



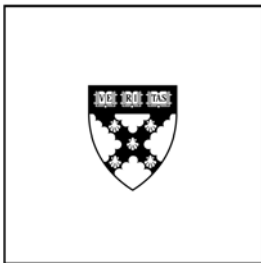
# DIGITAL ACCESS TO SCHOLARSHIP AT HARVARD

## Evaluating Firm-Level Expected-Return Proxies

The Harvard community has made this article openly available.  
[Please share](#) how this access benefits you. Your story matters.

<b>Citation</b>	Lee, Charles M.C., Eric C. So, and Charles C.Y. Wang. "Evaluating Firm-Level Expected-Return Proxies." Harvard Business School Working Paper, No. 15-022, October 2014.
<b>Accessed</b>	February 17, 2015 4:52:59 AM EST
<b>Citable Link</b>	<a href="http://nrs.harvard.edu/urn-3:HUL.InstRepos:13350443">http://nrs.harvard.edu/urn-3:HUL.InstRepos:13350443</a>
<b>Terms of Use</b>	This article was downloaded from Harvard University's DASH repository, and is made available under the terms and conditions applicable to Open Access Policy Articles, as set forth at <a href="http://nrs.harvard.edu/urn-3:HUL.InstRepos:dash.current.terms-of-use#OAP">http://nrs.harvard.edu/urn-3:HUL.InstRepos:dash.current.terms-of-use#OAP</a>

*(Article begins on next page)*



# Evaluating Firm-Level Expected-Return Proxies

Charles M. C. Lee  
Eric C. So  
Charles C.Y. Wang

Working Paper

15-022

October 9, 2014

Copyright © 2014 by Charles M. C. Lee, Eric C. So, and Charles C.Y. Wang

Working papers are in draft form. This working paper is distributed for purposes of comment and discussion only. It may not be reproduced without permission of the copyright holder. Copies of working papers are available from the author.

---

# Evaluating Firm-Level Expected-Return Proxies

---

## Abstract

We develop and implement a rigorous analytical framework for empirically evaluating the relative performance of firm-level expected-return proxies (ERPs). We show that superior proxies should closely track true expected returns both cross-sectionally and over time (that is, the proxies should exhibit lower measurement-error variances). We then compare five classes of ERPs nominated in recent studies to demonstrate how researchers can easily implement our two-dimensional evaluative framework. Our empirical analyses document a tradeoff between time-series and cross-sectional ERP performance, indicating the optimal choice of proxy may vary across research settings. Our results illustrate how researchers can use our framework to critically evaluate and compare a growing body of ERPs.

**JEL Classifications:** G10, G11, G12, G14, M41

## I Introduction

Expected rates of return play a central role in many managerial and investment decisions that affect the allocation of scarce resources. Recognition of this role has given rise to a substantial literature, spanning the fields of economics, finance, and accounting, about estimating expected rates of return for individual equities. The importance of firm-level estimates is widely understood, but consensus is lacking on how such estimates should be made. As a result, the specific estimation methods chosen by researchers vary widely across disciplines and studies, often without justification or discussion of alternative approaches.

Disagreement over how to estimate firm-level expected equity returns is exacerbated by the continued proliferation of new proxies proposed by researchers. One reason for this growth is the development of new asset-pricing models, each of which yields a specific theoretical formulation of expected returns. For each such formulation, furthermore, researchers propose innovations to the inputs used when empirically implementing expected-returns proxies, such as new forecasting techniques for earnings or inclusion of additional asset-pricing factors.<sup>1</sup> Thus, objectively comparing the relative merits of different firm-level proxies requires a rigorous evaluative framework for adjudicating between them. This paper offers such a framework.

Our central contribution is a two-dimensional framework for empirically assessing the relative quality of firm-level expected-return proxies (ERPs). Using a firm's true — but unobservable — expected-return as the normative benchmark, we define a given ERP's deviation from this benchmark as its measurement error. Although the measurement errors are themselves unobservable, we show that it is possible to derive characteristics of the distribution of errors for each ERP, such that researchers can compare the relative performance of alternative proxies.

---

<sup>1</sup>For example, [Gebhardt et al. \(2001\)](#) use a residual-income model and analysts' earnings forecasts to estimate firms' implied cost of equity capital. Subsequent researchers have modified this model by introducing the use of alternative growth forecasts (e.g., [Easton and Monahan, 2005](#)) and corrections for bias in analysts' forecasts (e.g., [Easton and Sommers, 2007](#)), and/or by replacing analysts' forecasts with mechanical earnings forecasts (e.g., [Hou et al., 2012](#)).

Our two-dimensional framework evaluates ERPs on the basis of their relative time-series and cross-sectional measurement-error variances. Prior studies on the performance of ERPs have focused almost exclusively on cross-sectional tests, with mixed results (See Section II for a discussion of this literature).<sup>2</sup> We advance this literature by introducing the time-series dimension into the performance evaluation of firm-level ERPs.

Our framework formalizes the intuition that well-performing ERPs should both track expected returns in the cross-section (that is, cross-sectional variation in ERPs should reflect cross-sectional variation in firms' expected returns) and track a given firm's expected returns closely over time (that is, time-series variation in a firm's ERP should reflect variation in its expected returns over time). Our framework allows researchers to characterize the cross-sectional and time-series dimensions of ERP performance for broad classes of firm-level proxies simultaneously and concisely. We show, both analytically and empirically, that the two dimensions of ERP performance are not redundant. We argue, further, that each can have a significant impact on research inferences in a given setting, and that researchers' preferences over these dimensions should depend on the particular application and/or research context. Thus, the optimal measure produced by this framework will depend on each proxy's cross-sectional and time-series performance and on the relative importance that a researcher assigns to each dimension.

To illustrate how researchers can implement our two-dimensional framework, we assess the relative performance of five families of ERPs (see Appendices I and II for a detailed description of each family). These five ERP families are based either on traditional equilibrium asset-pricing theory or on a variation of the implied-cost-of-capital (ICC) approach featured in accounting studies in recent years. Collectively, they encompass all of the prototype classes of ERPs nominated by the academic literature in both finance and accounting over the past 50 years.

---

<sup>2</sup>Most prior tests judge ERPs based on their ability to predict subsequent realized returns. Standard regression-based tests check whether the slope coefficient from a cross-sectional regression of ex-post returns on an ex-ante expected-returns proxy yields a coefficient of one (e.g., [Guay et al., 2011](#); [Easton and Monahan, 2005](#)).

Three of the ERPs we test originate in traditional equilibrium asset-pricing theory (the left-hand branch of the ERP tree depicted in Appendix I), in which non-diversifiable risk is priced, a firm's ERP is a linear function of its sensitivity to each risk factor (the  $\beta$ 's), and the risk premium associated with the factor (the  $\gamma$ 's). We test a single-factor version based on the Capital Asset Pricing Model (CAPM) and a similar multi-factor version based on four empirically inspired factors (FFF). We also test a characteristic-based expected-return estimate (CER), discussed in Lewellen (2014), in which a firm's factor loadings (the  $\beta$ 's) reflect its relative ranking in terms of each firm characteristic.

We also test two prototype ERPs from the ICC literature (the right-hand branch of the ERP tree in Appendix I). The implied-cost-of-capital is the internal rate of return that equates a firm's market value to the present value of its expected future cash-flows; ICCs have become increasingly popular as a class of ERPs. We test a commonly used method of estimating ICC drawn from Gebhardt et al. (2001) and Hou et al. (2012). Finally, we develop a new ERP prototype by computing a "fitted" version of the Gebhardt et al. (2001) measure that we refer to as FICC. This proxy is new to the literature, but it seems to us to reflect a natural progression in the evolution of ERPs. FICC is based on an instrumental variable approach, whereby each firm's ICC estimate is regressed on a vector of firm characteristics. Each firm's FICC estimate is therefore a "fitted" value from the regression — that is, it is a linear function of the firm's current characteristics.

Our results show that ICC, CER, and FICC dramatically outperform the traditional factor-based proxies (CAPM, FFF), both in the cross-section and in time-series. Among the three non-factor-based proxies, CER, the proxy nominated by Lewellen (2014), performs best at exhibiting the lowest variance in cross-sectional measurement errors; the two implied-cost-of-capital proxies, ICC and FICC, perform better in the time-series tests. The performance of the new proxy, FICC, reflects its hybrid nature in that it offers lower cross-sectional measurement-error variance than ICC and relatively lower time-series measurement-error variance than CER. Our evidence is consistent with the findings of Lewellen (2014), which

show that characteristic-based proxies exhibit good return-predictability in the cross-section and suggest that these proxies may be more reliable than ICC proxies. However, we show that ICC proxies outperform CER in time-series. These findings suggest that, in research contexts where the cross-sectional variation of expected returns is of greater importance, such as in investments or capital budgeting, CER may be preferable. In contexts where time-series tracking of expected returns is of greater importance, such as studying the impact of certain shocks on firms' expected returns (e.g., Callahan et al., 2012; Dai et al., 2013), ICC might perform best. When both time-series and cross-sectional performance are important, FICC might be the best option.

Overall, our empirical analyses give further credence to our two-dimensional framework by documenting a tradeoff between time-series and cross-sectional ERP performance, such that the optimal proxy may vary across research objectives and settings. We hope and expect that, by providing a rigorous tool for evaluating the relative performance of expected-return proxies, this framework will provide guidance on ERP selection and stimulate further thought and research on a matter of central import to researchers, investors, and corporate managers.

