_t}\frac{1}{N_t}\sigma_{\varepsilon_i,t}^2$. Similarly, the conditional VWIV can be defined as:

$$
VWIV_t = A_t^{VW}\sigma_{H,t}^2 + B_t^{VW}\sigma_{MH,t} + \Omega_t^{VW}, \tag{A.10}
$$

where $A_t^{VW} = \sum_{i=1}^{N_t} w_{i,t}\beta_{iH,t}b_{iH,t}$, $B_t^{VW} = \sum_{i=1}^{N_t} w_{i,t}\beta_{iH,t}b_{iM,t}$, $\Omega_t^{VW} = \sum_{i=1}^{N_t} w_{i,t}\sigma_{\varepsilon_i,t}^2$. $w_{i,t}$ is the market-cap weight of each firm at the end of time t. One can clearly see that either EWIV or VWIV is a linear function of at least three time-varying components: the conditional variance of the hedge portfolio, the conditional covariance of the market portfolio with the hedge portfolio, and the average true idiosyncratic variance. Three points are worthy of mentioning here:

(1) The reason why to decompose the variation of return into the risk exposures is that previous literature finds that most of the predictable variation in returns can be attributed to predictable variation in risk premia (Ferson and Harvey (1991); Evans (1994)) instead of time-varying risk exposures. The finding is consistent with my empirical results. For example, if I calculate the value-weighted average market beta and idiosyncratic volatility separately, the coefficients of variations are 0.0659 for the market beta and 0.3318 for VWIV. By decomposing the idiosyncratic variance into risk exposures and other variations, I am able to analyze which part determines the variation of time-varying risk premium.

(2) When choosing the weighting schemes, to void degeneracy for the future derivation, I need to choose the weights so that A_t^{EW} and B_t^{EW} are not perfectly linearly related to A_t^{VW} and B_t^{VW}. At the same time, I also need to choose the weights so that the average true idiosyncratic variance can be diversified enough and stable. In my case of EWIV and VWIV, this criterion is easy to be satisfied.

(3) Since either EWIV or VWIV is constructed from a bottom-up approach, the aggregate true idiosyncratic variance (Ω_t^{EW} and Ω_t^{VW}) cannot be negligible, thus leading to estimation biases. However, under the moderate assumption that the conditional covariance matrix of the true idiosyncratic shocks is relatively stable, when running regressions, the effect of the true idiosyncratic shocks can be captured by the intercept term in the empirical tests. Therefore, the effect caused by the true idiosyncratic variance can be ignored. The detailed effect of the aggregate true idiosyncratic variance is discussed in the simulation study in Section 6.

Since both $\sigma_{H,t}^2$ and $\sigma_{MH,t}$ are unobservable, I use two types of weighted average idiosyncratic variance to solve the linear system. Similar techniques are used in other papers such as Lo and Wang (2006). For simplicity, I define the following:

$$\begin{cases} \widetilde{EWIV}_t \equiv EWIV_t - \Omega_t^{EW} = A_t^{EW}\sigma_{H,t}^2 + B_t^{EW}\sigma_{MH,t} \\ \widetilde{VWIV}_t \equiv VWIV_t - \Omega_t^{VW} = A_t^{VW}\sigma_{H,t}^2 + B_t^{VW}\sigma_{MH,t} \end{cases} \tag{A.11}$$

It is a linear system with two linearly independent equations with two unknowns. One can easily figure out $\sigma^2_{H,t}$ and $\sigma_{MH,t}$:

$$\begin{cases} \sigma^2_{H,t} = \frac{B^{EW}_t}{A^{VW}_t B^{EW}_t - A^{EW}_t B^{VW}_t} \widetilde{VWIV}_t - \frac{B^{VW}_t}{A^{VW}_t B^{EW}_t - A^{EW}_t B^{VW}_t} \widetilde{EWIV}_t \\ \sigma_{MH,t} = \frac{A^{VW}_t}{A^{VW}_t B^{EW}_t - A^{EW}_t B^{VW}_t} \widetilde{EWIV}_t - \frac{A^{VW}_t}{A^{VW}_t B^{EW}_t - A^{EW}_t B^{VW}_t} \widetilde{VWIV}_t \end{cases} \tag{A.12}$$

This completes the Proof.

C Proof of Corollary 1.2

The proof is similar to Corollary 1.1, except that I use the first conditional moment instead of the second conditional moment. Similarly, I consider the equal-weighted and value-weighted average conditional expectation of the misspecified idiosyncratic:

$$EWAP_t = \sum_{i=1}^{N_t} \frac{1}{N_t} E_t(\eta_{i,t+1}) = F^{EW}_t \sigma^2_{H,t} + G^{EW}_t \sigma_{MH,t}, \tag{A.13}$$

where $F^{EW}_t = \sum_{i=1}^{N_t} \frac{1}{N_t} \gamma_H b_{iH,t}$, $G^{EW}_t = \sum_{i=1}^{N_t} \frac{1}{N_t} \gamma_H b_{iM,t}$. Similarly, define:

$$VWAP_t = \sum_{i=1}^{N_t} w_{i,t} E_t(\eta_{i,t+1}) = F^{VW}_t \sigma^2_{H,t} + G^{VW}_t \sigma_{MH,t}, \tag{A.14}$$

where $F^{VW}_t = \sum_{i=1}^{N_t} w_{i,t} \gamma_H b_{iH,t}$, $G^{VW}_t = \sum_{i=1}^{N_t} w_{i,t} \gamma_H b_{iM,t}$. Similar to the derivation in Corollary 1.1, one can obtain that:

$$\begin{cases} \sigma^2_{H,t} = \frac{G^{EW}_t}{F^{VW}_t G^{EW}_t - F^{EW}_t G^{VW}_t} VWAP_t - \frac{G^{VW}_t}{F^{VW}_t G^{EW}_t - F^{EW}_t G^{VW}_t} EWAP_t \\ \sigma_{MH,t} = \frac{F^{VW}_t}{F^{VW}_t G^{EW}_t - F^{EW}_t G^{VW}_t} EWAP_t - \frac{F^{EW}_t}{F^{VW}_t G^{EW}_t - F^{EW}_t G^{VW}_t} VWAP_t \end{cases} \tag{A.15}$$

This completes the Proof.

D Proof of Proposition 2

The proposition can be simply derived by plug the proposition 2 back into the stock market return equation, since only the covariance terms matters for predicting the market portfolio returns:

$$\mu_{M,t} = \gamma_M \sigma^2_{M,t} + \gamma_H \sigma_{MH,t} = \gamma_M \times \sigma^2_{M,t} + C_t \times \widetilde{EWIV}_t - D_t \times \widetilde{VWIV}_t, \tag{A.16}$$

where $C_t = \frac{\gamma_H A_t^{VW}}{A_t^{VW} B_t^{EW} - A_t^{EW} B_t^{VW}}$, $D_t = \frac{\gamma_H A_t^{EW}}{A_t^{VW} B_t^{EW} - A_t^{EW} B_t^{VW}}$. As to the expected return at the firm/portfolio level, based on (6), (A.5), (A.6), and Corollary 1.1, one can plug back $\sigma_{M,t}^2$ and use EWIV and VWIV to proxy for $\sigma_{H,t}^2$ and $\sigma_{MH,t}$:

$$\mu_{i,t} = \beta_{iM,t}\mu_{M,t} + \beta_{iH,t}(\gamma_H \sigma_{H,t}^2 + \gamma_M \sigma_{MH,t}). \tag{A.17}$$

After some algebra, the expected firm/portfolio return can be written as:

$$\mu_{i,t} = \beta_{iM,t}\mu_{M,t} + C_{i,t} \times \widetilde{EWIV}_t - D_{i,t} \times \widetilde{VWIV}_t, \tag{A.18}$$

where: $C_{i,t} = \beta_{iH,t} \frac{\gamma_M A_t^{VW} - \gamma_H B_t^{VW}}{A_t^{VW} B_t^{EW} - A_t^{EW} B_t^{VW}}$, $D_{i,t} = \beta_{iH,t} \frac{\gamma_M A_t^{EW} - \gamma_H B_t^{EW}}{A_t^{VW} B_t^{EW} - A_t^{EW} B_t^{VW}}$.

This completes the proof.

E Proof of Corollary 2.1

The first two equations ($C_{i,t}$ and $D_{i,t}$) are based on (A.18) directly. I only need to prove the last equation. Based on (D3) in the proof of Proposition 2, it can be easily shown that:

$$\beta_{iH,t} = \frac{A_t^{EW} C_{i,t} - A_t^{VW} D_{i,t}}{\gamma_H} \tag{A.19}$$

Similarly, based on Proposition 2 and the coefficients of (A.16), one can obtain that:

$$\begin{cases} A_t^{VW} = \frac{C_t(A_t^{VW} B_t^{EW} - A_t^{EW} B_t^{VW})}{\gamma_H} \\ A_t^{EW} = \frac{D_t(A_t^{VW} B_t^{EW} - A_t^{EW} B_t^{VW})}{\gamma_H} \end{cases} \tag{A.20}$$

Replace (A.20) into (A.19), the conditional risk exposures can be expressed as:

$$\beta_{iH,t} = \frac{(A_t^{VW} B_t^{EW} - A_t^{EW} B_t^{VW})}{\gamma_H^2}(D_t \times C_{i,t} - C_t \times D_{i,t}). \tag{A.21}$$

Since $(A_t^{VW} B_t^{EW} - A_t^{EW} B_t^{VW})$ is same for any firm at each time t,

$$\beta_{iH,t} \propto (D_t \times C_{i,t} - C_t \times D_{i,t}) \text{ at cross-section.} \tag{A.22}$$

This completes the proof.

F Proof of Proposition 3

Recall that the tail risk measure proposed by Kelly and Jiang (2014) is:

$$\lambda_{t+1}^{Hill} = \frac{1}{K_{t+1}} \sum_{k=1}^{K_{t+1}} \log\left(\frac{\eta_{k,t+1}}{u_{t+1}}\right), \tag{A.23}$$

where $\eta_{k,t+1}$ is the kth daily residual return that falls below an extreme value threshold u_{t+1} during month $t+1$, and K_{t+1} is the total number of such exceedances within month $t+1$. In Kelly and Jiang (2014), to avoid the bias in tail estimates arising from dependence among returns, they mitigate the effect by first removing common return factors and then estimating the tail process from return residuals. Based on (A.1) and (A.3), one can obtain:

$$\begin{cases} \eta_{k,t+1} = E_t(\eta_{k,t+1}) + \varepsilon_{k,t+1} \\ u_{t+1} = E_t(u_{t+1}) + \varepsilon_{u,t+1} \end{cases} \tag{A.24}$$

Proposition 1 gives that:

$$\begin{cases} E_t(\eta_{k,t+1}) = \gamma_H b_{kH,t}\sigma_{H,t}^2 + \gamma_H b_{kM,t}\sigma_{MH,t} \\ E_t(u_{t+1}) = \gamma_H b_{uH,t}\sigma_{H,t}^2 + \gamma_H b_{uM,t}\sigma_{MH,t} \end{cases} \tag{A.25}$$

Following the structure of Corollary 1.1:

$$\begin{cases} E_t(\bar{\eta}_{k,t+1}) = E_t\left(\sum_{k=1}^{K_{t+1}} \frac{1}{K_{t+1}}\eta_{k,t+1}\right) = \sum_{k=1}^{K_{t+1}} \frac{1}{K_{t+1}} E_t(\eta_{k,t+1}) = F_t^K \sigma_{H,t}^2 + G_t^K \sigma_{MH,t} \\ E_t(u_{t+1}) = F_t^u \sigma_{H,t}^2 + G_t^u \sigma_{MH,t} \end{cases}$$

$$\text{where} \begin{cases} F_t^K = \sum_{k=1}^{K_{t+1}} \frac{1}{K_{t+1}}\gamma_H b_{kH,t} \\ F_t^u = \gamma_H b_{uH,t}, \end{cases} \quad \begin{cases} G_t^K = \sum_{k=1}^{K_{t+1}} \frac{1}{K_{t+1}}\gamma_H b_{kM,t} \\ G_t^u = \gamma_H b_{uM,t} \end{cases} . \tag{A.26}$$

Therefore, similar to the derivation Corollary 1.2:

$$\begin{aligned} \sigma_{MH,t} &= \frac{F_t^u}{F_t^{VW}G_t^{EW} - F_t^{EW}G_t^{VW}} E_t(\bar{\eta}_{k,t+1}) - \frac{F_t^K}{F_t^{VW}G_t^{EW} - F_t^{EW}G_t^{VW}} E_t(u_{t+1}) \\ &\equiv J_t^u E_t(\bar{\eta}_{k,t+1}) - J_t^K E_t(u_{t+1}). \end{aligned} \tag{A.27}$$

Based on (A.25),

$$\begin{aligned}\sigma_{MH,t} &= J_t^u E_t(\bar{\eta}_{k,t+1}) - J_t^K E_t(u_{t+1}) \\ &= J_t^u \bar{\eta}_{k,t} - J_t^K u_{t+1} + (J_t^K \varepsilon_{u,t+1} - J_t^u \bar{\varepsilon}_{k,t+1}). \end{aligned} \tag{A.28}$$