The rest of the paper is organized as follows. Section II discusses related literature. Section III presents the theoretical underpinnings of our performance metrics. Section IV provides details on our sample construction and empirical results. Section V concludes.

## II Related Literature

A large and growing literature examines the impact of regulation, managerial decisions, and market design on firm-level expected returns. For example, firm-level expected returns have been used variously to study the effect of disclosure levels (Botosan, 1997; Botosan and Plumlee, 2005), information precision (Botosan et al., 2004), legal institutions and security laws (Hail and Leuz, 2006; Daouk et al., 2006), cross-listings (Hail and Leuz, 2009), corporate governance (Ashbaugh et al., 2004), accrual quality (Francis et al., 2004; Core et al., 2008), taxes and leverage (Dhaliwal et al., 2005), internal control deficiencies (Ashbaugh-Skaife

et al., 2009), voluntary disclosure (Francis et al., 2008), and accounting restatements (Hribar and Jenkins, 2004). In all of these studies, the research objective is to examine the effect of various elements in its information environment on a firm’s expected-return. Although these studies focus on factors affecting firm-level expected returns, they do not address the performance-evaluation problems that we identify here.

The specific methods of estimating expected returns chosen by researchers vary widely across studies and disciplines, often without justification or discussion of alternative approaches.<sup>3</sup> Furthermore, a related stream of research aims to develop new estimates of expected returns, often by modifying existing proxies via the introduction of new inputs, such as forecasts of earnings or growth. The evaluation framework that we present here provides a means to compare the relative merits of existing proxies within a given research context; it also provides a tool to gauge whether a new proxy represents an advancement by using the performance of existing proxies as a minimum benchmark. By establishing how to implement this benchmark, our framework introduces clarity into a muddled and continually growing pool of potential ERPs.

The value of assessing ERPs within a two-dimensional framework is intuitive. In many decision contexts, such as investment and capital budgeting, we would like ERPs to reflect cross-sectional differences in true expected returns. In numerous other research contexts, however, it is crucial for time-series variation in a firm’s ERPs from one period to the next to reflect variations in the firm’s true expected returns — for example, when researchers use a difference-in-differences research design to study the impact of a regulatory change on a firm’s expected returns. In these settings the time-series dimension is more relevant, but existing performance tests do not assess the quality of ERPs along this dimension. Unlike prior studies that focus on cross-sectional differences in ERP performance (e.g., Easton and

---

<sup>3</sup>In recent years a substantial literature on ICCs has developed, first in accounting, and now increasingly in finance. The collective evidence from these studies indicates that the ICC approach offers significant promise in dealing with a number of longstanding empirical asset-pricing conundrums. See Easton and Sommers (2007) for a summary of the accounting literature prior to 2007. In finance, the ICC methodology has been used to test the Intertemporal CAPM (Pástor et al., 2008), international asset-pricing models (Lee et al., 2009), and default risk (Chava and Purnanandam, 2010).



Monahan, 2005), our time-series tests allow researchers to identify the most suitable ERP for tracking a firm's expected-return variation over time in a particular context. Thus a key contribution of our paper is demonstrating how researchers can implement this critical second dimension of ERP performance evaluation.<sup>4</sup>

The paper most closely related to ours is Easton and Monahan (2005), hereafter referred to as EM, which derives a methodology for relative comparisons of cross-sectional measurement-error variance between alternative ERPs. Our paper complements and extends EM's analysis in several ways. First, we argue and demonstrate that better-performing ERPs should track true expected returns not only in the cross-section but also over time, thus allowing researchers to more comprehensively assess the relative performance of ERPs. Second, EM's framework is based on stricter assumptions, making it more difficult to apply their methodology to compare broad classes of ERPs. Specifically, EM assumes that a proxy's measurement errors are uncorrelated with true expected returns, making their measure inappropriate for a large class of ERPs. As a simple example, any ERP that is a multiple of true expected returns would violate the assumption necessary to use their approach because the ERP's error would be clearly correlated with the true expected-return. Third, our approach circumvents the requirement of the EM framework to estimate multiple firm-specific and cross-sectional parameters (e.g., cash-flow news); it is thus much simpler to implement empirically. Overall, though our cross-sectional measurement-error variance metric is conceptually similar to EM, our two-dimensional framework is more parsimonious, easier to implement, applicable to broad classes of ERPs, and it is more comprehensive by incorporating time-series performance evaluation.

In a related study, Botosan et al. (2011) proposes an alternative approach to evaluating ERPs, on the basis of their associations with risk proxies. A central difference between our

---

<sup>4</sup>We note that other papers have also examined time-series properties, in particular of ICCs. For example, Easton and Sommers (2007) examines the properties of aggregate risk premiums implied by ICCs and the role of analyst biases. Pástor et al. (2008) assess the time-series relations between aggregate risk premiums and market volatility. Whereas these studies focus on the properties of aggregate expected returns, our framework focuses on the time-series performance of firm-level ERPs under a unified measurement-errors-based framework.

approach and that of [Botosan et al. \(2011\)](#) has to do with which construct is assumed to be valid. [Botosan et al. \(2011\)](#) assume that expected returns, as an economic construct, must entail certain associations with assumed risk proxies. Our approach relies on expected returns as a statistical construct, and our method relies on the properties of conditional expectations. Similarly, [Larocque and Lyle \(2013\)](#) proposes a framework for assessing ERPs on the basis of their ability to predict accounting returns. The central difference between our approach and theirs is the assumed normative benchmark. Whereas our framework assumes the normative benchmark to be a firm’s true expected returns, their framework assumes the benchmark to be a firm’s future returns on equity.

### III Theoretical Underpinnings

This section begins with a simple decomposition of returns and then derives our two-dimensional evaluation framework for expected-return proxies.

#### Return Decomposition

We begin with a simple decomposition of realized returns:

$$r_{i,t+1} = er_{i,t} + \delta_{i,t+1}, \tag{1}$$

where  $r_{i,t+1}$  is firm  $i$ ’s realized return in period  $t + 1$ , and  $\delta_{i,t+1}$  is the firm’s unanticipated news or forecast error.<sup>5</sup> In this framework, we define  $er_{i,t}$  as the firm’s true but unobserved expectation of future returns conditional on publicly available information at time  $t$ , capturing all ex-ante predictability (on the basis of the information set) in returns. By the property of conditional expectations,  $er_{i,t}$  is optimal or efficient in the sense of minimizing mean squared errors.<sup>6</sup>

<sup>5</sup>In [Campbell \(1991\)](#), [Campbell and Shiller \(1988a,b\)](#), and [Vuolteenaho \(2002\)](#), the news term is further decomposed into cash-flow and expected-returns components. This is not necessary for our purposes.

<sup>6</sup>Known as the Prediction Property of conditional expectations (e.g., [Angrist and Pischke, 2008](#)).

It follows from this definition and the property of conditional expectations that a firm's expected returns ( $er_{i,t}$ ) cannot be systematically correlated with its forecast errors ( $\delta_{i,t+1}$ ), in time-series or in the cross-section.<sup>7</sup> Intuitively, if expected returns were correlated with subsequent forecast errors, one could always improve on the expected-return measure by taking into account such systematic predictability, thereby violating the definition of an optimal forecast.<sup>8</sup>

Having thus defined our normative benchmark, we abstract away from the market-efficiency debate. If one subscribes to market efficiency,  $er_{i,t}$  should only be a function of risk factors and of expected risk premia associated with these factors. Conversely, in a behavioral framework,  $er_{i,t}$  can also be a function of other non-risk-related behavioral factors.

Next, we introduce the idea of ERPs ( $\widehat{er}_{i,t+1}$ ), defined as the unobserved expected-return ( $er_{i,t+1}$ ) measured with error ( $\omega_{i,t+1}$ ):

$$\widehat{er}_{i,t+1} = er_{i,t+1} + \omega_{i,t+1}. \quad (2)$$

In concept,  $\widehat{er}_{i,t+1}$  need not be an ICC estimate as defined in the accounting literature — it can be any ex-ante expected-return measure, including a firm's Beta, its book-to-market ratio, or its market capitalization at the beginning of the period. The key is that, whatever the “true” expected-return may be, we do not observe it. What we can observe are empirical proxies that contain measurement error. Our goal here is to evaluate how good a job these proxies do at capturing or tracking  $er_{i,t}$ .

Differences between alternative ERPs are reflected in the properties (time-series and cross-sectional) of their  $\omega$  terms. Comparisons between different ERPs are, therefore, comparisons of the distributional properties of the  $\omega$ 's they generate, over time and across firms. Statements we make about the desirability of one ERP over another are, in essence, expressions of preference with regard to the properties of the alternative measurement errors (i.e.,

<sup>7</sup>Known as the Decomposition Property of conditional expectations (e.g., [Angrist and Pischke, 2008](#)).

<sup>8</sup>[Fama and Gibbons \(1982\)](#) make a similar argument in relating observed ex-post real interest rates to unobserved ex ante expected real interest rate.

the  $\omega$  terms) that each is expected to generate. In other words, when we choose one ERP over another, we are specifying the loss function (in terms of measurement error) that we find least distasteful or problematic. The choice of ERPs thus becomes a choice between the attractiveness of alternative loss functions, expressed over  $\omega$  space.

Under this setup, what would superior ERPs look like? We cannot nominate a single criterion by which all ERPs should be judged — doing so is impossible without specifying the researcher’s preference function over the properties of measurement errors. In our setup, however, “well-behaved” ERPs will exhibit certain empirical attributes. The extent to which they do so thus becomes a basis for comparison.

## Comparing ERPs

Recall that our main objective is to produce ERPs that track true expected returns well, both across firms and over time. Equivalently, we would like measurement errors to be small at all times (i.e.,  $\omega_{i,t} \sim 0$  for all  $i$  and all  $t$ ). Because these expectations are unlikely to hold, we must choose between alternative error distributions, and specify those properties of  $\omega$  that are most important to us as researchers.

If the measurement errors ( $\omega$ ’s) are non-trivial, two further properties become important. First, we want *measurement errors for a given firm that are stable over time*. That is, all else equal, ERPs with lower time-series variance in measurement errors are preferred. If  $\omega_{i,t}$  is stable over time, the ERP for a given firm will track its true expected returns more closely in time-series. Consequently, changes in a firm’s ERP over time will be informative about changes in its underlying expected returns, rather than merely reflecting changes in ERP measurement errors. For example, an ERP with constant measurement errors over time is ideal, since its time-series variations will precisely reflect variations in the underlying unobserved expected returns. This is particularly useful in research contexts when studying the impact of regulation or disclosure policy on a firm’s expected returns (e.g., [Callahan et al., 2012](#); [Dai et al., 2013](#)).

Second, we prefer *measurement errors that are stable across firms at a given point in time*. If this property holds, cross-sectional differences in ERPs are more informative about differences in expected returns between firms. For example, an ERP with constant measurement errors across firms is ideal, since differences in ERPs precisely capture differences in the underlying expected returns. This is particularly desirable in investment or capital budgeting decisions.

The stability of measurement errors over time and across firms is captured by the notion of lower measurement-error variance (in both time-series and cross-section). Thus, to capture how well ERPs track the underlying unobserved expected returns we propose two empirical properties by which to assess expected-return proxies: lower measurement-error variance in time-series and in the cross-section. Note that these two properties do not necessarily imply each other. As shown below, an ERP that exhibits perfect time-series tracking ability could exhibit noisy measurement errors in the cross-section; similarly, an ERP exhibiting perfect cross-sectional tracking ability could exhibit great intertemporal variations in measurement errors. This non-redundancy can also be seen in our empirical tests in Section IV.

## **Time-Series and Cross-Sectional Measurement-Error Variances**

This section formalizes the two dimensions of ERP performance evaluation. The subsequent discussion presents the basic foundation and intuition of our evaluative framework. For brevity, most of the technical details appear in the [Technical Appendix](#).

### ***Time-Series Error Variance***

To assess the stability of ERP measurement errors over time, we must be able to empirically identify and compare the time-series variance of the error terms,  $Var_i(\omega_{i,t})$ , generated by different ERPs. A key objective (and, we believe, contribution) of this paper is to analytically disentangle the time-series properties of ERP measurement errors from those of the true expected returns, when both are time-varying and persistent over time.

We show in this sub-section (and in the [Technical Appendix](#)), that it is possible to derive an empirically estimable and firm-specific measure, Scaled Time-Series Variance, denoted as  $SVar_i(\omega_{i,t})$ , that allows us to compare alternative ERPs in terms of their time-series measurement-error variance, even when the errors themselves are not observable. This analysis then provides the foundation for our comparison of ERPs.

The [Technical Appendix](#) provides a detailed derivation of  $Var_i(\omega_{i,t})$  and  $SVar_i(\omega_{i,t})$ . Under the assumptions that (a) expected returns and measurement errors are jointly covariance-stationary, and (b) future news is unforecastable, we show in equation (T7) in the Technical Appendix that the time-series measurement-error variance of a given ERP for firm  $i$  can be expressed as:

$$Var_i(\omega_{i,t}) = Var_i(\hat{er}_{i,t}) - 2Cov_i(r_{i,t+1}, \hat{er}_{i,t}) + Var_i(er_{i,t}), \quad (3)$$

where  $Var_i(\hat{er}_{i,t})$  is the time-series variance of a given ERP for firm  $i$ ,  $Var_i(er_{i,t})$  is the time-series variance of firm  $i$ 's expected returns, and  $Cov_i(r_{i,t+1}, \hat{er}_{i,t})$  is the time-series covariance between a given ERP and realized returns for firm  $i$  in period  $t + 1$ .

The first term on the right-hand-side shows that the (time-series) variance of a given ERP's measurement error is increasing in the variance of the ERP [ $Var_i(\hat{er}_{i,t})$ ]. This is intuitive: as the time-series variance of measurement errors of a given ERP for firm  $i$  increases, all else equal, so will the observed variance of the ERP.

The second term on the right-hand side shows that the variance of the error terms for a given ERP is decreasing in the covariance of the ERP and future returns [ $Cov_i(r_{i,t+1}, \hat{er}_{i,t})$ ]. This is also intuitive: to the extent the within-firm covariance between a given ERP and future realized returns is consistently positive over time, the variance of the errors will be smaller. In other words, to the extent a given expected-return proxy consistently predicts variation in future returns for the same firm, time-series variation in that proxy is more likely to reflect variation in the firm's true expected returns than in measurement errors.

Finally, notice that the third term,  $Var_i(er_{i,t})$ , is the time-series variance of firm  $i$ 's true (but unobserved) expected returns. For a given firm, this variable is constant across alternative ERPs and therefore does not play a role in relative performance comparisons. In other words, we need only the first two terms of (2) to determine which expected-return proxy exhibits lower time-series variance in measurement errors. Accordingly, we define the sum of the first two terms of (2) as the Scaled TS Variance:

$$SVar_i(\omega_{i,t}) = Var_i(\hat{er}_{i,t}) - 2Cov_i(r_{i,t+1}, \hat{er}_{i,t}). \quad (4)$$

In our empirical tests, we compute for each ERP and each firm the relative error variance measure using (4), and then assess the time-series performance of ERPs based on the average of  $SVar_i$  across the  $N$  firms in our sample:

$$AvgSVar^{TS} = \frac{1}{N} \sum_i SVar_i(\omega_{i,t}). \quad (5)$$

For a given sample, ERPs that exhibit lower average time-series measurement-error variances [ $Var_i(\omega_{i,t})$ ] also exhibit lower average scaled TS variance ( $AvgSVar^{TS}$ ). All else equal, ERPs with lower time-series measurement-error variances for a given sample are deemed to be of higher quality because time-series variation in the expected-return proxy is more likely to reflect changes in firms' expected returns than is time-series variation in measurement errors.

Note that  $AvgSVar^{TS}$  facilitates relative comparisons across ERPs. If we impose additional structure (that is, if we make stricter assumption about the time-series behavior of expected returns), it is possible to obtain an empirically estimable absolute measure of the time-series measurement-error variance.<sup>9</sup> Note too that our time-series framework allows researchers to pick the best ERP for a specific firm on the basis of  $SVar_i(\omega_{i,t})$ .

---

<sup>9</sup>This can be done, for example, by assuming that expected returns and ERP measurement errors follow AR(1) processes (e.g., Wang, 2014). The empirical tests in this paper, however, do not require such assumptions.

### *Cross-Sectional Error Variance*

Although low time-series measurement-error variance is desirable, this criterion alone is not sufficient to assess the quality of an ERP. When choosing between two ERPs that track true expected-return equally well in time-series, we will unambiguously prefer the one whose errors are more stable in the cross-section, since cross-sectional variations in such proxies are likely to reflect the cross-sectional variation in true expected returns.

Employing similar logic, we show in [Part B](#) of the Technical Appendix that it is possible to derive an empirically estimable and proxy-specific measure, Average Scaled Cross-Sectional Variance ( $AvgSVar^{CS}$ ), that allows us to compare the cross-sectional measurement-error variance of alternative ERPs. We show that the cross-sectional measurement-error variance of a given ERP for a given cross-section  $t$  can be expressed as:

$$Var_t(\omega_{i,t}) = Var_t(\hat{er}_{i,t}) - 2[Var_t(er_{i,t}) + Cov_t(er_{i,t}, \omega_{i,t})] + Var_t(er_{i,t}), \quad (6)$$

where  $Var_t(\hat{er}_{i,t})$  is a given ERP's cross-sectional variance at time  $t$ ,  $Var_t(er_{i,t})$  is the cross-sectional variance in firms' expected returns at time  $t$ , and  $Cov_t(er_{i,t}, \omega_{i,t})$  is the cross-sectional covariance between firms' ERPs at time  $t$  and their realized returns in period  $t+1$ .