I first use the Taylor series to approximate the equation above to logarithm format. Recall that the first order of Taylor series for $\log(x)$ at some $a = x_0$ is:

$$\log(x) \approx \log(x_0) + \frac{1}{x_0}(x - x_0) + \mathcal{O}(x^2) \tag{A.29}$$

The approximation is reasonable when x is close to x_0. In my case, $J_t^u \bar{\eta}_{k,t+1}$ is close to $J_t^K u_{t+1}$ on average. For example, the average $J_t^u \bar{\eta}_{k,t+1}$ within the sample period is around 0.0598, while the average $J_t^K u_{t+1}$ is around 0.0586. And the corresponding standard deviations are 0.0190 and 0.0181. It is reasonable that one can define the following point as x_0:

$$x_0 \equiv \frac{E(J_t^u)E(\bar{\eta}_{k,t+1}) + E(J_t^K)E(u_{t+1})}{2} \tag{A.30}$$

Suppose $x_0 > 0$, then apply the Taylor expansion:

$$\begin{cases} \log(J_t^u \bar{\eta}_{K,t+1}) \approx \log(x_0) + \frac{1}{x_0} J_t^u \bar{\eta}_{k,t+1} - 1 \\ \log(J_t^K u_{t+1}) \approx \log(x_0) + \frac{1}{x_0} J_t^K u_{t+1} - 1 \end{cases}. \tag{A.31}$$

Take the difference between the two:

$$J_t^u \bar{\eta}_{k,t+1} - J_t^K u_{t+1} \approx x_0 \log\left(\frac{J_t^u}{J_t^K}\right) + x_0 \log\left(\frac{\bar{\eta}_{k,t+1}}{u_{t+1}}\right). \tag{A.32}$$

Therefore, the conditional covariance risk can be approximated by:

$$\begin{aligned}\sigma_{MH,t} &= J_t^u E_t(\bar{\eta}_{k,t+1}) - J_t^K E_t(u_{t+1}) \\ &= J_t^u \bar{\eta}_{k,t} - J_t^K u_{t+1} + (J_t^K \varepsilon_{u,t+1} - J_t^u \bar{\varepsilon}_{k,t+1}) \\ &\approx x_0 \log\left(\frac{J_t^u}{J_t^K}\right) + x_0 \log\left(\frac{\bar{\eta}_{k,t+1}}{u_{t+1}}\right) + (J_t^K \varepsilon_{u,t+1} - J_t^u \bar{\varepsilon}_{k,t+1}). \end{aligned} \tag{A.33}$$

The tail risk of (A.23) can also be approximated by similar techniques. Suppose $\frac{\eta_{k,t+1}}{u_{t+1}} > 0$

and $y_{0,t+1} = E\left(\frac{\eta_{k,t+1}}{u_{t+1}}\right)$, the logarithm of the ratio can be approximated by:

$$\log\left(\frac{\eta_{k,t+1}}{u_{t+1}}\right) \approx \log(y_{0,t+1}) + \frac{1}{y_{0,t+1}}\left(\frac{\eta_{k,t+1}}{u_{t+1}} - y_{0,t+1}\right). \tag{A.34}$$

Then the tail risk measure can be approximated by:

$$\begin{aligned}
\lambda_{t+1}^{Hill} &\approx \frac{1}{K_{t+1}}\sum_{k=1}^{K_{t+1}}\left(\log(y_{0,t+1}) + \frac{1}{y_{0,t+1}}\left(\frac{\eta_{k,t+1}}{u_{t+1}} - y_{0,t+1}\right)\right) \\
&= \log(y_{0,t+1}) + \frac{1}{y_{0,t+1}}\left(\frac{\bar{\eta}_{k,t+1}}{u_{t+1}} - y_{0,t+1}\right) \\
&\approx \log\left(\frac{\bar{\eta}_{k,t+1}}{u_{t+1}}\right).
\end{aligned} \tag{A.35}$$

Therefore, the conditional covariance risk based on (A.33) can be expressed as:

$$\begin{aligned}
\sigma_{MH,t} &\approx x_0 \log\left(\frac{J_t^u}{J_t^K}\right) + x_0 \log\left(\frac{\bar{\eta}_{k,t+1}}{u_{t+1}}\right) \\
&\approx x_0 \log\left(\frac{J_t^u}{J_t^K}\right) + x_0 \lambda_{t+1}^{Hill} + (J_t^K \varepsilon_{u,t+1} - J_t^u \bar{\varepsilon}_{k,t+1}).
\end{aligned} \tag{A.36}$$

Since both $\varepsilon_{u,t+1}$ and $\bar{\varepsilon}_{k,t+1}$ follow normal distributions, I can reorganize the formula to:

$$\lambda_{t+1}^{Hill} \approx \log\left(\frac{J_t^K}{J_t^u}\right) + \frac{1}{x_0}\sigma_{MH,t} + \varepsilon_{t+1}. \tag{A.37}$$

where ε_{t+1} follows a normal distribution. This completes the proof. To illustrate accuracy of this approximation, I conduct a numerical study in Section 6 to understand the link.

References

Allen, L., T. G. Bali, and Y. Tang. 2012. Does systemic risk in the financial sector predict future economic downturns? *Review of Financial Studies* 25:3000–36.

Amihud, Y., C. M. Hurvich, and Y. Wang. 2008. Multiple-predictor regressions: Hypothesis testing. *Review of Financial Studies* 22:413–34.

Andersen, T. G., T. Bollerslev, F. X. Diebold, and P. Labys. 2003. Modeling and forecasting realized volatility. *Econometrica* 71:579–625.

Ang, A., and G. Bekaert. 2007. Stock return predictability: Is it there? *Review of Financial Studies* 20:651–707.

Ang, A., R. J. Hodrick, Y. Xing, and X. Zhang. 2006. The cross-section of volatility and expected returns. *Journal of Finance* 61:259–99.

Baker, M., R. Greenwood, and J. Wurgler. 2003. The maturity of debt issues and predictable variation in bond returns. *Journal of Financial Economics* 70:261–91.

Baker, M., and J. Wurgler. 2007. Investor sentiment in the stock market. *Journal of Economic Perspectives* 21:129–52.

Bali, T. G., N. Cakici, X. Yan, and Z. Zhang. 2005. Does idiosyncratic risk really matter? *Journal of Finance* 60:905–29.

Bali, T. G., and R. F. Engle. 2010. The intertemporal capital asset pricing model with dynamic conditional correlations. *Journal of Monetary Economics* 57:377–90.

Bali, T. G., and L. Peng. 2006. Is there a risk–return trade-off? Evidence from high-frequency data. *Journal of Applied Econometrics* 21:1169–98.

Bansal, R., and I. Shaliastovich. 2011. Learning and asset-price jumps. *Review of Financial Studies* 24:2738–80.

Bansal, R., and A. Yaron. 2004. Risks for the long run: A potential resolution of asset pricing puzzles. *Journal of Finance* 59:1481–509.

Barro, R. J. 2006. Rare disasters and asset markets in the twentieth century. *Quarterly Journal of Economics* 121:823–66.

Bartram, S. M., G. W. Brown, and R. M. Stulz. 2018. Why Has Idiosyncratic Risk Been Historically Low in Recent Years? Working Paper, National Bureau of Economic Research.

Bekaert, G., R. J. Hodrick, and X. Zhang. 2012. Aggregate idiosyncratic volatility. *Journal of Financial and Quantitative Analysis* 47:1155–85.

Bloom, N. 2014. Fluctuations in uncertainty. *Journal of Economic Perspectives* 28:153–76.

Bollerslev, T., G. Tauchen, and H. Zhou. 2009. Expected stock returns and variance risk premia. *Review of Financial Studies* 22:4463–92.

Brandt, M. W., A. Brav, J. R. Graham, and A. Kumar. 2010. The idiosyncratic volatility puzzle: Time trend or speculative episodes? *Review of Financial Studies* 23:863–99.

Campbell, J. Y., M. Lettau, B. G. Malkiel, and Y. Xu. 2001. Have individual stocks become more volatile? An empirical exploration of idiosyncratic risk. *Journal of Finance* 56:1–43.

Campbell, J. Y., and G. B. Taksler. 2003. Equity volatility and corporate bond yields. *Journal of Finance* 58:2321–50.

Campbell, J. Y., and S. B. Thompson. 2008. Predicting excess stock returns out of sample: Can anything beat the historical average? *Review of Financial Studies* 21:1509–31.

Cao, J., and B. Han. 2016. Idiosyncratic risk, costly arbitrage, and the cross-section of stock returns. *Journal of Banking & Finance* 73:1–15.

Chapman, D. A., M. F. Gallmeyer, and J. S. Martin. 2018. Aggregate tail risk and expected returns. *Review of Asset Pricing Studies* 8:36–76.

Chen, L. 2009. On the reversal of return and dividend growth predictability: A tale of two periods. *Journal of Financial Economics* 92:128–51.

Chen, Y., G. W. Eaton, and B. S. Paye. 2018. Micro (structure) before macro? The predictive power of aggregate illiquidity for stock returns and economic activity. *Journal of Financial Economics* 130:48–73.

Chen, Z., and R. Petkova. 2012. Does idiosyncratic volatility proxy for risk exposure? *Review of Financial Studies* 25:2745–87.

Clark, T. E., and K. D. West. 2007. Approximately normal tests for equal predictive accuracy in nested models. *Journal of Econometrics* 138:291–311.

Cox, J. C., J. E. Ingersoll Jr, and S. A. Ross. 1985. An intertemporal general equilibrium model of asset prices. *Econometrica: Journal of the Econometric Society* 363–84.

Drechsler, I., and A. Yaron. 2011. What's vol got to do with it. *Review of Financial Studies* 24:1–45.

Engle, R. F., and V. K. Ng. 1993. Measuring and testing the impact of news on volatility. *Journal of Finance* 48:1749–78.

Evans, M. D. 1994. Expected returns, time-varying risk, and risk premia. *Journal of Finance* 49:655–79.

Fama, E. F., and J. D. MacBeth. 1973. Risk, return, and equilibrium: Empirical tests. *Journal of Political Economy* 81:607–36.

Ferreira, M. A., and P. Santa-Clara. 2011. Forecasting stock market returns: The sum of the parts is more than the whole. *Journal of Financial Economics* 100:514–37.

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

Ferson, W. E., S. Kandel, and R. F. Stambaugh. 1987. Tests of asset pricing with time-varying expected risk premiums and market betas. *Journal of Finance* 42:201–20.

French, K. R., G. W. Schwert, and R. F. Stambaugh. 1987. Expected stock returns and volatility. *Journal of Financial Economics* 19:3–.

Fu, F. 2009. Idiosyncratic risk and the cross-section of expected stock returns. *Journal of Financial Economics* 91:24–37.

Gabaix, X. 2012. Variable rare disasters: An exactly solved framework for ten puzzles in macro-finance. *Quarterly Journal of Economics* 127:645–700.

Garcia, R., D. Mantilla-Garcia, and L. Martellini. 2014. A model-free measure of aggregate idiosyncratic volatility and the prediction of market returns. *Journal of Financial and Quantitative Analysis* 49:1133–65.

Ghysels, E., P. Santa-Clara, and R. Valkanov. 2005. There is a risk-return trade-off after all. *Journal of Financial Economics* 76:509–48.

Giglio, S., and D. Xiu. 2018. Asset pricing with omitted factors. *Chicago Booth Research Paper* .

Glosten, L. R., R. Jagannathan, and D. E. Runkle. 1993. On the relation between the expected value and the volatility of the nominal excess return on stocks. *Journal of Finance* 48:1779–801.

Goetzmann, W. N., R. G. Ibbotson, and L. Peng. 2001. A new historical database for the NYSE 1815 to 1925: Performance and predictability. *Journal of Financial Markets* 4:1–32.

Golez, B., and P. Koudijs. 2018. Four centuries of return predictability. *Journal of Financial Economics* 127:248–63.

Goyal, A., and P. Santa-Clara. 2003. Idiosyncratic risk matters! *Journal of Finance* 58:975–1007.

Goyal, A., and I. Welch. 2008. A comprehensive look at the empirical performance of equity premium prediction. *Review of Financial Studies* 21:1455–508.

Guo, H., and R. Savickas. 2008. Average idiosyncratic volatility in G7 countries. *Review of Financial Studies* 21:1259–96.

Guo, H., and R. F. Whitelaw. 2006. Uncovering the risk–return relation in the stock market. *Journal of Finance* 61:1433–63.

Han, B., and G. Li. 2017. Aggregate implied volatility spread and stock market returns. *Working Paper* .

Hansen, L. P., and S. F. Richard. 1987. The role of conditioning information in deducing testable restrictions implied by dynamic asset pricing models. *Econometrica: Journal of the Econometric Society* 587–613.

Harvey, C. R. 2001. The specification of conditional expectations. *Journal of Empirical Finance* 8:573–637.

Hedegaard, E., and R. J. Hodrick. 2014. Measuring the Risk-Return Tradeoff with Time-Varying Conditional Covariances. Working Paper, National Bureau of Economic Research.

Herskovic, B., B. Kelly, H. Lustig, and S. Van Nieuwerburgh. 2016. The common factor in idiosyncratic volatility: Quantitative asset pricing implications. *Journal of Financial Economics* 119:249–83.

Ingersoll, J. E. 1987. *Theory of financial decision making*, vol. 3. Rowman & Littlefield.

Jagannathan, R., and Z. Wang. 1996. The conditional CAPM and the cross-section of expected returns. *Journal of Finance* 51:3–53.

Jurado, K., S. C. Ludvigson, and S. Ng. 2015. Measuring uncertainty. *American Economic Review* 105:1177–216.

Kandel, S., and R. F. Stambaugh. 1996. On the predictability of stock returns: an asset-allocation perspective. *Journal of Finance* 51:385–424.

Kelly, B., and H. Jiang. 2014. Tail risk and asset prices. *Review of Financial Studies* 27:2841–71.

Kelly, B., and S. Pruitt. 2013. Market expectations in the cross-section of present values. *Journal of Finance* 68:1721–56.

———. 2015. The three-pass regression filter: A new approach to forecasting using many predictors. *Journal of Econometrics* 186:294–316.

Lettau, M., and S. Ludvigson. 2001a. Consumption, aggregate wealth, and expected stock returns. *Journal of Finance* 56:815–49.

———. 2001b. Resurrecting the (C) CAPM: A cross-sectional test when risk premia are time-varying. *Journal of Political Economy* 109:1238–87.

Levy, H. 1978. Equilibrium in an Imperfect Market: A Constraint on the Number of Securities in the Portfolio. *American Economic Review* 68:643–58.

Light, N., D. Maslov, and O. Rytchkov. 2017. Aggregation of information about the cross section of stock returns: A latent variable approach. *Review of Financial Studies* 30:1339–81.

Lo, A. W., and A. C. MacKinlay. 1988. Stock market prices do not follow random walks: Evidence from a simple specification test. *Review of Financial Studies* 1:41–66.

Lo, A. W., and J. Wang. 2006. Trading volume: Implications of an intertemporal capital asset pricing model. *Journal of Finance* 61:2805–40.

Long Jr, J. B. 1974. Stock prices, inflation, and the term structure of interest rates. *Journal of Financial Economics* 1:131–70.

Ludvigson, S. C., and S. Ng. 2007. The empirical risk–return relation: A factor analysis approach. *Journal of Financial Economics* 83:171–222.

Malkiel, B. G., and Y. Xu. 2002. Idiosyncratic risk and security returns. *University of Texas at Dallas (November 2002)* .

Martin, I. 2017. What is the Expected Return on the Market? *Quarterly Journal of Economics* 132:367–433.

Menzly, L., T. Santos, and P. Veronesi. 2004. Understanding predictability. *Journal of Political Economy* 112:1–47.

Merton, R. C. 1973. An intertemporal capital asset pricing model. *Econometrica: Journal of the Econometric Society* 867–87.

Nelson, D. B. 1991. Conditional heteroskedasticity in asset returns: A new approach. *Econometrica: Journal of the Econometric Society* 347–70.

Newey, W. K., and K. D. West. 1987. A simple, positive semi-definite, heteroskedasticity and autocorrelationconsistent covariance matrix.

Pollet, J. M., and M. Wilson. 2010. Average correlation and stock market returns. *Journal of Financial Economics* 96:364–80.

Rapach, D. E., M. C. Ringgenberg, and G. Zhou. 2016. Short interest and aggregate stock returns. *Journal of Financial Economics* 121:46–65.

Rapach, D. E., J. K. Strauss, and G. Zhou. 2010. Out-of-sample equity premium prediction: Combination forecasts and links to the real economy. *Review of Financial Studies* 23:821–62.

Rossi, A. G., and A. Timmermann. 2015. Modeling covariance risk in Merton's ICAPM. *Review of Financial Studies* 28:1428–61.

Santos, T., and P. Veronesi. 2006. Labor income and predictable stock returns. *Review of Financial Studies* 19:1–44.

Schwert, G. W. 1990. Indexes of United States stock prices from 1802 to 1987. Working Paper 3, National Bureau of Economic Research.

Scruggs, J. T. 1998. Resolving the puzzling intertemporal relation between the market risk premium and conditional market variance: A two-factor approach. *Journal of Finance* 53:575–603.

Shumway, T. 1997. The delisting bias in CRSP data. *Journal of Finance* 52:327–40.

Van Binsbergen, J. H., and R. S. Koijen. 2010. Predictive regressions: A present-value approach. *Journal of Finance* 65:1439–71.

Wachter, J. A. 2013. Can time-varying risk of rare disasters explain aggregate stock market volatility? *Journal of Finance* 68:987–1035.

Wei, S. X., and C. Zhang. 2005. Idiosyncratic risk does not matter: A re-examination of the relationship between average returns and average volatilities. *Journal of Banking & Finance* 29:603–21.

Table 1
Summary Statistics of Monthly EWIV, VWIV, SMV, and MKTRF

			Autocorrelation at Lag (Number of Months)					Pearson Correlation			
Variable	**Mean**	**STD**	**1**	**3**	**6**	**12**	**24**	**EWIV**	**VWIV**	**SMV**	**MKTRF**
Panel A. Time Period: 181501 to 187012											
EWIV	0.049	0.024	0.975	0.952	0.925	0.880	0.778	1			
VWIV	0.040	0.018	0.893	0.835	0.752	0.679	0.560	0.853	1		
SMV	0.030	0.013	0.995	0.981	0.955	0.883	0.705	0.566	0.535	1	
MKTRF	0.004	0.033	0.058	0.051	0.123	0.040	-0.022	0.129	0.082	0.069	1
Panel B. Time Period: 187101 to 192512											
EWIV	0.064	0.011	0.949	0.897	0.845	0.760	0.569	1			
VWIV	0.050	0.011	0.949	0.904	0.842	0.737	0.547	0.826	1		
SMV	0.032	0.005	0.987	0.950	0.884	0.734	0.489	0.529	0.307	1	
MKTRF	0.002	0.033	0.285	-0.001	0.040	-0.045	-0.031	0.116	0.038	0.073	1
Panel C. Time Period: 192601 to 196212											
EWIV	0.026	0.016	0.948	0.856	0.760	0.565	0.470	1			
VWIV	0.021	0.012	0.935	0.808	0.654	0.401	0.360	0.959	1		
SMV	0.065	0.046	0.886	0.646	0.647	0.578	0.383	0.626	0.514	1	
MKTRF	0.008	0.065	0.135	-0.166	-0.031	-0.012	0.035	0.015	0.017	-0.111	1
Panel D. Time Period: 196301 to 201812											
EWIV	0.027	0.008	0.970	0.876	0.777	0.693	0.520	1			
VWIV	0.020	0.005	0.948	0.850	0.727	0.567	0.277	0.893	1		
SMV	0.038	0.023	0.672	0.480	0.357	0.234	0.054	0.443	0.457	1	
MKTRF	0.005	0.044	0.072	0.022	-0.051	0.028	-0.012	-0.016	-0.045	-0.299	1

The table reports the descriptive statistics for the monthly time-series EWIV, VWIV, SMV, and MKTRF, as well as their correlations. The sample periods are separated based on data resources and quality. The data before 192601 is collected from the Global Financial Data. The data after 192601 is collected from CRSP. Detailed descriptions of the variables and their constructions are provided in Section 3.

Table 2
Univariate and Bivariate Monthly Regression

Panel A. Univariate Regression						
196301 to 201812	**Coefficient**	**K=1**	**K=3**	**K=6**	**K=12**	**K=24**
EWIV	b	0.000	0.001	0.001	0.001	0.001
	t-stat	(0.21)	(0.34)	(0.50)	(0.34)	(0.39)
	adj.R^2(%)	-0.14	-0.11	-0.03	-0.01	0.38
VWIV	b	-0.003	-0.002	-0.002	-0.002	-0.001
	t-stat	(-1.32)	(-1.45)	(-1.28)	(-1.18)	(-0.77)
	adj.R^2(%)	0.24	0.64	0.81	1.69	1.88
Panel B. Bivariate Regression						
196301 to 201812	**Coefficient**	**K=1**	**K=3**	**K=6**	**K=12**	**K=24**
EWIV	b	0.014	0.013	0.012	0.013	0.011
	t-stat	(2.96)	(3.57)	(3.58)	(4.07)	(4.54)
	mARM *t*-stat	(2.11)	(2.72)	(2.43)	(3.97)	(3.99)
VWIV	b	-0.015	-0.014	-0.013	-0.014	-0.012
	t-stat	(-3.03)	(-3.74)	(-3.68)	(-4.52)	(-5.13)
	mARM *t*-stat	(-2.62)	(-3.61)	(-3.66)	(-3.89)	(-5.09)
	adj.R^2(%)	1.29	3.48	6.10	13.93	23.30
Panel C. Bivariate Regression with Different Sample Periods						
181501 to 187012	**Coefficient**	**K=1**	**K=3**	**K=6**	**K=12**	**K=24**
EWIV	b	0.009	0.009	0.010	0.010	0.008
	t-stat	(3.23)	(4.69)	(4.71)	(4.03)	(2.99)
VWIV	b	-0.006	-0.007	-0.007	-0.008	-0.007
	t-stat	(-2.13)	(-3.84)	(-4.20)	(-3.98)	(-2.81)
	adj.R^2(%)	2.16	5.88	11.16	16.41	20.42
187101 to 192512	**Coefficient**	**K=1**	**K=3**	**K=6**	**K=12**	**K=24**
EWIV	b	0.005	0.006	0.007	0.006	0.004
	t-stat	(2.18)	(2.54)	(3.44)	(3.15)	(3.49)
VWIV	b	-0.004	-0.005	-0.006	-0.006	-0.005
	t-stat	(-1.79)	(-2.31)	(-3.43)	(-3.44)	(-3.49)
	adj.R^2(%)	0.49	1.68	5.31	8.12	9.09

Table 2 Continued:

192601 to 196212	Coefficient	K=1	K=3	K=6	K=12	K=24
EWIV	b	0.029	0.043	0.043	0.038	0.033
	t-stat	(1.26)	(2.15)	(2.62)	(3.42)	(3.74)
VWIV	b	-0.027	-0.043	-0.044	-0.038	-0.034
	t-stat	(-1.32)	(-2.46)	(-2.84)	(-3.50)	(-3.35)
	adj.R^2(%)	0.48	4.35	10.36	13.99	22.90
196301 to 200112	**Coefficient**	**K=1**	**K=3**	**K=6**	**K=12**	**K=24**
EWIV	b	0.009	0.007	0.006	0.006	0.004
	t-stat	(3.14)	(3.23)	(2.89)	(2.91)	(2.49)
VWIV	b	-0.007	-0.006	-0.006	-0.006	-0.003
	t-stat	(-2.12)	(-2.53)	(-2.36)	(-2.55)	(-1.39)
	adj.R^2(%)	1.29	3.13	4.83	10.16	11.38
200201 to 201812	**Coefficient**	**K=1**	**K=3**	**K=6**	**K=12**	**K=24**
EWIV	b	0.023	0.031	0.032	0.022	0.014
	t-stat	(1.63)	(2.38)	(2.42)	(2.04)	(2.46)
VWIV	b	-0.030	-0.036	-0.034	-0.021	-0.012
	t-stat	(-2.16)	(-2.82)	(-2.65)	(-1.96)	(-2.48)
	adj.R^2(%)	4.42	14.48	17.61	15.19	19.17
181501 to 201812	**Coefficient**	**K=1**	**K=3**	**K=6**	**K=12**	**K=24**
EWIV	b	0.008	0.009	0.009	0.009	0.007
	t-stat	(3.37)	(4.26)	(4.26)	(3.92)	(3.24)
VWIV	b	-0.009	-0.010	-0.010	-0.010	-0.008
	t-stat	(-3.75)	(-5.02)	(-4.91)	(-4.55)	(-3.73)
	adj.R^2(%)	0.47	1.65	3.28	5.56	7.21

This table reports the results of univariate and bivariate monthly predictive regressions. The dependent variable is the average monthly value-weighted market excess returns in logarithm (MKTRF) over the relevant forecast horizon. All predictors are normalized to have zero mean and one standard deviation. K stands for the forecast horizon in number of months. "b" is the slope coefficient on the predictor. When $K > 1$, to adjust for the overlapping dependent variable, the t-stat is computed using the GMM standard errors with $K - 1$ Newey-West lag correction. The sample periods are specified in the table.