Since the cross-sectional variance in firms' expected returns — the last term — is invariant across different ERPs, relative comparisons of cross-sectional ERP measurement-error variance can be made by comparing the Scaled Cross-Sectional Variance:

$$SVar_t(\omega_{i,t}) = Var_t(\hat{er}_{i,t}) - 2[Var_t(er_{i,t}) + Cov_t(er_{i,t}, \omega_{i,t})]. \quad (7)$$

In particular, our empirical tests assess the cross-sectional performance of ERPs based on the average of  $SVar_t$  across the  $T$  cross-sections in our sample:

$$AvgSVar^{CS} = \frac{1}{T} \sum_t Var_t(\hat{er}_{i,t}) - 2[Var_t(er_{i,t}) + Cov_t(er_{i,t}, \omega_{i,t})]. \quad (8)$$



Part B of the Technical Appendix shows that  $AvgSVar^{CS}$  can be estimated by

$$\frac{1}{T} \sum_t \text{Var}_t(\hat{e}r_{i,t}) - 2\text{Cov}_t(r_{i,t+1}, \hat{e}r_{i,t}), \quad (9)$$

following assumptions similar to those that characterize the time-series case above. Equation (9) indicates that, all else equal, an ERP's average cross-sectional measurement-error variance is increasing in the cross-sectional variance of the ERP. In other words, when ERPs are noisier, there is more measurement error. The second term suggests that, all else equal, an ERP's average measurement-error variance is decreasing to the degree that ERPs predict future returns in the cross-section. In other words, when the cross-sectional variation of an ERP captures more of the cross-sectional variation in realized returns, variations in proxies are more likely to reflect the true variations in expected returns.

### *The Two-Dimensional Framework*

The two dimensions of our evaluation framework are not redundant. An ERP that performs well in time-series may perform very poorly in the cross-section. For example, an ERP can have firm-specific measurement errors that are constant across time, resulting in zero time-series measurement-error variance, but these measurement errors can obscure the cross-sectional ordering of expected returns across firms. Consider two stocks, A and B, with constant true expected returns of 10 percent and 2 percent respectively. Suppose that a particular ERP model produces expected-returns proxies of 2 percent and 10 percent for stocks A and B respectively. Such an ERP produces zero time-series measurement-error variance for both stocks, since the measurement errors are constant across time for each firm, but such an ERP mis-orders the stocks' expected returns in the cross-section and produces cross-sectional error variance.

Conversely, an ERP that performs well in the cross-section may perform poorly in time-series. For example, an ERP can have time-specific measurement errors that are constant

across firms but vary over time, resulting in zero cross-sectional measurement-error variance but potentially substantial time-series measurement-error variance. Suppose again that the true expected returns of stocks A and B are always 10 percent and 2 percent, respectively. Now consider an ERP model that produces expected-return proxies for A and B of 13 percent and 5 percent in certain years, and 10 percent and 2 percent in other years. Such an ERP always correctly orders expected returns in the cross-section and exhibits constant measurement errors for each firm in the cross-section (i.e., zero measurement-error variance in each cross-section), but will produce time-series error variance. In this case, time-series variation in the proxies does not reflect variations in true expected returns, but it reflects variations in measurement errors. In sum, an ERP that is equal to true expected returns is clearly a “perfect” ERP in our framework. More broadly, as shown above, a perfect ERP in our framework may have non-zero measurement error, so long as these errors are (a) constant across time for a given firm and (b) constant in the cross-section for all firms.

As noted earlier, EM also derive a methodology to rank ERPs on the basis of their cross-sectional measurement errors using a measure they call the modified noise variable. Like us, their measurement facilitates relative comparisons of cross-sectional measurement-error variance between alternative ERPs. However, their measure is based on stricter assumptions. Specifically, EM assume that ERP measurement errors are uncorrelated with true expected returns, which makes their measure inappropriate for a large class of ERPs. For example, any ERPs of the form  $C \times er_{i,t}$  for some constant  $C$  would violate the assumptions necessary to use this “modified noise variable” to compare cross-sectional measurement error variance. Interestingly, EM also assume, to facilitate the use of their empirical metric, that on average the difference in the cross-sectional covariance of ERP measurement errors ( $\omega$ ) and news ( $\delta$ ) is “second-order” between any pair of ERPs; taken to the extreme, this assumption is equivalent to our assumption of zero cross-sectional correlation between news and expected returns (and therefore their proxies and measurement errors), on average. Overall, our cross-sectional performance metric is conceptually similar to EM. But we believe that ours is more

parsimonious, easier to implement empirically, and applicable to broad classes of ERPs, and that it extends the scope of analysis to include time-series performance evaluation.

In sum, we have provided a rationale for a two-dimensional evaluation framework that compares ERPs under a set of minimalistic assumptions. Researchers can use equations (5) and (9) to gauge the relative performance of ERPs to determine the optimal choice for a given research context. The following section applies this evaluation framework to assess the merits of five representative ERP measures nominated by prior literature.

## IV Empirical Implementation

A key strength of our two-dimensional evaluation framework is that it can be implemented in a small set of empirical analyses and is thus easily portable across research settings. To illustrate how researchers can empirically implement our framework, this section supplements our theoretical analyses by evaluating the relative performance of five families of monthly expected-return proxies. These five ERP groups are based either on traditional equilibrium asset-pricing theory or on some variation of the ICC approach featured in accounting studies in recent years. Collectively, they span all the prototype classes of ERPs nominated by the academic finance and accounting literatures over the past 50 years.

Three of the ERPs we test originate in traditional equilibrium asset-pricing theory, where non-diversifiable risk is priced. Specifically, we test a single-factor version based on the Capital Asset Pricing Model (CAPM) and similar multi-factor version based on [Fama and French \(1993\)](#). We also test a characteristic-based ERP presented in [Lewellen \(2014\)](#) in which a firm's factor loadings (the  $\beta$ 's) embody its relative ranking in terms of each firm characteristic.

We also test two prototype ERPs from the ICC literature (the right-hand branch of the ERP tree in Appendix I). ICC is the internal rate of return that equates a firm's market value to the present value of its expected future cash-flows. Finally, we develop a new ERP prototype by computing a "fitted" version of the [Gebhardt et al. \(2001\)](#) measure based on

an instrumental-variable approach, whereby each firm’s ICC estimate is therefore a “fitted” value from the regression — i.e., a linear function of the firm’s current characteristics. This section outlines our sample-selection process and the methodologies underlying our estimates of expected-return proxies.

## Sample Selection

We obtain market-related data on all U.S.-listed firms (excluding ADRs) from the Center for Research in Security Prices (CRSP) and annual accounting data from Compustat for the period 1977-2011. For each firm-month, we estimate five expected-return proxies using data from the CRSP Monthly Stock file and, when applicable, firms’ most recent annual financial statements. To be included in our sample, each firm-month observation must include information on stock price, shares outstanding, book values, earnings, dividends, and industry identification (SIC) codes. We also require each firm-month observation to include valid, non-missing values for each of our five expected-return proxies, detailed below. Our final sample consists of 1,549,530 firm-month observations, corresponding to 12,022 unique firms.

## Factor-Based Expected-Return Proxies

Our empirical tests include two estimates of expected returns to those derived from standard factor models: CAPM and a four-factor model based on [Fama and French \(1993\)](#) that adds a momentum factor (the UMD factor obtained from Ken French’s [data library](#)). At the end of each calendar month  $t$ , we estimate the expected one-month-ahead returns as

$$\widehat{\mathbb{E}}_t[r_{i,t+1}] = rf_{t+1} + \sum_{j=1}^J \hat{\beta}_i \widehat{\mathbb{E}}_t[f_{j,t}] \quad (10)$$

for each factor model (with  $J = 1, 4$  factors), where  $rf_{t+1}$  is the risk-free rate in period  $t + 1$ ,  $\hat{\beta}_i$  are the estimated factor sensitivities (estimated in time-series for each firm using monthly

stock and factors' returns over the 60 months prior to the forecast date), and  $f_{j,t}$  are the corresponding factors in period  $t$ . Expected monthly factor returns are estimated based on trailing average 60-month factor returns.

We denote the capital-asset pricing model and a four-factor Fama-French type model as CAPM and FFF respectively. We estimate CAPM for each firm at the end of each calendar month using historical factor sensitivities. Specifically, we first estimate each firm's Beta to the market factor using the prior 60 months' data (from  $t - 1$  to  $t - 60$ ). CAPM is then obtained by multiplying the estimated Beta by the most recent 12 months' compounded annualized market-risk premium (provided by Fama and French) and adding the risk-free rate. Similarly, FFF represents a four-factor based ERP computed using the Mkt-Rf, HML, SMB, and UMD factors and 60-month rolling Beta estimates.

### Implied-Cost-of-Capital (ICC)

We use the methodology in Gebhardt et al. (2001) to estimate a firm's implied-cost-of-capital. ICC is a practical implementation of the residual income valuation model that employs a specific forecast methodology, forecast period, and terminal value assumption.<sup>10</sup> Specifically, the time- $t$  ICC expected-returns proxy for firm  $i$  is the  $\widehat{er}_{i,t}^{ICC}$  that solves

$$P_{i,t} = B_{i,t} + \sum_{n=1}^{11} \frac{\frac{\mathbb{E}_t[NI_{i,t+n}]}{\mathbb{E}_t[B_{i,t+n-1}]} - \widehat{er}_{i,t}^{ICC}}{\left(1 + \widehat{er}_{i,t}^{ICC}\right)^n} \mathbb{E}_t[B_{i,t+n-1}] + \frac{\frac{\mathbb{E}_t[NI_{i,t+12}]}{\mathbb{E}_t[B_{i,t+11}]} - \widehat{er}_{i,t}^{ICC}}{\widehat{er}_{i,t}^{ICC} \left(1 + \widehat{er}_{i,t}^{ICC}\right)^{11}} \mathbb{E}_t[B_{i,t+11}], \quad (11)$$

where  $\mathbb{E}_t[NI_{i,t+n}]$  is the  $n$ -year-ahead forecast of earnings estimated using the approach in Hou et al. (2012). We estimate the book value per share,  $B_{i,t+n}$ , using the clean surplus relation, and apply the most recent fiscal year's dividend-payout ratio ( $k$ ) to all future expected earnings to obtain forecasts of expected future dividends, i.e.,  $\mathbb{E}_t[D_{i,t+n+1}] = \mathbb{E}_t[NI_{i,t+n+1}] \times k$ .

<sup>10</sup>Also known as the Edwards-Bell-Ohlson model, the residual income model simply re-expresses the dividend discount model by assuming that book value forecasts satisfy the clean surplus relation,  $\mathbb{E}_t[B_{i,t+n+1}] = \mathbb{E}_t[B_{i,t+n}] + \mathbb{E}_t[NI_{i,t+n+1}] - \mathbb{E}_t[D_{i,t+n+1}]$ , where  $\mathbb{E}_t[B_{i,t+n}]$ ,  $\mathbb{E}_t[NI_{i,t+n}]$ , and  $\mathbb{E}_t[D_{i,t+n}]$ , are the time  $t$  expectation of book values, net income, and dividends in  $t + n$ .

We compute ICC as of the last trading day of each calendar month for all U.S. firms (excluding ADRs and those in the Miscellaneous category in the Fama-French 48-industry classification scheme), combining monthly prices and total-shares data from CRSP and annual financial-statements data from Compustat.

### **Characteristic-Based Expected-Return Proxies**

Following [Lewellen \(2014\)](#), we calculate a characteristic-based ERP, which we denote as CER, by first estimating a firm’s factor loadings to three characteristics. This measure is based on an instrumental-variable approach, whereby each firm’s returns are regressed on a vector of firm characteristics in each cross-section (i.e., using Fama-MacBeth regressions); then the historical average of the estimated slope coefficients is applied to a given forecast period’s observed firm characteristics to obtain a proxy of expected future returns.

We also compute a “fitted” expected-return proxy, which uses ICCs instead of returns as the dependent variable. We refer to this fitted version of ICC, which represents a “fitted” value using historically estimated Fama-MacBeth coefficients, as FICC. We apply rolling 10-year Fama-MacBeth coefficients in our implementations of FICC and CER. [Figure 1](#) reports rolling average Fama-MacBeth slope estimates from cross-sectional regressions of expected-returns proxies on firms’ size, book-to-market, and return momentum. The left-hand panel reports rolling average coefficients using realized returns as the dependent variable; the right-hand panel uses the ICC as the dependent variable. The coefficients from the ICC regressions are noticeably smoother than those from returns, consistent with FICC’s circumvention of some of the noise in returns by equating prices to estimates of future earnings.

### **Descriptive Statistics**

[Table 1](#) reports the medians of five monthly expected-return proxies for each year from 1977 through 2011. We compute expected-return proxies for each firm-month in our main sample based on the stock price and on publicly available information as of the last trading

day of each month. Our sample consists of firm-months for which all five expected-return proxies are non-missing. The number of firm-months varies by year, ranging from a low of 30,930 in 1978 to a high of 61,487 in 1998. The average number of firm-months per year is 44,272, indicating that expected-return proxies are available for a broad cross-section of stocks in any given year. The time-series means of the monthly median expected-return proxies range from 0.93 percent (for ICC), to 1.59 percent (for CAPM).

It is instructive to compare the results for the two factor-based proxies (CAPM and FFF) with their non-factor-based counterparts (ICC, CER, and FICC). Recall that we compute CAPM and FFF using firm-specific Betas estimated over the previous 60 months and a continuously updated market-risk premium provided by Fama-French. The monthly means for CAPM and FFF (1.59 percent and 1.55 percent) are similar to those of the non-factor-based proxies. However, the time-series standard deviation of the factor-based proxies is 3 to 5 times larger than the standard deviation of the non-factor-based proxies. Annual medians for CAPM range from -1.88 percent to 4.53 percent; in 5 out of 35 years the median of CAPM is negative, indicating that more than half of the monthly observations signal expected returns below zero. The volatility of CAPM and FFF reflects the instability of the market equity risk premium estimated on the basis of historical realized returns.

Table 2 reports the average monthly Spearman correlations among the five expected-return proxies. We calculate correlations by month and then average them over the sample period. The table shows that the three non-factor-based proxies are highly correlated among themselves, as are the two factor-based proxies. However, we find no positive correlation across the two groups — that is, none of the three non-factor-based proxies is positively correlated with the two factor-based proxies. In fact, the correlations between the non-factor-based and factor-based proxies are generally negative (finding consistent with earlier findings reported by [Gebhardt et al., 2001](#)).

## Comparison of Measurement-Error Variances

As noted in Section III, better expected-return proxies should generate measurement errors with lower cross-sectional variance. Table 3 presents descriptive statistics for the cross-sectional variances of scaled measurement errors of the five expected-return proxies. Scaled measurement-error variances are calculated for each unique calendar-month/expected-return proxy pair using equation (7), as follows:

$$SVar_t(\omega_{i,t}) = Var_t(\hat{e}r_{i,t}) - 2Cov_t(r_{i,t+1}, \hat{e}r_{i,t}),$$

where  $SVar_t(\omega_{i,t})$  is the cross-sectional scaled measurement-error variance in month  $t$ ,  $\hat{e}r_{i,t}$  is the expected-return proxy in month  $t$ , and  $r_{i,t+1}$  is the realized return in month  $t + 1$ .

Panel A of Table 3 reports summary statistics for the error variance from each model, using a sample of data for 418 calendar months during our 1977-2011 sample period. Table values in this panel represent descriptive statistics for the error variance from each expected-return proxy computed across these 418 months.

Recall from Section III that we do not estimate the third component of  $Var_t(\omega)$  in Equation (6); instead, we estimate the scaled measurement-error variances as in Equation (7). One implication of omitting this third term, which is a variance and therefore non-negative, is that the resulting estimates of scaled measurement-error variances can be negative. Hence, to ensure that  $SVar_t(\omega_{i,t})$  is positive, we multiply our estimates by 100 and add an arbitrary constant of 10 when reporting summary statistics.

Panel A of Table 3 shows that the three non-factor-based proxies (ICC, CER, and FICC) generate smaller cross-sectional error variances than the other two. Panel B reports t-statistics based on Newey-West-adjusted standard errors corresponding to the pair-wise comparisons of cross-sectional scaled measurement-error variances within the sample of 418 months used in Panel A.

The reported values in Panel B of Table 3 are negative (positive) when the expected-



return proxy displayed in the leftmost column has a larger (smaller) scaled measurement-error variance than the expected-return proxy displayed in the topmost row. The [Panel B](#) findings indicate that all three non-factor-based proxies significantly outperform CAPM and FFF. Among the non-factor-based proxies, furthermore, both CER and FICC outperform the remaining three proxies, indicating that CER and FICC are best suited to rank firms in our sample in terms of their true expected returns.

According to the second dimension of our evaluative framework, better expected-return proxies should also generate measurement errors with lower time-series variance. [Table 4](#) reports a time-series measure of error variance for each of the five expected-return proxies, calculated for each unique firm/expected-return proxy pair using [equation \(4\)](#) as follows:

$$SVar_i(\omega_{i,t}) = Var_i(\hat{e}r_{i,t}) - 2Cov_i(r_{i,t+1}, \hat{e}r_{i,t}),$$

where  $SVar_i(\omega_{i,t})$  is the scaled measurement-error variance of firm  $i$ ,  $\hat{e}r_{i,t}$  is the expected-return proxy in month  $t$ , and  $r_{i,t+1}$  is the realized return in month  $t + 1$ .

[Panel A](#) provides descriptive statistics for the variance of the error terms. To construct this panel, we require each firm to have a minimum of 20 (not necessarily consecutive) months of data during our 1977-2011 sample period. A total of 12,022 unique firms met this data requirement. Table values in this panel represent summary statistics for the error variance (multiplied by 100) from each expected-return proxy computed across these 12,022 firms; again, we add a constant of 10 to each estimate before calculating summary statistics. [Panel B](#) reports t-statistics corresponding to the pair-wise comparison of firm-specific measurement errors across the sample of 12,022 firms used in [Panel A](#).

The results in [Table 4](#) show that the three non-factor-based proxies (ICC, CER, and FICC) generate lower time-series error variances than the other two. [Panel B](#) shows that ICC in particular generates measurement-error variances that are more stable over time than all other expected return proxies. Thus, the three non-factor-based proxies are not

merely better in the cross-section than the Beta-based proxies; they also behave better in time-series, on average.

Figure 2 illustrates the results from our evaluative framework. The Y-axis depicts the median cross-sectional scaled measurement-error variance and the X-axis depicts the median time-series scaled measurement-error variances. Therefore, the upper left-hand corner of the figure demarcates an efficiency frontier where error variances are minimized.