Table 3
Univariate and Bivariate Daily Regression

Panel A. Univariate Regression						
19630103 to 20181231	**Coefficient**	**D=1**	**D=3**	**D=6**	**D=12**	**D=24**
EWIV	b	0.001	0.001	0.001	0.001	0.001
	t-stat	(0.36)	(0.45)	(0.48)	(0.49)	(0.52)
	adj.R^2(%)	0.26	0.02	0.00	0.00	0.02
VWIV	b	-0.002	-0.002	-0.002	-0.002	-0.002
	t-stat	(-0.61)	(-0.83)	(-0.91)	(-1.01)	(-1.07)
	adj.R^2(%)	0.27	0.04	0.03	0.08	0.17
Panel B. Bivariate Regression						
19630103 to 20181231	**Coefficient**	**D=1**	**D=3**	**D=6**	**D=12**	**D=24**
EWIV	b	0.017	0.018	0.018	0.017	0.017
	t-stat	(2.97)	(3.53)	(3.61)	(3.67)	(3.74)
	mARM t-stat	(1.97)	(3.20)	(2.76)	(2.68)	(2.41)
VWIV	b	-0.018	-0.018	-0.018	-0.018	-0.017
	t-stat	(-2.78)	(-3.36)	(-3.54)	(-3.74)	(-3.86)
	mARM t-stat	(-1.96)	(-3.12)	(-3.24)	(-3.43)	(-2.95)
	adj.R^2(%)	0.34	0.25	0.46	0.94	1.80
Panel C. Bivariate Regression with Different Sample Periods						
19280303 to 19621231	**Coefficient**	**D=1**	**D=3**	**D=6**	**D=12**	**D=24**
EWIV	b	0.024	0.024	0.024	0.028	0.034
	t-stat	(1.06)	(1.23)	(1.27)	(1.44)	(1.62)
VWIV	b	-0.023	-0.024	-0.023	-0.027	-0.035
	t-stat	(-1.07)	(-1.24)	(-1.30)	(-1.55)	(-1.84)
	adj.R^2(%)	0.67	0.08	0.20	0.58	1.27

Table 3 Continued:

19630103 to 20011231	**Coefficient**	**D=1**	**D=3**	**D=6**	**D=12**	**D=24**
EWIV	b	0.014	0.015	0.015	0.015	0.015
	t-stat	(2.71)	(3.32)	(3.42)	(3.49)	(3.50)
VWIV	b	-0.012	-0.013	-0.013	-0.013	-0.012
	t-stat	(-1.95)	(-2.47)	(-2.65)	(-2.87)	(-2.98)
	adj.R^2(%)	2.29	0.68	0.69	1.08	1.91
20020103 to 20181231	**Coefficient**	**D=1**	**D=3**	**D=6**	**D=12**	**D=24**
EWIV	b	0.046	0.043	0.042	0.041	0.040
	t-stat	(1.72)	(2.16)	(2.22)	(2.24)	(2.40)
VWIV	b	-0.052	-0.049	-0.048	-0.046	-0.046
	t-stat	(-1.77)	(-2.32)	(-2.47)	(-2.55)	(-2.73)
	adj.R^2(%)	0.73	1.04	2.02	3.29	6.34
19280303 to 20181231	**Coefficient**	**D=1**	**D=3**	**D=6**	**D=12**	**D=24**
EWIV	b	0.020	0.021	0.020	0.020	0.021
	t-stat	(2.43)	(2.88)	(3.00)	(3.10)	(3.11)
VWIV	b	-0.019	-0.020	-0.020	-0.020	-0.021
	t-stat	(-2.39)	(-2.86)	(-3.03)	(-3.28)	(-3.59)
	adj.R^2(%)	0.49	0.15	0.27	0.57	1.12

This table reports the results of univariate and bivariate daily predictive regressions. The dependent variable is the average daily value-weighted market excess returns in logarithm (MKTRF) over the relevant forecast horizon. All predictors are normalized to have zero mean and one standard deviation. D stands for the forecast horizon in number of days. I also include the lagged stock market return as a regressor (whose coefficient is omitted for brevity) to control for the market return autocorrelation. "b" is the slope coefficient on the predictor. When $D > 1$, to adjust for the overlapping dependent variable, the t-stat is computed using the GMM standard errors with $D-1$ Newey-West lag correction. The sample periods are specified in the table.

Table 4
Multiple Predictive Monthly Regression

	K=1							K=12						
	Control		**EWIV**		**VWIV**			**Control**		**EWIV**		**VWIV**		
Predictor	**b**	***t*-stat**	**b**	***t*-stat**	**b**	***t*-stat**	**adj.R^2(%)**	**b**	***t*-stat**	**b**	***t*-stat**	**b**	***t*-stat**	**adj.R^2(%)**
Log dividend-price ratio	0.002	(0.80)	0.015	(3.11)	-0.016	(-3.12)	1.38	0.002	(1.34)	0.014	(4.18)	-0.014	(-4.43)	16.37
Log dividend yield	0.002	(0.95)	0.015	(3.11)	-0.016	(-3.11)	1.42	0.002	(1.33)	0.014	(4.17)	-0.014	(-4.42)	16.35
Dividend-payout ratio	-0.000	(-0.08)	0.015	(3.10)	-0.016	(-3.20)	1.26	0.000	(0.35)	0.013	(3.87)	-0.014	(-4.45)	14.68
Book-to-market ratio	0.001	(0.68)	0.016	(3.20)	-0.016	(-3.25)	1.35	0.002	(1.27)	0.015	(4.52)	-0.015	(-4.82)	16.25
Net equity expansion	-0.000	(-0.22)	0.015	(3.02)	-0.016	(-3.00)	1.27	0.000	(-0.16)	0.013	(3.95)	-0.014	(-4.28)	14.63
Treasury bill rate	-0.002	(-0.90)	0.014	(2.90)	-0.015	(-2.86)	1.40	-0.001	(-0.41)	0.013	(4.10)	-0.014	(-4.39)	14.76
Long-term yield	-0.002	(-0.89)	0.015	(3.11)	-0.016	(-3.15)	1.40	-0.000	(-0.11)	0.014	(4.12)	-0.014	(-4.57)	14.60
Long-term return	0.004	(2.20)	0.014	(2.86)	-0.015	(-3.03)	2.17	0.001	(2.48)	0.013	(4.04)	-0.014	(-4.50)	15.24
Term spread	0.004	(2.08)	0.012	(2.57)	-0.013	(-2.59)	2.14	0.001	(1.23)	0.013	(3.99)	-0.014	(-4.34)	15.34
Default yield spread	0.002	(0.96)	0.014	(3.03)	-0.015	(-3.13)	1.47	0.002	(1.70)	0.013	(4.00)	-0.014	(-4.52)	16.77
Default return spread	0.005	(2.49)	0.014	(2.91)	-0.014	(-2.91)	2.69	0.001	(1.18)	0.013	(4.14)	-0.014	(-4.51)	15.13
Inflation	-0.002	(-1.12)	0.014	(2.79)	-0.015	(-2.92)	1.51	-0.001	(-1.27)	0.013	(4.15)	-0.014	(-4.60)	15.41
Average correlation	0.003	(1.30)	0.014	(2.89)	-0.015	(-2.90)	1.64	0.002	(2.01)	0.013	(3.82)	-0.013	(-4.14)	16.91
SMV	-0.000	(-0.11)	0.008	(2.85)	-0.009	(-2.33)	1.08	0.004	(3.91)	0.007	(3.38)	-0.010	(-4.85)	13.65
ILIQ	0.002	(0.75)	0.010	(1.35)	-0.012	(-1.66)	1.19	0.002	(1.37)	0.009	(2.10)	-0.010	(-2.68)	14.74
VRP	0.010	(3.68)	0.007	(0.91)	-0.011	(-1.32)	6.02	0.001	(1.06)	0.016	(3.26)	-0.018	(-4.14)	20.39
IVS	0.011	(2.85)	0.016	(1.12)	-0.018	(-1.30)	5.77	0.001	(0.90)	0.026	(2.85)	-0.028	(-3.54)	23.97
SVIX	0.002	(0.45)	0.014	(1.75)	-0.018	(-2.01)	1.00	0.003	(3.18)	0.017	(3.39)	-0.020	(-4.39)	26.03
CAY	-0.004	(-1.93)	0.022	(3.47)	-0.021	(-3.57)	1.81	0.001	(0.60)	0.012	(3.08)	-0.013	(-3.90)	14.88
Sentiment	-0.003	(-1.23)	0.015	(3.02)	-0.015	(-2.89)	1.42	-0.002	(-1.10)	0.014	(4.26)	-0.014	(-4.49)	15.90
SII	-0.006	(-3.08)	0.021	(3.36)	-0.021	(-3.20)	2.84	-0.006	(-4.10)	0.019	(5.04)	-0.019	(-5.52)	31.77

This table reports the results of multivariate predictive regressions with both EWIV and VWIV and each control predictor. The dependent variable is monthly market excess returns (MKTRF) and the forecast horizon is one month ($K = 1$) and one year ($K = 12$). I normalize all predictors to have zero mean and one standard deviation. When $K > 1$, to adjust for the overlapping dependent variable, the t-stat is computed using the GMM standard errors with $K - 1$ Newey-West correction. The sample period is from 196301 to 201812.

Table 5
Multiple Predictive Daily Regression

	Panel A. Control for SMV, SMB, and HML					
19630103 to 20181231	**Coefficient**	**D=1**	**D=3**	**D=6**	**D=12**	**D=24**
EWIV	b	0.017	0.017	0.017	0.017	0.017
	t-stat	(2.92)	(3.49)	(3.57)	(3.63)	(3.71)
VWIV	b	-0.019	-0.019	-0.019	-0.019	-0.019
	t-stat	(-2.92)	(-3.56)	(-3.73)	(-3.95)	(-4.09)
SMV	b	0.003	0.003	0.003	0.003	0.003
	t-stat	(0.60)	(0.86)	(1.01)	(1.26)	(1.55)
SMB	b	-0.005	-0.003	-0.001	0.000	0.000
	t-stat	(-1.23)	(-1.53)	(-1.06)	(0.46)	(0.62)
HML	b	-0.015	-0.006	-0.004	-0.002	-0.001
	t-stat	(-3.99)	(-3.42)	(-3.27)	(-2.14)	(-1.70)
Total	adj.R^2(%)	0.84	0.52	0.72	1.18	2.21
	Panel B. Control for SMV, VRP, IVS, SMB, and HML					
19960103 to 20181231	**Coefficient**	**D=1**	**D=3**	**D=6**	**D=12**	**D=24**
EWIV	b	0.050	0.064	0.072	0.072	0.072
	t-stat	(1.01)	(1.82)	(2.24)	(2.51)	(2.87)
VWIV	b	-0.063	-0.075	-0.083	-0.084	-0.083
	t-stat	(-1.48)	(-2.53)	(-3.09)	(-3.44)	(-3.63)
SMV	b	-0.032	-0.017	-0.003	0.011	0.011
	t-stat	(-1.02)	(-0.82)	(-0.13)	(0.63)	(1.23)
SMB	b	0.035	0.011	0.007	0.006	0.003
	t-stat	(2.94)	(2.17)	(1.78)	(2.27)	(1.74)
HML	b	0.013	0.000	0.000	0.002	0.001
	t-stat	(0.98)	(0.09)	(0.10)	(0.73)	(0.84)
VRP	b	0.055	0.029	0.012	0.000	-0.001
	t-stat	(1.63)	(1.43)	(0.67)	(0.03)	(-0.14)
IVS	b	0.052	0.025	0.015	0.016	0.012
	t-stat	(2.96)	(2.37)	(1.73)	(2.42)	(2.65)
	adj.R^2(%)	3.55	3.87	5.11	9.30	16.13

This table reports the results of multiple daily predictive regressions. Each column in this table corresponds to one multiple predictive regression, labeled by the forecast horizons (D=days). The dependent variable is the average daily value-weighted market excess returns (MKTRF) over the relevant forecast horizon, and all predictors are normalized to have zero mean and one standard deviation. For all the regressions, I include the lagged stock market return as a regressor (whose coefficient is omitted for brevity) to control for the market return autocorrelation. "b" is the slope coefficient on the predictor. When $D > 1$, to adjust for the overlapping dependent variable, the t-stat is computed using the GMM standard errors with $D - 1$ Newey-West lag correction. The sample periods are specified in each panel.

Table 6
Out-of-Sample Performance of the Combination of EWIV and VWIV

Panel A. Out-of-Sample R^2 Statistics (Monthly)						
Predictor	**Statistic**	**K=1**	**K=3**	**K=6**	**K=12**	**K=24**
EWIV	OOS R^2(%)	-0.808	-2.193	-2.974	-6.671	-19.067
198101 to 201812	z-stat	(-0.77)	(-0.95)	(-0.71)	(-0.01)	(-0.36)
VWIV	OOS R^2(%)	-0.431	-1.278	-2.895	-8.682	-30.728
198101 to 201812	z-stat	(-0.24)	(-0.44)	(-0.24)	(-0.83)	(-1.33)
EWIV+VWIV	OOS R^2(%)	1.015	2.363	5.238	7.888	8.985
186001 to 192512	z-stat	(2.35)	(2.52)	(2.72)	(2.49)	(1.97)
EWIV+VWIV	OOS R^2(%)	0.640	2.217	5.459	11.899	20.579
198101 to 201812	z-stat	(2.14)	(2.52)	(2.95)	(3.45)	(5.00)
Panel B. Out-of-Sample R^2 Statistics (Daily)						
Predictor	**Statistic**	**D=1**	**D=3**	**D=6**	**D=12**	**D=24**
EWIV	OOS R^2(%)	-0.061	-0.167	-0.296	-0.610	-1.199
19810103 to 20181231	z-stat	(-0.76)	(-0.92)	(-1.18)	(-1.29)	(-1.38)
VWIV	OOS R^2(%)	-0.048	-0.121	-0.189	-0.363	-0.733
19810103 to 20181231	z-stat	(-0.08)	(-0.12)	(-0.36)	(-0.48)	(-0.62)
EWIV+VWIV	OOS R^2(%)	0.055	0.154	0.316	0.654	1.346
19810103 to 20181231	z-stat	(2.70)	(3.01)	(3.10)	(3.12)	(3.12)
Panel C: Optimal Portfolio Sharpe Ratio and CER Gain (Monthly)						
Predictor	**Statistic**	**K=1**	**K=3**	**K=6**	**K=12**	**K=24**
EWIV	Sharpe Ratio	0.378	0.386	0.392	0.395	0.351
198101 to 201812	CER Gain (%)	-0.770	-0.115	0.467	0.482	0.077
VWIV	Sharpe Ratio	0.396	0.397	0.361	0.355	0.271
198101 to 201812	CER Gain (%)	-0.090	0.250	0.196	0.094	-0.618
EWIV+VWIV	Sharpe Ratio	0.642	0.601	0.591	0.564	0.466
186001 to 192512	CER Gain (%)	1.488	1.910	1.637	0.857	1.528
EWIV+VWIV	Sharpe Ratio	0.497	0.509	0.514	0.553	0.414
198101 to 201812	CER Gain (%)	1.358	1.928	2.431	3.272	1.325
Historical Average	Sharpe Ratio	0.401	0.377	0.345	0.350	0.299
	CER Gain (%)	-	-	-	-	-

The forecast target is the market excess returns (MKTRF). The construction process is specified in section 3. The z-stat is computed based on Clark and West (2007), I reject the null hypothesis if this z-stat is greater than 1.282 (for a one-sided test at 10% confidence), 1.645 (for a one-sided test at 5% confidence), or 2.334 (for a one-sided test at 1% confidence). The out-of-sample evaluation periods are specified in the table.

Table 7
Predictive Regression based on Alternative Weighting Schemes and Variables

Panel A. In-Sample Performance

Predictor	**Coefficient**	**K=1**	**K=3**	**K=6**	**K=12**	**K=24**
EWIV	b	0.012	0.011	0.011	0.011	0.010
	t-stat	(2.64)	(3.25)	(3.35)	(3.98)	(4.40)
PWIV	b	-0.012	-0.011	-0.011	-0.012	-0.010
	t-stat	(-2.70)	(-3.47)	(-3.40)	(-4.43)	(-5.00)
	adj.R^2(%)	1.01	2.87	5.34	12.83	21.60
Predictor	**Coefficient**	**K=1**	**K=3**	**K=6**	**K=12**	**K=24**
EWIV	b	-0.006	-0.006	-0.006	-0.006	-0.005
	t-stat	(-2.07)	(-2.54)	(-2.57)	(-2.60)	(-1.94)
IWIV	b	0.008	0.008	0.008	0.008	0.006
	t-stat	(2.64)	(3.62)	(3.50)	(3.86)	(3.48)
	adj.R^2(%)	0.73	2.50	4.38	8.77	11.92
Predictor	**Coefficient**	**K=1**	**K=3**	**K=6**	**K=12**	**K=24**
SWIV	b	0.006	0.005	0.005	0.006	0.005
	t-stat	(2.45)	(3.11)	(3.20)	(3.51)	(3.74)
BWIV	b	-0.006	-0.006	-0.005	-0.006	-0.005
	t-stat	(-2.44)	(-3.41)	(-3.29)	(-4.70)	(-5.28)
	adj.R^2(%)	0.99	3.12	5.58	13.75	23.95
Predictor	**Coefficient**	**K=1**	**K=3**	**K=6**	**K=12**	**K=24**
EWAP	b	0.014	0.014	0.013	0.014	0.012
	t-stat	(2.44)	(3.18)	(3.16)	(3.84)	(4.52)
VWAP	b	-0.016	-0.014	-0.014	-0.015	-0.013
	t-stat	(-2.64)	(-3.49)	(-3.26)	(-4.33)	(-5.23)
	adj.R^2(%)	0.95	2.64	4.78	12.00	19.21

Panel B. Out-of-Sample Performance

Predictor	**Statistic**	**K=1**	**K=3**	**K=6**	**K=12**	**K=24**
EWIV+PWIV	OOS R^2(%)	0.769	2.292	5.491	12.067	20.135
198101 to 201812	z-stat	(2.23)	(2.55)	(2.99)	(3.55)	(5.09)
EWIV+IWIV	OOS R^2(%)	0.704	2.172	4.196	9.116	12.900
198101 to 201812	z-stat	(2.20)	(2.74)	(2.94)	(3.47)	(3.68)
SWIV+BWIV	OOS R^2(%)	0.911	3.026	6.697	13.949	23.949
198101 to 201812	z-stat	(2.65)	(3.15)	(3.61)	(4.19)	(5.30)
EWAP+VWAP	OOS R^2(%)	0.257	1.608	3.772	10.416	12.216
198101 to 201812	z-stat	(0.99)	(2.46)	(2.68)	(2.84)	(2.47)

This table summarizes various robustness checks for the predictive regression using monthly time-series data. K stands for the forecast time horizon in number of months. For forecast horizons beyond one period, the t-stat is computed using the GMM standard errors with $K-1$ Newey-West lag correction in the monthly time-series regressions. The sample period is from 196301 to 201812.

Table 8
Portfolio Return Predictability

			Time Horizon K=6					Time Horizon K=12				
Fama-French Port	**Predictor**	**Coefficient**	**Port 1**	**Port 2**	**Port 3**	**Port 4**	**Port 5**	**Port 1**	**Port 2**	**Port 3**	**Port 4**	**Port 5**
Market Size	EWIV	b	0.003	0.001	0.001	0.002	0.001	0.002	0.001	0.001	0.002	0.001
		t-stat	(1.78)	(0.89)	(0.72)	(1.31)	(0.38)	(1.56)	(0.65)	(0.59)	(1.10)	(0.35)
		adj.R^2(%)	1.02	0.24	0.20	0.72	0.04	1.87	0.39	0.29	1.27	0.04
	VWIV	b	0.001	-0.000	-0.001	0.000	-0.001	0.002	-0.000	-0.001	0.000	-0.001
		t-stat	(1.03)	(-0.09)	(-0.53)	(0.21)	(-0.80)	(1.09)	(-0.08)	(-0.36)	(0.22)	(-0.55)
		adj.R^2(%)	0.18	-0.15	-0.13	-0.07	0.04	0.41	-0.15	0.06	-0.12	0.69
	EWIV	b	0.007	-0.005	0.007	-0.006	0.008	0.005	-0.003	0.005	-0.005	0.007
		t-stat	(1.65)	(-1.26)	(1.85)	(-1.89)	(2.64)	(1.51)	(-1.13)	(1.84)	(-2.05)	(2.78)
	VWIV	b	-0.008	0.007	-0.006	0.008	-0.009	-0.007	0.007	-0.006	0.008	-0.008
		t-stat	(-2.85)	(2.66)	(-2.44)	(3.36)	(-3.86)	(-3.24)	(2.92)	(-3.00)	(3.56)	(-4.07)
		adj.R^2(%)	3.42	3.11	3.98	2.16	2.49	3.09	2.93	5.77	5.05	8.61
Book-to-Market	EWIV	b	0.000	0.002	0.002	0.001	0.002	0.000	0.002	0.002	0.001	0.002
		t-stat	(0.20)	(1.50)	(1.63)	(1.16)	(1.53)	(0.14)	(1.25)	(1.57)	(1.10)	(1.13)
		adj.R^2(%)	-0.06	0.83	0.89	0.36	1.05	-0.10	1.37	1.44	0.69	2.00
	VWIV	b	-0.002	0.000	0.000	0.001	0.001	-0.002	0.000	0.001	0.001	0.001
		t-stat	(-0.92)	(0.35)	(0.52)	(0.58)	(1.03)	(-0.81)	(0.31)	(0.70)	(0.64)	(0.56)
		adj.R^2(%)	0.10	-0.09	-0.04	-0.06	0.13	1.08	-0.09	-0.04	0.01	0.53
	EWIV	b	0.009	-0.010	0.007	-0.006	0.006	0.008	-0.009	0.006	-0.005	0.006
		t-stat	(2.88)	(-3.37)	(2.64)	(-2.43)	(2.30)	(3.11)	(-3.89)	(3.10)	(-3.08)	(2.85)
	VWIV	b	-0.005	0.004	-0.003	0.005	-0.004	-0.004	0.004	-0.003	0.005	-0.004
		t-stat	(-1.85)	(1.38)	(-1.05)	(1.68)	(-1.27)	(-2.76)	(1.74)	(-1.52)	(2.25)	(-1.85)
		adj.R^2(%)	2.83	2.79	2.43	0.80	1.94	8.51	4.92	4.42	1.40	3.04
Operating Profitability	EWIV	b	0.000	0.001	0.001	0.001	0.001	0.000	0.001	0.001	0.001	0.001
		t-stat	(0.14)	(0.84)	(0.45)	(0.74)	(0.53)	(0.02)	(0.75)	(0.36)	(0.64)	(0.45)
		adj.R^2(%)	0.00	0.22	-0.04	0.32	0.15	-0.11	0.51	0.06	0.42	0.13
	VWIV	b	-0.002	-0.000	-0.001	-0.001	-0.001	-0.002	-0.000	-0.001	-0.001	-0.001
		t-stat	(-0.71)	(-0.30)	(-0.77)	(-0.55)	(-0.62)	(-0.70)	(-0.11)	(-0.56)	(-0.32)	(-0.43)
		adj.R^2(%)	0.00	-0.15	0.07	-0.12	-0.13	1.05	-0.07	0.44	0.20	0.14
	EWIV	b	0.011	-0.012	0.007	-0.007	0.007	0.009	-0.011	0.007	-0.006	0.007
		t-stat	(2.85)	(-2.85)	(2.98)	(-2.99)	(3.07)	(3.35)	(-2.89)	(3.32)	(-3.35)	(3.52)
	VWIV	b	-0.008	0.008	-0.008	0.007	-0.007	-0.007	0.007	-0.007	0.007	-0.007
		t-stat	(-3.50)	(3.35)	(-3.53)	(2.43)	(-2.89)	(-4.02)	(3.45)	(-3.87)	(2.72)	(-3.50)
		adj.R^2(%)	2.70	1.94	2.86	3.57	2.08	7.94	5.52	6.98	8.41	5.21
Investment	EWIV	b	0.002	0.002	0.001	0.002	-0.001	0.002	0.002	0.001	0.002	-0.001
		t-stat	(1.68)	(1.88)	(0.95)	(1.44)	(-0.24)	(1.26)	(1.44)	(0.86)	(1.25)	(-0.31)
		adj.R^2(%)	1.30	1.03	0.20	1.06	-0.15	2.13	2.16	0.53	1.83	-0.05
	VWIV	b	0.001	0.001	-0.000	0.001	-0.003	0.001	0.001	0.000	0.001	-0.003
		t-stat	(0.52)	(0.86)	(-0.28)	(0.53)	(-1.32)	(0.38)	(0.65)	(0.06)	(0.42)	(-1.33)
		adj.R^2(%)	0.04	-0.01	-0.14	0.16	0.65	0.03	0.21	-0.11	0.11	3.22
	EWIV	b	0.008	-0.007	0.006	-0.004	0.006	0.007	-0.006	0.005	-0.004	0.005
		t-stat	(2.90)	(-2.73)	(2.72)	(-2.29)	(2.60)	(3.11)	(-2.88)	(2.97)	(-2.80)	(2.45)
	VWIV	b	-0.006	0.006	-0.005	0.011	-0.014	-0.005	0.007	-0.005	0.010	-0.013
		t-stat	(-2.78)	(2.54)	(-2.11)	(3.10)	(-3.62)	(-2.73)	(3.05)	(-2.91)	(3.82)	(-4.76)
		adj.R^2(%)	3.25	2.68	2.19	1.83	4.03	6.27	4.91	4.91	4.50	11.87

This table summarizes bivariate predictive regressions for portfolio returns on EWIV and VWIV. The dependent variable is various monthly portfolio excess returns and the forecast horizon is one month and one year in all regressions. I normalize all predictors to have zero mean and one standard deviation. For forecast horizons beyond one period, the t-stat is computed using the GMM standard errors with $K-1$ Newey-West lag correction in the monthly time-series regressions. The sample period is from 196301 to 201812.