This efficient frontier depicts the two-dimensional framework that researchers can use to compare alternative ERPs within their specific research setting. The figure shows that ICC, FICC, and CER are the best-performing proxies in terms of minimal error variance. The efficient frontier defined by these three proxies illustrates that the choice of expected-return proxies depends on the researcher's particular loss function with respect to the stability of cross-sectional measurement errors versus the stability of time-series measurement errors. Strikingly, ICC, FICC, and CER all perform much better than the two factor-based measures, FFF and CAPM. The figure also illustrates the tendency for models that perform well on one dimension to do so on the other as well.

Why do certain proxies outperform others? Figure 2 helps to identify the relative strengths of the CER and ICC. CER performs better in terms of cross-section; ICC performs better in the time-series. One possible explanation for this phenomenon is that CER assumes a stable risk premium, relative to firm characteristics, based on a long panel of historical data (the estimated rolling coefficients on firm characteristics shown in Appendix II). These estimated risk premiums are relatively stable over time, while firm characteristics (i.e., risk factors) vary substantially. To the extent that risk premiums on these firm characteristics are moving over time, we would expect greater time-series measurement-error variances (though the cross-sectional ordering of expected returns may be less affected because firms are still sorted in each period according to their current risk factors).

Additionally, we find that ICC performs better in the time-series but not in the cross-section. To understand this result, recall that ICCs are fully forward-looking and do not rely

on estimating historical average premiums as CER does. Because estimating ICCs does not require calculating historical premiums, ICCs are more fluid across periods than CER; thus, ICC trades off decreased measurement errors in the time-series for increased measurement errors in the cross-section. These tradeoffs give further credence to our two-dimensional framework by empirically documenting a tradeoff between time-series and cross-sectional ERP performance, indicating that the appropriate proxy may vary across research objectives and settings.

Overall, our results offer a much more sanguine assessment of the implied-cost-of-capital approach than does some prior literature. A large and growing literature uses implied-cost-of-capital estimates as de-facto proxies for firm-level expected returns. However, prior research provides limited assurance that implied-cost-of-capital estimates are in fact useful proxies for expected returns. After examining seven implied-cost-of-capital estimates, for example, EM concluded that “for the entire cross-section of firms, these proxies are unreliable.” By contrast, we show that ERPs based on an implied-cost-of-capital approach (ICC, FICC) are attractive in terms of their cross-sectional performance, and also strongly outperform alternative ERPs in time-series tests.

Like those of [Botosan et al. \(2011\)](#), our findings raise questions about the EM assertion that ICC estimates are unreliable. Unlike [Botosan et al. \(2011\)](#), however, our primary evaluation criteria do not necessarily require superior ERPs to exhibit stronger empirical correlations with estimated Beta or other presumed risk proxies.

One caveat is that the conclusions drawn from our empirical analyses can depend on the sample used in the analysis (e.g., [Ecker et al., 2013](#)) and on the variations of ERPs considered. The empirical exercise presented here is intended to illustrate the implementation of our two-dimensional evaluation framework for ERPs, but we hope that this framework will help researchers to determine which ERPs are appropriate for their intended research setting.

## V Conclusion

Estimates of expected returns play a central role in many managerial and investment decisions that affect the allocation of scarce resources in society. This study addresses a key problem in the literature that relies on estimates of expected returns: how to assess the relative performance of expected-returns proxies (ERPs) when prices are noisy.

Our paper demonstrates the importance of evaluating the time-series performance of ERPs, whereas most prior studies focus on cross-sectional performance evaluation. Evaluating an ERP's time-series performance is crucial in numerous research contexts, such as when researchers use a difference-in-differences research design to study the impact of a regulatory change on a firm's expected returns.

We derive a two-dimensional evaluation framework that explicitly models both time-series and cross-sectional measurement-error variances for ERPs. Using a firm's true but unobservable expected-return as the normative benchmark, we define an ERP's deviation from this benchmark as its measurement error. Although the measurement errors themselves are unobservable, we show that it is possible to derive characteristics of the distribution of errors for each ERP such that researchers can compare the relative performance of alternative proxies.

Our main goal is less to establish the superiority of specific ERPs, than to demonstrate the value of an easily implementable performance-evaluation framework derived from a minimalistic set of assumptions. We do not assume here that Beta or future realized returns are normative benchmarks by which ERPs should be measured. By establishing a rigorous evaluative framework, our findings help researchers select the appropriate ERP for a given context, and thus establish a minimum bar for what should be demanded from new entrants in the vast and still growing pool of firm-level expected-return proxies.

## References

- Angrist, J. D. and Pischke, J. 2008. Mostly harmless econometrics: An empiricist's companion. Princeton University Press.
- Ashbaugh, H., Collins, D., and LaFond, R. 2004. Corporate governance and the cost of equity capital. Working Paper, University of Wisconsin and University of Iowa.
- Ashbaugh-Skaife, H., Collins, D., Kinney Jr., W., and Lafond, R. 2009. The effect of sox internal control deficiencies on firm risk and cost of equity. Journal of Accounting Research, 47(1):1–43.
- Botosan, C. 1997. Disclosure level and the cost of equity capital. The Accounting Review, 72(3):323–349.
- Botosan, C., Plumlee, M., and Wen, H. 2011. The relation between expected returns, realized returns, and firm risk characteristics. Contemporary Accounting Research, 28(4):1085–1122.
- Botosan, C., Plumlee, M., and Yuan, X. 2004. The role of information precision in determining the cost of equity capital. Review of Accounting Studies, 9:233–259.
- Botosan, C. A. and Plumlee, M. A. 2005. Assessing alternative proxies for the expected risk premium. The Accounting Review, 80(1):21–53.
- Callahan, C. M., Smith, R. E., and Spencer, A. W. 2012. An examination of the cost of capital implications of fin 46. The Accounting Review, 87(4):1105–1134.
- Campbell, J. 1991. A variance decomposition for stock returns. The Economic Journal, 101(405):157–179.
- Campbell, J. and Shiller, R. 1988a. The dividend-price ratio and expectations of future dividends and discount factors. Review of financial studies, 1(3):195–228.
- Campbell, J. and Shiller, R. 1988b. Stock prices, earnings, and expected dividends. Journal of Finance, 43(3):661–676.
- Chattopadhyay, A., Lyle, M. R., and Wang, C. C. 2014. The cross section of expected returns around the world. Working Paper.
- Chava, S. and Purnanandam, A. 2010. Is default risk negatively related to stock returns? Review of Financial Studies, 23(6):2523–2559.
- Core, J. E., Guay, W. R., and Verdi, R. 2008. Is accruals quality a priced risk factor? Journal of Accounting and Economics, 46(1):2 – 22.
- Dai, Z., Shackelford, D. A., Zhang, H. H., and Chen, C. 2013. Does financial constraint affect the relation between shareholder taxes and the cost of equity capital? The Accounting Review, 88(5):1603–1627.

- Daouk, H., Lee, C., and Ng, D. 2006. Capital market governance: How do security laws affect market performance? Journal of Corporate Finance, 12(3):560–593.
- Dhaliwal, D., Krull, L., Li, O., and Moser, W. 2005. Dividend taxes and implied cost of equity capital. Journal of Accounting Research, 43(5):675–708.
- Easton, P. and Monahan, S. 2005. An evaluation of accounting-based measures of expected returns. The Accounting Review, 80:501–538.
- Easton, P. and Sommers, G. 2007. Effect of analysts' optimism on estimates of the expected rate of return implied by earnings forecasts. Journal of Accounting Research, 45(5):983–1015.
- Ecker, F., Francis, J., Olsson, P., and Schipper, K. 2013. Associations between realized returns and risk proxies using non-random samples. Working Paper.
- Fama, E. and French, K. 1993. Common risk factors in the returns on stocks and bonds. Journal of Financial Economics, 33(1):3–56.
- Fama, E. and French, K. 1996. Multifactor explanations of asset pricing anomalies. The Journal of Finance, 51:55–84.
- Fama, E. and Gibbons, M. 1982. Inflation, real returns and capital investment. Journal of Monetary Economics, 9(3):297–323.
- Francis, J., LaFond, R., Olsson, P. M., and Schipper, K. 2004. Costs of equity and earnings attributes. The Accounting Review, 79(4):967–1010.
- Francis, J., Nanda, D., and Olsson, P. 2008. Voluntary disclosure, earnings quality, and cost of capital. Journal of Accounting Research, 46.1:53–99.
- Gebhardt, W., Lee, C., and Swaminathan, B. 2001. Toward an implied cost of capital. Journal of Accounting Research, 39(1):135–176.
- Guay, W., Kothari, S., and Shu, S. 2011. Properties of implied cost of capital using analysts' forecasts. Australian Journal of Management, 36(2):125–149.
- Hail, L. and Leuz, C. 2006. International differences in the cost of equity capital: Do legal institutions and securities regulation matter? Journal of Accounting Research, 44(3):485–531.
- Hail, L. and Leuz, C. 2009. Cost of capital effects and changes in growth expectations around u.s. cross-listings. Journal of Financial Economics, 93.3:428–454.
- Hamilton, J. 1994. Time series analysis, volume 2. Cambridge University Press.
- Hou, K., Van Dijk, M., and Zhang, Y. 2012. The implied cost of capital: A new approach. Journal of Accounting and Economics, 53:504–526.