Table 9
Reconcile with Goyal and Santa-Clara (2003) and Bali, et al. (2005)

Panel A. Goyal and Santa-Clara (2003)						
196308 to 199912	**Coefficient**	**K=1**	**K=3**	**K=6**	**K=12**	**K=24**
EWIV	b	0.006	0.005	0.004	0.004	0.003
	t-stat	(3.39)	(3.90)	(3.58)	(3.19)	(2.80)
	adj.R^2(%)	1.70	4.07	5.14	7.64	14.67
VWIV	b	0.004	0.004	0.003	0.002	0.002
	t-stat	(2.18)	(2.69)	(2.13)	(1.90)	(1.94)
	adj.R^2(%)	0.72	1.91	2.01	2.44	5.35
EWIV	b	0.012	0.010	0.009	0.009	0.007
	t-stat	(2.63)	(2.78)	(2.56)	(2.53)	(2.57)
VWIV	b	-0.006	-0.006	-0.006	-0.006	-0.005
	t-stat	(-1.35)	(-1.48)	(-1.51)	(-1.72)	(-1.77)
	adj.R^2(%)	1.85	4.63	6.35	10.62	18.44
Panel B. Bali et al. (2005)						
196308 to 200112	**Coefficient**	**K=1**	**K=3**	**K=6**	**K=12**	**K=24**
EWIV	b	0.003	0.003	0.002	0.002	0.003
	t-stat	(1.65)	(1.73)	(1.16)	(0.95)	(2.01)
	adj.R^2(%)	0.39	1.03	0.86	1.39	8.42
VWIV	b	0.001	0.001	0.000	-0.001	0.001
	t-stat	(0.34)	(0.34)	(-0.19)	(-0.54)	(0.59)
	adj.R^2(%)	-0.19	-0.17	-0.19	0.37	0.63
EWIV	b	0.014	0.012	0.011	0.011	0.009
	t-stat	(3.10)	(3.38)	(3.27)	(3.70)	(3.04)
VWIV	b	-0.012	-0.010	-0.010	-0.011	-0.008
	t-stat	(-2.43)	(-2.97)	(-2.98)	(-3.50)	(-2.54)
	adj.R^2(%)	1.42	3.36	5.55	13.92	18.59

Table 9 Continued:

Panel C. Small-Stock Effect (Only Use NYSE Stocks)						
196301 to 201812	**Coefficient**	**K=1**	**K=3**	**K=6**	**K=12**	**K=24**
EWIV	b	0.008	0.008	0.007	0.007	0.006
	t-stat	(3.00)	(3.36)	(3.21)	(3.10)	(2.95)
VWIV	b	-0.010	-0.009	-0.008	-0.007	-0.006
	t-stat	(-2.94)	(-3.41)	(-3.12)	(-3.27)	(-2.84)
	adj.R^2(%)	1.29	3.29	5.18	9.16	14.52

Panel D. Aggregate Idiosyncratic Variance						
196301 to 201812	**Coefficient**	**K=1**	**K=3**	**K=6**	**K=12**	**K=24**
EWIV	b	0.009	0.008	0.007	0.007	0.006
	t-stat	(3.17)	(3.65)	(3.51)	(3.16)	(2.77)
VWIV	b	-0.010	-0.009	-0.008	-0.007	-0.006
	t-stat	(-3.00)	(-3.55)	(-3.57)	(-3.29)	(-2.67)
	adj.R^2(%)	1.37	3.43	5.02	8.45	13.17

This table reports the results of univariate and bivariate monthly predictive regressions. The dependent variable is the average monthly value-weighted market excess returns in logarithm (MKTRF) over the relevant forecast horizon. All predictors are normalized to have zero mean and one standard deviation. K stands for the forecast horizon in number of months. "b" is the slope coefficient on the predictor. When $K > 1$, to adjust for the overlapping dependent variable, the t-stat is computed using the GMM standard errors with $K - 1$ Newey-West lag correction. The sample periods are specified in the table.

Table 10
VAR(1) based on the Conditional Covariance Risk and Other State Variables

Panel A. The Coefficients and t-stat from VAR(1)

	$\hat{\sigma}_{MH,t}$	Consumption	Income	CAY	Unrate	IndPro	TMS	DFY	DP
$\hat{\sigma}_{MH,t}$	0.928	-0.047	-0.003	0.010	-0.001	0.002	0.003	0.029	0.000
	(64.82)	(-2.89)	(-0.24)	(2.78)	(-0.14)	(0.18)	(1.65)	(1.56)	(1.26)
Consumption	-0.075	-0.094	0.065	-0.018	0.016	-0.002	-0.015	-0.187	0.003
	(-2.13)	(-2.34)	(2.48)	(-2.16)	(1.22)	(-0.11)	(-3.67)	(-4.13)	(6.64)
Income	-0.182	0.196	-0.188	-0.009	0.026	0.078	-0.016	-0.223	0.003
	(-3.43)	(3.24)	(-4.79)	(-0.66)	(1.37)	(2.44)	(-2.61)	(-3.28)	(4.06)
CAY	0.051	0.101	0.014	0.962	0.028	0.026	-0.007	-0.029	-0.001
	(1.09)	(1.90)	(0.40)	(84.84)	(1.65)	(0.91)	(-1.31)	(-0.49)	(-1.67)
Unrate	0.014	-0.052	0.001	-0.003	0.970	-0.059	-0.002	0.111	0.000
	(0.89)	(-2.98)	(0.02)	(-0.78)	(174.78)	(-6.37)	(-0.88)	(5.62)	(2.25)
IndPro	-0.033	0.137	0.047	-0.009	0.140	0.178	-0.018	-0.654	-0.001
	(-0.52)	(1.85)	(0.97)	(-0.55)	(5.97)	(4.54)	(-2.35)	(-7.88)	(-1.68)
TMS	1.223	-1.077	-0.672	-0.364	0.407	0.137	0.291	-0.878	-0.031
	(3.70)	(-2.85)	(-2.74)	(-4.54)	(3.40)	(0.68)	(7.57)	(-2.07)	(-7.25)
DFY	-0.050	-0.018	-0.007	0.010	-0.002	-0.015	0.004	0.949	0.001
	(-4.55)	(-1.42)	(-0.91)	(3.73)	(-0.51)	(-2.18)	(3.06)	(67.37)	(4.20)
DP	-0.804	-0.256	0.289	0.085	0.037	-0.533	-0.110	-0.830	0.992
	(-1.92)	(-0.54)	(0.93)	(0.83)	(0.24)	(-2.10)	(-2.25)	(-1.55)	(180.38)

Panel B. Correlation Matrix of the Residuals from VAR(1)

	$\hat{\sigma}_{MH,t}$	Consumption	Income	CAY	Unrate	IndPro	TMS	DFY	DP
$\hat{\sigma}_{MH,t}$	1								
Consumption	-0.15	1							
Income	-0.01	0.18	1						
CAY	0.10	0.14	-0.19	1					
Unrate	0.02	-0.06	-0.07	-0.02	1				
IndPro	-0.10	0.15	0.18	0.00	-0.28	1			
TMS	0.09	-0.21	-0.10	-0.07	0.17	-0.09	1		
DFY	0.01	-0.08	-0.05	-0.07	0.17	-0.12	0.19	1	
DP	0.20	-0.11	-0.08	0.18	-0.02	0.01	-0.10	0.04	1

The table reports the results of the Vector Autoregression (VAR) based on the conditional covariance risk and other state variables. In Panel A, I report the coefficients and the corresponding t-stats in the parentheses for each VAR. In Panel B, I report the Pearson correlation matrix of the residuals of each independent variable in the VAR (1). The sample period is from 196301 to 201812.

Table 11
Portfolios Sorted by Risk Exposures to EWIV and VWIV

Panel A. Equal-Weighted Portfolio (%)

	$\hat{\beta}_{iM,t}$	$\hat{\beta}_{iH,t} \propto \hat{C}_{i,t}$	$\hat{\beta}_{iH,t} \propto \hat{D}_{i,t}$	$\hat{\beta}_{iH,t} \propto (\hat{D}_t \times \hat{C}_{i,t} - \hat{C}_t \times \hat{D}_{i,t})$
Port 1	1.394	1.511	1.553	1.387
Port 2	1.469	1.296	1.293	1.323
Port 3	1.552	1.344	1.339	1.359
Port 4	1.604	1.470	1.491	1.517
Port 5	1.897	2.053	2.000	2.044
Port 5−1	0.503	0.542	0.448	0.657
t-stat	(2.33)	(5.75)	(4.30)	(5.27)
FF 5-factor alpha	-0.031	0.493	0.410	0.526
t-stat of alpha	(-0.29)	(5.13)	(3.94)	(4.62)

Panel B. Value-Weighted Portfolio (%)

	$\hat{\beta}_{iM,t}$	$\hat{\beta}_{iH,t} \propto \hat{C}_{i,t}$	$\hat{\beta}_{iH,t} \propto \hat{D}_{i,t}$	$\hat{\beta}_{iH,t} \propto (\hat{D}_t \times \hat{C}_{i,t} - \hat{C}_t \times \hat{D}_{i,t})$
Port 1	1.054	1.196	1.212	1.205
Port 2	1.123	1.168	1.109	1.181
Port 3	1.180	1.238	1.239	1.216
Port 4	1.264	1.307	1.327	1.301
Port 5	1.489	1.684	1.626	1.675
Port 5−1	0.436	0.488	0.414	0.470
t-stat	(1.93)	(3.62)	(2.93)	(3.55)
FF 5-factor alpha	0.054	0.382	0.352	0.363
t-stat of alpha	(0.38)	(2.87)	(2.47)	(2.78)

Panel C. Double Sorting by $\hat{\beta}_{iM,t}$ and $\hat{\beta}_{iH,t}$ (%)

Sort by $\hat{\beta}_{iH,t}$	Sort by $\hat{\beta}_{iM,t}$				
	Port 1	Port 2	Port 3	Port 4	Port 5
Port 1	1.158	1.223	1.271	1.290	1.626
Port 2	1.186	1.270	1.339	1.370	1.600
Port 3	1.102	1.298	1.415	1.387	1.616
Port 4	1.276	1.424	1.449	1.501	1.842
Port 5	1.719	1.803	1.633	1.790	2.256
Port 5−1	0.561	0.580	0.362	0.500	0.631
t-stat	(5.13)	(3.67)	(3.33)	(4.36)	(4.57)
FF 5-factor alpha	0.466	0.305	0.277	0.469	0.576
t-stat of alpha	(4.33)	(2.92)	(2.53)	(4.01)	(4.14)

This table summarizes the average returns and alphas in monthly frequencies for portfolios sorted by the risk loadings on corresponding independent variables and hold for one month. The firm risk exposures to both the market and the hedge portfolio are calculated based on Section 4.5. Column 3 to 5 provide the sorting performance of $\hat{\beta}_{iH,t}$ based on the three different proxies specified in Corollary 2.1. Panel A reports the equal-weighted average returns and alphas, while panel B reports the value-weighted (based on the market capitalization) average returns and alphas. Panel C exhibits the portfolio performnce double sorted by $\hat{\beta}_{iM,t}$ and $\hat{\beta}_{iH,t}$. The alpha is calculated based on the Fama-French five-factor model. The sample period is from 196301 to 201812. The t-stat is computed with Newey-West one-lag correction.

Table 12
Fama-Macbeth Regression

Factor	Coefficient	I	II	III	IV	V
Intercept	b	0.163	0.171	0.175	0.175	0.175
	t-stat	(6.92)	(7.38)	(7.26)	(7.26)	(7.26)
$\hat{\beta}_{iH,t}$	b	0.015	0.017	0.014	0.013	0.016
	t-stat	(4.16)	(4.94)	(4.06)	(3.83)	(3.59)
$\hat{\beta}_{iM,t}$	b		0.019	0.015	0.013	0.009
	t-stat		(2.13)	(1.73)	(1.62)	(1.32)
Market Size	b			-0.012	-0.012	-0.009
	t-stat			(-4.77)	(-4.95)	(-4.48)
Book-to-Market	b			0.020	0.020	0.018
	t-stat			(5.69)	(5.44)	(5.18)
Momentum	b				0.014	0.020
	t-stat				(2.43)	(4.25)
CIV Beta	b				-0.005	-0.012
	t-stat				(-2.49)	(-2.64)
Idiosyncratic Volatility	b					-0.013
	t-stat					(-1.28)
	Average adj.R^2(%)	0.81	2.49	3.85	5.08	6.05

This table reports the Fama-MacBeth regressions for all eligible stocks. The dependent variables are the cross-section of stock excess returns. The independent variables are the cross-sectional conditional risk exposures of the market and hedge portfolio based on Section 2, and other control variables in Section 3. The CIV Beta is constructed based on Herskovic, et al. (2016). All portfolios are formed at monthly frequency and held for one month. All dependent and independent variables are expressed as monthly values. The coefficients in the table are calculated by taking the time-series average of the cross-sectional regressions over time. The t-stat is computed with Newey-West one-lag correction. The sample period is from 196301 to 201812.

Table 13
Tail Risk and Conditional Covariance Risk

Panel A. Tail Risk based on Kelly and Jiang (2014)						
Predictor	**Coefficient**	**K=1**	**K=3**	**K=6**	**K=12**	**K=24**
$\bar{\eta}_{k,t}$	b	-0.001	-0.001	-0.002	-0.002	-0.002
	t-stat	(-0.76)	(-0.93)	(-1.23)	(-1.15)	(-1.14)
	adj.R^2(%)	-0.05	0.16	0.59	1.30	3.46
u_t	b	-0.001	-0.001	-0.001	-0.001	-0.001
	t-stat	(-0.29)	(-0.32)	(-0.52)	(-0.42)	(-0.52)
	adj.R^2(%)	-0.14	-0.11	-0.01	0.07	0.77
$\bar{\eta}_{k,t}$	b	-0.017	-0.019	-0.018	-0.019	-0.017
	t-stat	(-2.25)	(-2.92)	(-2.92)	(-3.23)	(-3.27)
	mARM t-stat	(-1.90)	(-2.46)	(-2.67)	(-2.54)	(-2.67)
u_t	b	0.016	0.018	0.017	0.018	0.016
	t-stat	(2.07)	(2.71)	(2.69)	(2.93)	(2.77)
	mARM t-stat	(1.99)	(2.55)	(2.26)	(2.38)	(2.45)
	adj.R^2(%)	0.49	2.36	4.63	10.72	18.97
Panel B. Compare Tail Risk with Conditional Covariance						
Predictor	**Coefficient**	**K=1**	**K=3**	**K=6**	**K=12**	**K=24**
λ_t^{Hill}	b	0.004	0.004	0.004	0.004	0.004
	t-stat	(2.40)	(2.93)	(2.93)	(3.21)	(3.69)
	adj.R^2(%)	0.62	2.29	4.47	9.87	18.13
$\hat{\sigma}_{MH,t}$	b	0.005	0.005	0.005	0.005	0.004
	t-stat	(3.12)	(3.64)	(3.63)	(3.72)	(4.17)
	adj.R^2(%)	1.30	3.50	5.97	12.34	21.70
λ_t^{Hill}	b	-0.001	-0.000	0.001	0.001	0.001
	t-stat	(-0.44)	(-0.07)	(0.25)	(0.34)	(0.69)
$\hat{\sigma}_{MH,t}$	b	0.007	0.005	0.005	0.005	0.004
	t-stat	(2.21)	(2.51)	(2.21)	(2.54)	(2.53)
	adj.R^2(%)	1.40	3.69	6.55	14.47	24.00
Panel C: Correlation between Tail Risk and Conditional Covariance						
Predictor	$\boldsymbol{\lambda_t^{Hill}}$	$\boldsymbol{\hat{\sigma}_{MH,t}}$				
λ_t^{Hill}	1	0.795				
$\hat{\sigma}_{MH,t}$	0.795	1				

This table reports the results of univariate and bivariate monthly predictive regressions. The dependent variable is the average monthly value-weighted market excess returns in logarithm (MKTRF) over the relevant forecast horizon. All predictors are normalized to have zero mean and one standard deviation. λ_t^{Hill} and $\hat{\sigma}_{MH,t}$ are defined in Section 5 and 4. K stands for the forecast horizon in number of months. "b" is the slope coefficient on the predictor. When $K > 1$, to adjust for the overlapping dependent variable, the t-stat is computed using the GMM standard errors with $K - 1$ Newey-West lag correction. The sample period is from 196301 to 201812.

Table 14
Simulation Analysis based on the ICAPM

Panel A: Simulated Data based on the Empirical Moments

	Summary Statistics			Pearson Correlation		
	Mean	**STD**	**AR(1)**	**EWIV**	**VWIV**	**SMV**
EWIV	0.033	0.008	0.88	1		
VWIV	0.021	0.006	0.874	0.926	1	
SMV	0.056	0.013	0.533	0.576	0.54	1

Panel B: Regression Summary based on Simulated Data

	Coefficient	**I**	**II**	**III**	**IV**	**V**	**VI**
EWIV	$\bar{\beta}$	0.163			2.013	2.008	
	% positive β	89.80			99.50	100.00	
	Average t-stat	1.225			5.436	5.408	
	% + significant t-stat	20.00			98.00	99.20	
VWIV	$\bar{\beta}$		-0.176		-3.014	-3.233	
	% negative β		80.80		99.50	100.00	
	Average t-stat		-0.876		-5.545	-5.776	
	% $-$ significant t-stat		12.50		98.00	99.20	
SMV	$\bar{\beta}$			1.909		3.012	
	% positive β			93.20		96.80	
	Average t-stat			1.401		1.837	
	% + significant t-stat			93.20		96.80	
λ^{Hill}	$\bar{\beta}$						0.032
	% positive β						92.80
	Average t-stat						5.010
	% + significant t-stat						92.80
	Average adj.R^2(%)	0.07	0.04	0.09	1.48	1.92	1.25

This table reports the simulated results of univariate and multiple regressions. Both independent and dependent variables are simulated based on the empirical moments described in Section 6. The simulation is run for 10,000 times. In Panel B, $\bar{\beta}$ is the average of the coefficients of the corresponding regression in each column. "Average t-stat" is the average of the t-stat of the corresponding regression for each simulation. "% + ($-$) Significance t-stat' is the proportion of positive (negative) t-stat greater (less) than 2 (-2). "Average adj.R^2 " is the average of adjusted R^2 obtained from the time-series regressions.

Figure 1
Monthly Time Series of EWIV and VWIV

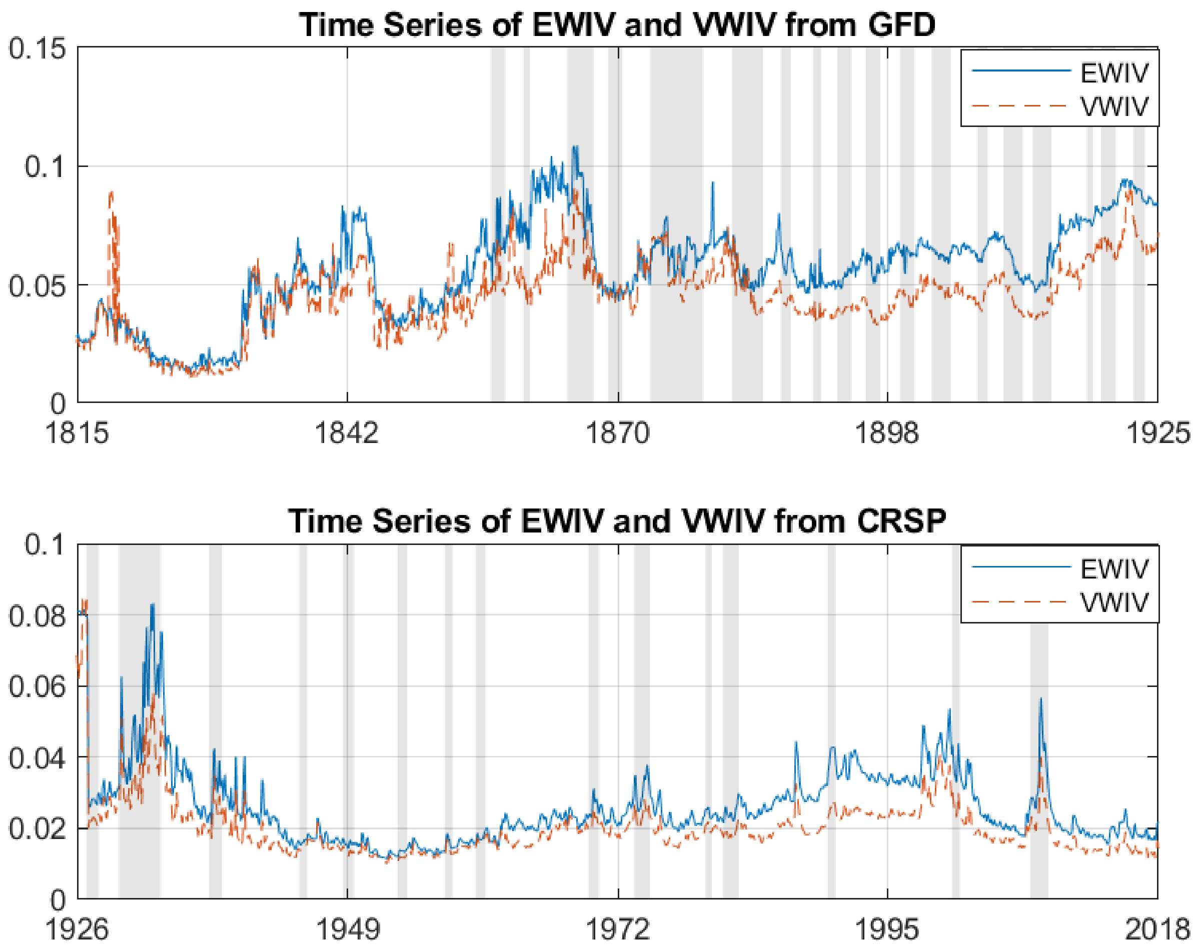

Figure 1 depicts the time series of EWIV and VWIV from 1815 to 2018. The data before 1926 is collected from Global Financial Data (GFD), while the data after 1926 is collected from CRSP. The variable frequency is at the monthly level. The grey areas indicate the National Bureau of Economic Research (NBER) recession periods.

Figure 2
Daily Time Series of EWIV and VWIV

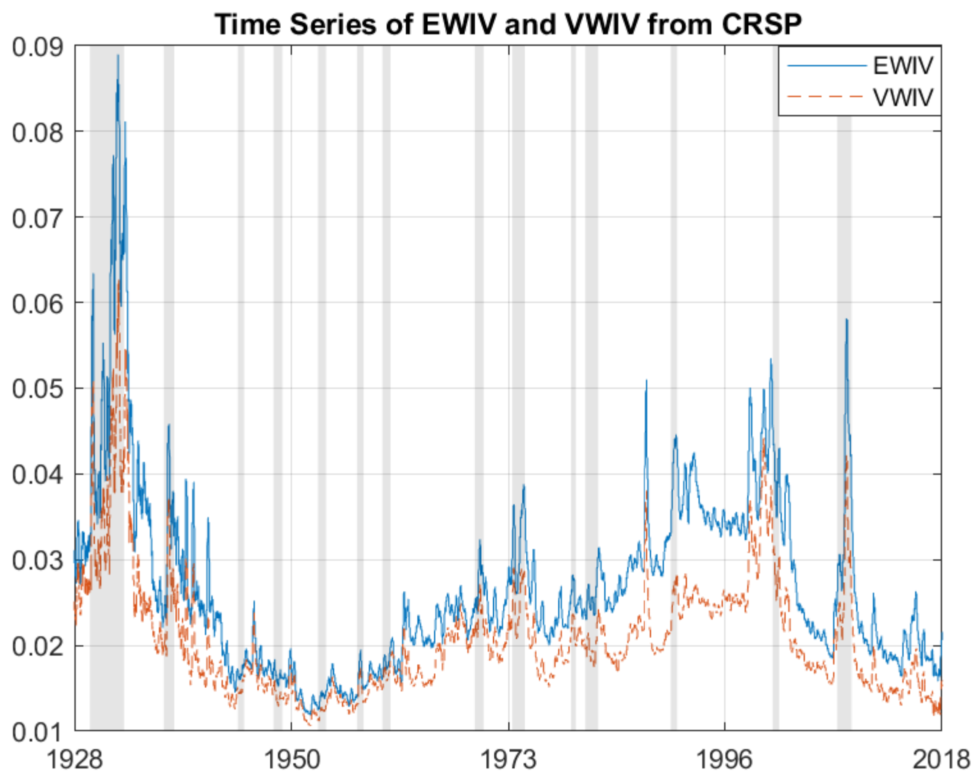

Figure 2 depicts the time series of EWIV and VWIV from 1928 to 2018. The data is collected from CRSP. The variable frequency is at the daily level. The grey areas indicate the National Bureau of Economic Research (NBER) recession periods.