- 
- Hribar, P. and Jenkins, N. 2004. The effect of accounting restatements on earnings revisions and the estimated cost of capital. Review of Accounting Studies, 9(2):337–356.
- Larocque, S. and Lyle, M. R. 2013. Implied cost of equity capital estimates as predictors of accounting returns. Working Paper, Mendoza College of Business and Kellogg School of Management.
- Lee, C., Ng, D., and Swaminathan, B. 2009. Testing international asset pricing models using implied costs of capital. Journal of Financial and Quantitative Analysis, 44(2):307–335.
- Lewellen, J. 2014. The cross section of expected stock returns. Critical Finance Review.
- Lintner, J. 1965. The valuation of risk assets and the selection of risky investments in stock portfolios and capital budgets. The review of economics and statistics, pages 13–37.
- Pástor, L., Sinha, M., and Swaminathan, B. 2008. Estimating the intertemporal risk–return tradeoff using the implied cost of capital. The Journal of Finance, 63(6):2859–2897.
- Sharpe, W. F. 1964. Capital asset prices: A theory of market equilibrium under conditions of risk. The Journal of Finance, 19(3):425–442.
- Vuolteenaho, T. 2002. What drives firm-level stock returns? Journal of Finance, 57(1):233–264.
- Wang, C. C. 2014. Measurement errors of expected return proxies and the implied cost of capital. Working Paper, Harvard Business School.

## Technical Appendix

### Part A. Ranking Firm-Specific ERP Measurement-Error Variance

Part A of this appendix derives a measure to rank ERP models on the basis of average time-series measurement-error variance, a measure that is ERP-specific and empirically estimable. We call this measure Average Scaled TS Variance ( $AvgSVar^{TS}$ ). Our derivation proceeds in three steps. In [Step 1](#) we decompose a firm's time-series ERP measurement-error variance and define a firm-specific Scaled TS Variance measure. In [Step 2](#) we decompose realized returns and derive an expression for the time-series return-ERP covariance. In [Step 3](#) we show how to estimate Scaled TS Variance using the time-series return-ERP covariance and define the Average Scaled TS Variance.

We make the following assumptions throughout:

- A1** Expected returns ( $er_{i,t+1}$ ), ERP measurement error ( $\omega_{i,t+1}$ ), and realized returns ( $r_{i,t+1}$ ) are jointly covariance stationary.<sup>11</sup>
- A2** Expected-returns forecast errors (or news, i.e.,  $\delta_{i,t+1} = r_{i,t+1} - er_{i,t}$ ) is not ex-ante forecastable, and is not systematically correlated with expected returns (in time-series or cross-section).

#### Step 1. Decomposing a Firm's Time-Series Variance in ERP Measurement Errors and Defining $MVar_i(\omega_{i,t})$

We define an ERP as the sum of the true expected-return and its measurement error ([T1](#)):

$$\widehat{er}_{i,t+1} = er_{i,t+1} + \omega_{i,t+1}. \quad (\text{T1})$$

Taking the time-series variance on both sides of ([T1](#)) and re-organizing terms, a firm  $i$ 's time-series variance in ERP measurement errors can be written as

$$Var_i(\omega_{i,t}) = Var_i(\widehat{er}_{i,t}) + Var_i(er_{i,t}) - 2Cov_i(er_{i,t}, \widehat{er}_{i,t}), \quad (\text{T2})$$

which can be re-expressed as

$$Var_i(\omega_{i,t}) = Var_i(\widehat{er}_{i,t}) - 2[Var_i(er_{i,t}) + Cov_i(er_{i,t}, \omega_{i,t})] + Var_i(er_{i,t}). \quad (\text{T3})$$

The last right-hand-side term, firm  $i$ 's time-series variance in expected returns, does not depend on the choice of ERP model. Therefore, in comparing the time-series variance of ERP measurement errors for firm  $i$ , one needs only to compare the first two terms of ([T3](#)), which we refer to collectively as the Scaled Time-Series Variance of an ERP's measurement

<sup>11</sup>A stochastic vector process  $\{\mathbf{y}_t\}_{t \geq 1}$  is covariance-stationary if (a)  $E[\mathbf{y}_t] = \mu$  for all  $t$ , and (b)  $E(\mathbf{y}_t - \mu)(\mathbf{y}_{t-j} - \mu) = \Sigma_j$  for all  $t$  and any  $j$ . That is, the mean and autocovariances do not depend on the date  $t$ .



errors of firm  $i$ 's expected returns  $[SVar_i(\omega_{i,t})]$ :<sup>12</sup>

$$SVar_i(\omega_{i,t}) = Var_i(\hat{er}_{i,t}) - 2[Var_i(er_{i,t}) + Cov_i(er_{i,t}, \omega_{i,t})]. \quad (\text{T4})$$

Notice that the first right-hand-side term is firm  $i$ 's time-series variance in the ERP, which can be empirically observed. The second right-hand-side term involves unobservables: specifically, firm  $i$ 's variance in expected returns  $[Var_i(er_{i,t})]$  and the time-series covariance between the firm's expected returns and the ERP measurement errors  $[Cov_i(er_{i,t}, \omega_{i,t})]$ . In what follows, we re-express the second term on the right-hand side in terms of variables that can be empirically observed.

## Step 2. Decomposing Realized Returns and Time-Series Return-ERP Covariance

In this step we show that  $Cov_i(r_{i,t+1}, \hat{er}_{i,t}) = Var_i(er_{i,t}) + Cov_i(er_{i,t}, \omega_{i,t})$ . To obtain this result, note that ex-post realized returns is the sum of the ex-ante expected returns and news or forecast errors:

$$r_{i,t+1} = er_{i,t} + \delta_{i,t+1}. \quad (\text{T5})$$

We define  $er_{i,t}$  to be firm  $i$ 's true but unobserved expected returns ( $er_{i,t}$ ) conditional on publicly available information at time  $t$ , capturing all ex-ante predictability (with respect to the information set) in returns. By the property of conditional expectations, it follows that  $er_{i,t}$  is optimal or efficient in the sense of minimizing mean squared errors. This is known as the Prediction Property of conditional expectations (e.g., [Angrist and Pischke, 2008](#)).

It follows from this definition and the property of conditional expectations that a firm's expected returns ( $er_{i,t}$ ) is uncorrelated with its forecast errors ( $\delta_{i,t+1}$ ) in time-series; this is also known as the Decomposition Property of conditional expectations (e.g., [Angrist and Pischke, 2008](#)). Intuitively, if expected returns were correlated with subsequent forecast errors, one could always improve on the expected-return measure by taking into account such systematic predictability, thereby violating the definition of an optimal forecast. This justifies assumption [A2](#).

We can thus write the time-series covariance between returns and ERPs as:

$$\begin{aligned} Cov_i(r_{i,t+1}, \hat{er}_{i,t}) &= Cov_i(er_{i,t} + \delta_{i,t+1}, er_{i,t} + \omega_{i,t}) \\ &= Var_i(er_{i,t}) + Cov_i(er_{i,t}, \omega_{i,t}), \end{aligned} \quad (\text{T6})$$

where the first equality follows from the return decomposition of [\(T5\)](#) and the definition of ERP [\(T1\)](#), and the last equality follows from assumption [A2](#) (i.e., "news is news"), which implies that  $Cov_i(\delta_{i,t+1}, er_{i,t}) = Cov_i(\omega_{i,t+1}, \delta_{i,t+1}) = 0$ .

## Step 3. Estimating $AvgSVar^{TS}$

Substituting [\(T6\)](#) into [\(T3\)](#) and [\(T4\)](#), we obtain:

$$Var_i(\omega_{i,t}) = Var_i(\hat{er}_{i,t}) - 2Cov_i(r_{i,t+1}, \hat{er}_{i,t}) + Var_i(er_{i,t}), \quad (\text{T7})$$

<sup>12</sup>Note that  $Var_i(\omega_{i,t}) = SVar_i(\omega_{i,t}) + Var_i(er_{i,t})$ .

so that

$$SVar_i(\omega_{i,t}) = Var_i(\hat{er}_{i,t}) - 2Cov_i(r_{i,t+1}, \hat{er}_{i,t}). \quad (\text{T8})$$

The first term of  $SVar_i$  shows that, all else equal, an ERP's measurement-error variance is increasing in the variance of the ERP. The second term of  $SVar_i$  shows that, all else equal, an ERP's measurement-error variance is decreasing in the degree to which ERPs predict future returns in time-series.

Notice that (T8) expresses  $SVar_i(\omega_{i,t})$  in terms of two empirically observable variables  $\{\hat{er}_{i,t}, r_{i,t+1}\}$ . These variables can be computed empirically, with consistency achieved under standard regularity conditions.<sup>13</sup> Our empirical tests compute, for each ERP and each firm, the relative error-variance measure using (T8), and assess the time-series performance of ERPs based on the average of  $SVar_i$  across the  $N$  firms in our sample:

$$AvgSVar^{TS} = \frac{1}{N} \sum_i SVar_i(\omega_{i,t}). \quad (\text{T9})$$

Notice also that  $SVar_i(\omega_{i,t}) \geq -Var_i(er_{i,t})$ , because  $Var_i(\omega_{i,t}) = SVar_i(\omega_{i,t}) + Var_i(er_{i,t})$  and  $Var_i(\omega_{i,t}) \geq 0$ . Therefore,  $-Var_i(er_{i,t})$  is the minimum bound for our empirically estimable Scaled Time-Series Variance measure. In other words, if we have an ICC that measures expected returns perfectly, then  $SVar_i(\omega_{i,t}) = -Var_i(er_{i,t})$ .<sup>14</sup>

## Part B. Ranking Cross-Sectional ICC Measurement-Error Variance

Here derive a measure to rank ERP models on the basis of their average cross-sectional measurement-error variance. We call this measure Average Scaled CS Variance ( $AvgSVar^{TS}$ ). Our derivation proceeds in two steps. In [Step 1](#), we decompose a firm's cross-sectional ERP measurement-error variance and define our cross-section-specific Scaled CS Variance measure. In [Step 2](#), we show how to estimate Average Scaled CS Variance using the average cross-sectional return-ERP covariance. We make the same assumptions as in [Part A](#).