Figure 3
Optimal Combination Weights of Aggregate Idiosyncratic Volatility (Alpha)

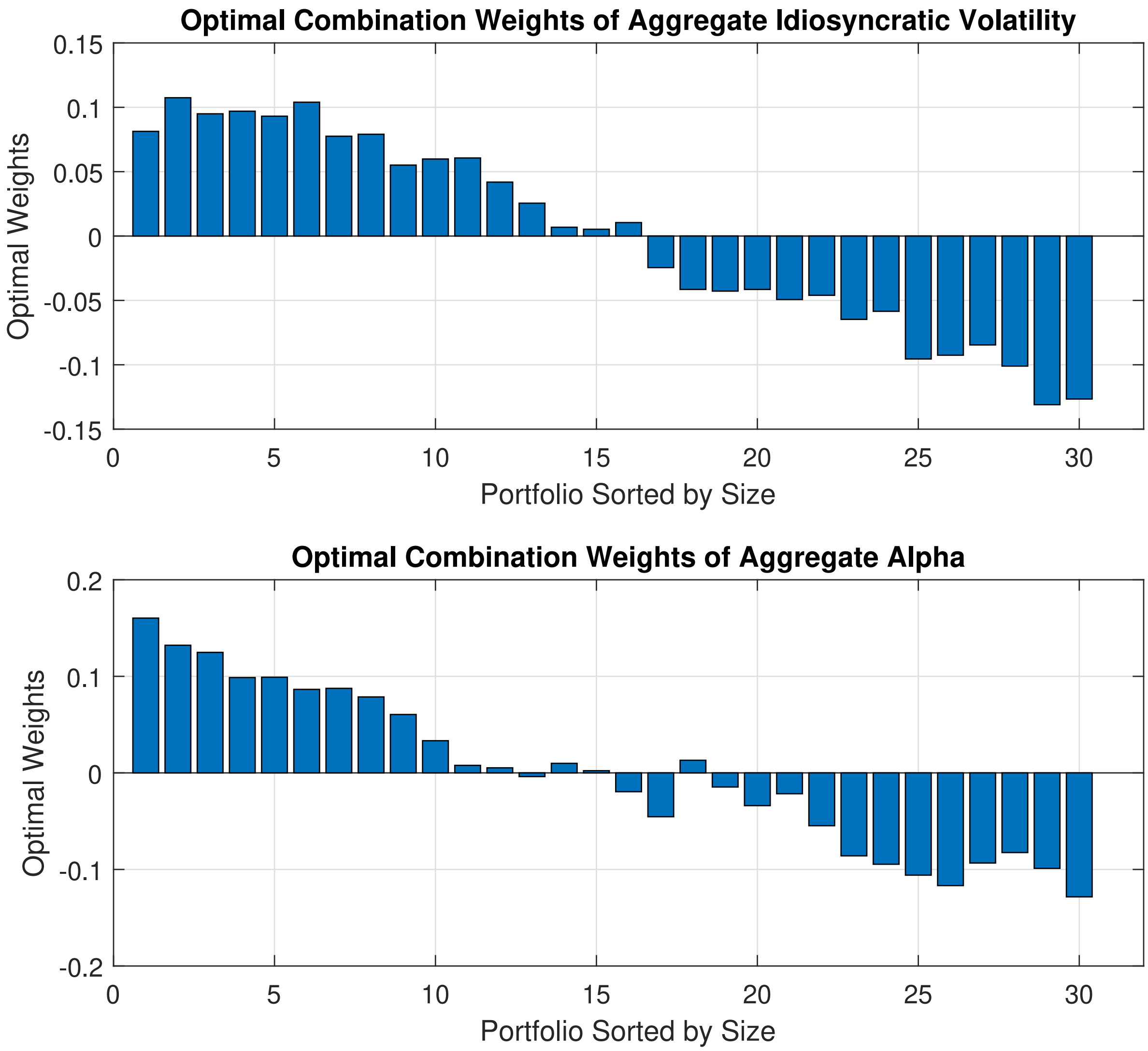

Figure 3 depicts the bar chats of the optimal weights to aggregate idiosyncratic volatility (alpha) for the thirty portfolios sorted by size in order to compute the condtional covariance risk. The optimal weights are estimated through the methodology proposed by Kelly and Pruitt (2015) in sample. The weights are scaled to sum up to 1 (-1) for positive (negative) values. The sample period is from 1963 to 2018. The variable frequency is at the monthly level.

Figure 4
Conditional Covariance Risk Estimated Through EWIV and VWIV

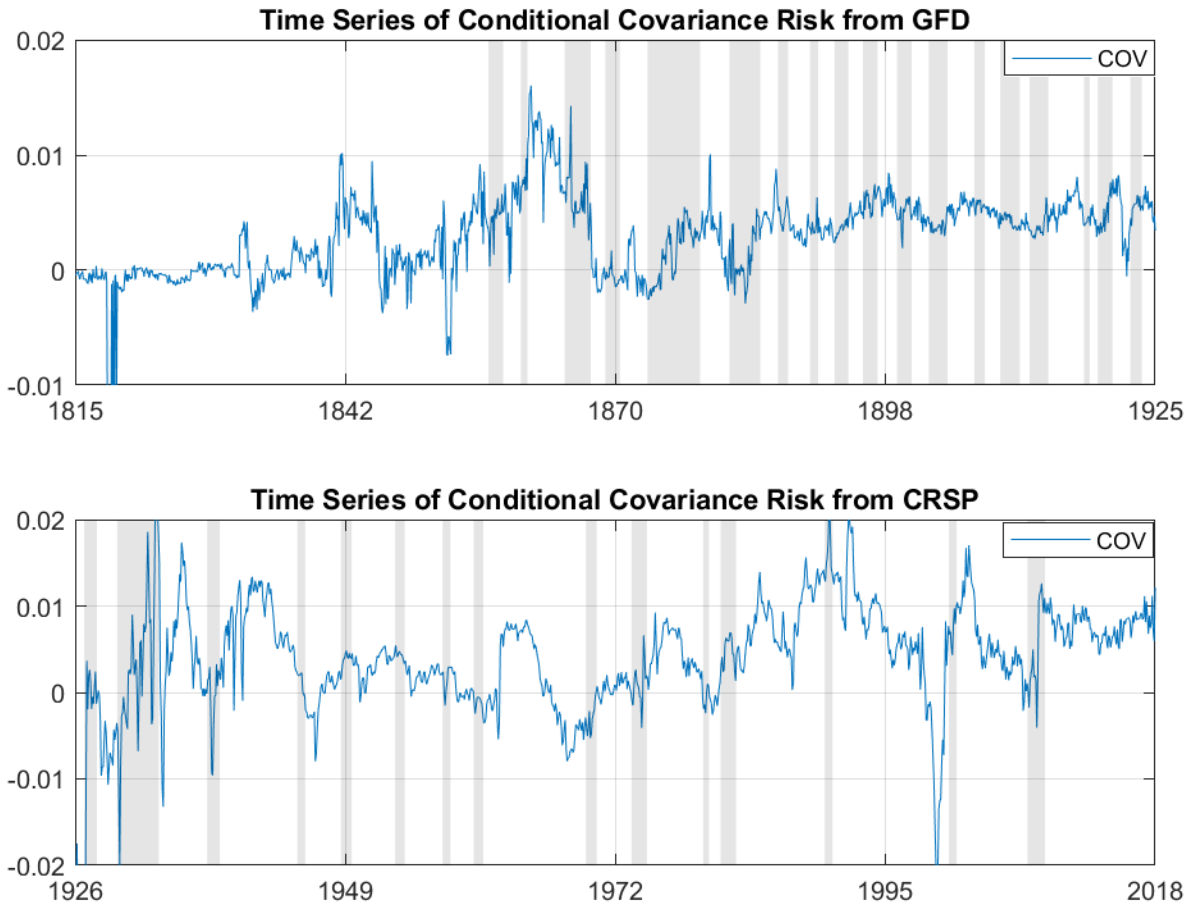

Figure 4 depicts the time series of conditional covariance risk from 1815 to 2018. The data before 1926 is collected from the Global Financial Data (GFD), while the data after 1926 is collected from CRSP. The conditional covariance risk is estimated based on the specification in Section 4. The variable frequency is at the monthly level. The grey areas indicate the National Bureau of Economic Research (NBER) recession periods.

Figure 5
Time Series of Conditional Covariance Risk through EWIV (EWAP) and VWIV (VWAP)

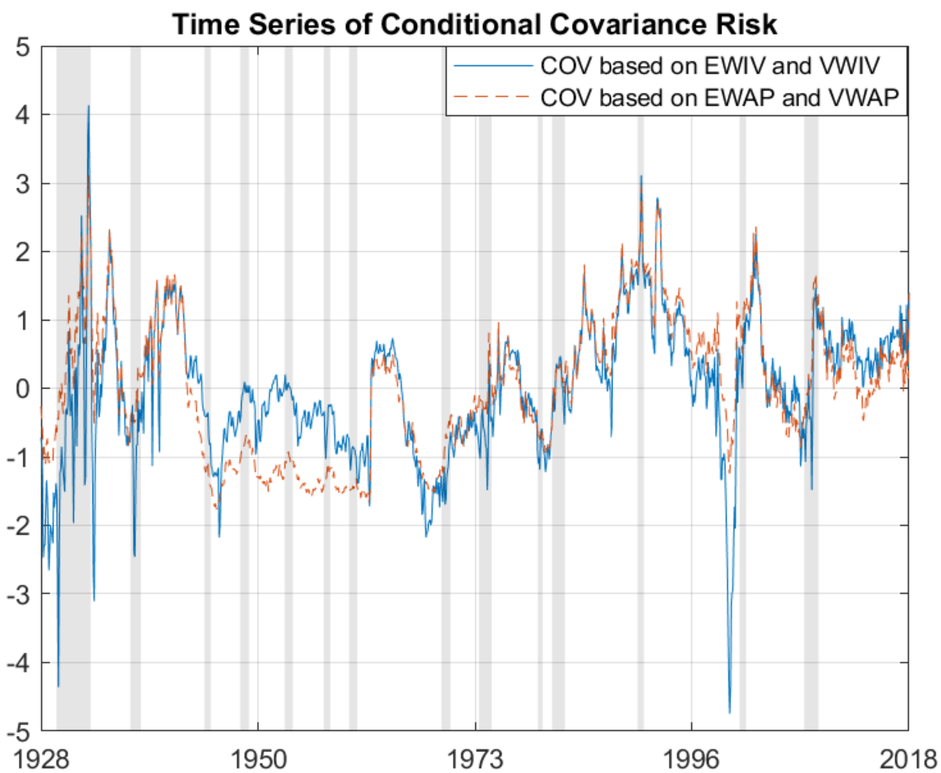

Figure 5 depicts the time series of conditional covariance risk from 1928 to 2018. The data is collected from CRSP. The conditional covariance risk is estimated based on either the aggregate idiosyncratic volatility or the aggregate first conditional moment of misspecified idiosyncratic shocks. The variable frequency is at the monthly level. The grey areas indicate the National Bureau of Economic Research (NBER) recession periods.

Figure 6
Monthly Time Series of Tail Risk and Conditional Covariance Risk

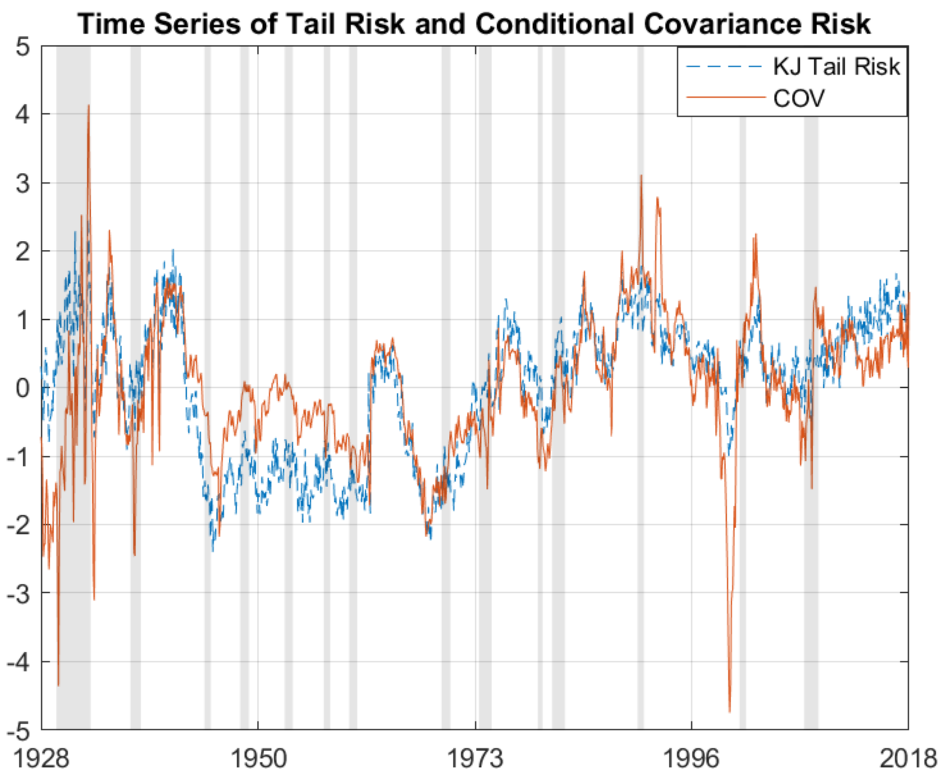

Figure 6 compares the time series of conditional covariance risk and tail risk proposed by Kelly and Jiang (2014) from 1928 to 2018. The data is from CRSP. The conditional covariance risk is estimated based on the specification in Section 4. The tail risk measure is constructed following Kelly and Jiang (2014). For comparability, I scale the value of both variables with zero mean and one standard deviation. The variable frequency is at the monthly level. The grey areas indicate the National Bureau of Economic Research (NBER) recession periods.