---

<sup>13</sup>The following regularity conditions are sufficient to ensure that sample time-series variances and covariances will converge in probability to population variances and covariances ([Hamilton, 1994](#)): for a covariance stationary stochastic process  $\{Y_t\}_{t \geq 1}$ , if there exists an MA( $\infty$ ) representation ( $Y_t = \sum_{j=0}^{\infty} \psi_j \epsilon_j$ ) where the MA coefficients are absolutely summable ( $\sum_{j=0}^{\infty} |\psi_j| < \infty$ ) and  $\{\epsilon_t\}_{t \geq 1}$  is an iid sequence with  $E|\epsilon_t|^r < \infty$ , then the sample covariance converges to the population covariance in probability:

$$\frac{1}{T} \sum_{t=1}^T (Y_t - \bar{Y}_T)(Y_{t-k} - \bar{Y}_T) \xrightarrow{p} E(Y_t - \mu)(Y_{t-k} - \mu), \text{ where } \bar{Y}_T = \frac{1}{T} \sum_{t=1}^T Y_t \text{ and } \mu = E(Y_t).$$

<sup>14</sup>Note that by comparing the resulting  $SVar$  for various ICC estimates, we only achieve a "relative" performance assessment; we cannot conduct an "absolute" comparison to the true expected-returns measure without imposing further structure on the time-series process governing expected returns. In [Part C](#) of this Technical Appendix, we illustrate how this can be done under the assumption that both the measurement error and the true expected returns follow AR(1) processes.

### Step 1. Decomposing an ERP's Cross-Sectional Measurement-Error Variance and Defining $SVar_t(\omega_{i,t})$

As in the case of time-series, the cross-sectional variance in ERP measurement errors can be written as

$$Var_t(\omega_{i,t}) = Var_t(\hat{er}_{i,t}) + Var_t(er_{i,t}) - 2Cov_t(er_{i,t}, \hat{er}_{i,t}), \quad (C1)$$

which can be re-expressed as

$$Var_t(\omega_{i,t}) = Var_t(\hat{er}_{i,t}) - 2[Var_t(er_{i,t}) + Cov_t(er_{i,t}, \omega_{i,t})] + Var_t(er_{i,t}). \quad (C2)$$

The final right-hand-side term, the cross-sectional variance in expected returns at time  $t$ , does not depend on the choice of ERP model. Therefore in comparing the cross-sectional variance of ERP measurement errors at time  $t$ , one needs only to compare the first two terms of (C2), which we refer to collectively as the Scaled CS Variance of an ERP's measurement errors [ $SVar_t(\omega_{i,t})$ ]:<sup>15</sup>

$$SVar_t(\omega_{i,t}) = Var_t(\hat{er}_{i,t}) - 2[Var_t(er_{i,t}) + Cov_t(er_{i,t}, \omega_{i,t})]. \quad (C3)$$

Notice that the first right-hand-side term is the cross-sectional variance in the ERP, which can be empirically estimated. The second right-hand-side term involves unobservables — specifically, the cross-sectional variance in expected returns [ $Var_t(er_{i,t})$ ], and the cross-sectional covariance between the firm's expected returns and the ERP measurement errors [ $Cov_t(er_{i,t}, \omega_{i,t})$ ].

### Step 2. Defining and Estimating $AvgSVar^{CS}$

In our empirical tests, we assess the cross-sectional performance of ERPs based on the average of  $SVar_t$  across the  $T$  cross-sections in our sample:

$$AvgSVar^{CS} = \frac{1}{T} \sum_t Var_t(\hat{er}_{i,t}) - 2[Var_t(er_{i,t}) + Cov_t(er_{i,t}, \omega_{i,t})]. \quad (C4)$$

To estimate  $AvgSVar^{TS}$ , we note that the average cross-sectional covariance between returns and ERPs can be expressed as:

$$\begin{aligned} \frac{1}{T} \sum_t Cov_t(r_{i,t+1}, \hat{er}_{i,t}) &= \frac{1}{T} \sum_t Cov_t(er_{i,t} + \delta_{i,t+1}, er_{i,t} + \omega_{i,t}) \\ &= \frac{1}{T} \sum_t [Var_t(er_{i,t}) + Cov_t(er_{i,t}, \omega_{i,t})], \end{aligned} \quad (C5)$$

where the first equality follows from the realized returns decomposition (T5) and the definition of expected-returns proxy (T1), and the last equality follows from the assumption (A2)

<sup>15</sup>Note that  $Var_t(\omega_{i,t}) = SVar_t(\omega_{i,t}) + Var_t(er_{i,t})$

that “news” cannot exhibit systematic forecastability.<sup>16</sup>

## Part C. Completing the Evaluative Framework

Here, we develop further intuition on the importance of the time-series error variance as a performance criterion for ERPs, and show how this approach augments prior studies that rely solely on cross-sectional realized returns as a performance benchmark.

### 1. What does a “zero-error-variance” ERP look like?

We begin by demonstrating that our benchmark is the true expected-return, and then show why several closely related ERPs do not result in zero-error-variance.

Suppose we have the perfect proxy for expected returns, so that  $\widehat{er}_{i,t} = er_{i,t} \forall i, t$ . Then from (T3) and (C2), since  $\omega_{i,t} = 0 \forall i, t$ , we have:

$$\begin{aligned} \text{Var}_i(\omega_{i,t}) &= \text{Var}_i(er_{i,t}) - 2[\text{Var}_i(er_{i,t}) + \text{Cov}_i(er_{i,t}, 0)] + \text{Var}_i(er_{i,t}) = 0, \text{ and} \\ \text{Var}_t(\omega_{i,t}) &= \text{Var}_t(er_{i,t}) - 2[\text{Var}_t(er_{i,t}) + \text{Cov}_t(er_{i,t}, 0)] + \text{Var}_t(er_{i,t}) = 0. \end{aligned}$$

As expected, an ERP that measures true expected-return without error will yield zero error variance. In fact, it should be clear from the above that any ERP that always differs from the true expected returns by the same fixed constant (i.e., takes the form  $\widehat{er}_{i,t} = er_{i,t} + C$  for some constant  $C$ ) will also have zero error variance (in time-series and in the cross-section). Our empirically estimable  $\text{AvgSVar}^{TS}$  and  $\text{AvgSVar}^{TS}$  measures will achieve their minimum bounds with such an ERP: i.e.,  $-\text{Var}_i(er_{i,t})$  and  $-\text{Var}_t(er_{i,t})$  respectively.

### 2. What about a “noisy but unbiased” ERP?

What about an ERP that is “on average right” (e.g.,  $\widehat{er}_{i,t} = er_{i,t} + \epsilon_{i,t}$  for some white noise process  $\epsilon_{i,t}$ )? Since the measurement error is simply white noise, i.e.,  $\omega_{i,t} = \widehat{er}_{i,t} - er_{i,t} = \epsilon_{i,t}$ , we have:

$$\begin{aligned} \text{Var}_i(\omega_{i,t}) &= \text{Var}_i(\epsilon_{i,t}) \text{ for any } i, \text{ and} \\ \text{Var}_t(\omega_{i,t}) &= \text{Var}_t(\epsilon_{i,t}) \text{ for any } t. \end{aligned}$$

This result is quite intuitive: for an unbiased ERP estimate that measures expected-return with random noise, the variance in measurement error is simply the variance of the white noise.

---

<sup>16</sup>Note that for any given cross-section it might be possible for realized news and measurement errors to be correlated, but this cannot be true systematically (i.e., across many cross-sections) by the definition of news.

### 3. Using a fixed constant as an ERP

Next, consider an ERP that is a fixed constant  $C$  (e.g., 7 percent each year). In this case,  $\omega_{i,t} = C - er_{i,t}$ , so that:

$$\begin{aligned} \text{Var}_i(\omega_{i,t}) &= \text{Var}_i(C - er_{i,t}) = \text{Var}_i(er_{i,t}), \text{ and} \\ \text{Var}_t(\omega_{i,t}) &= \text{Var}_t(C - er_{i,t}) = \text{Var}_t(er_{i,t}). \end{aligned}$$

This result is also quite intuitive: if the ERP estimate is a fixed constant, variations in measurement errors will be driven entirely by variations in expected returns.

### 4. Using future realized returns as ERPs

Finally, it is instructive to consider how one-period-ahead realized returns would fare in our framework — i.e., what measurement-error variances would look like if we used future realized returns as a proxy for expected returns. Note that doing so would violate a key assumption that underpins the derivation of equation (T7) or (T8), since future realized returns are highly correlated with future news. We therefore cannot apply these equations to this expected-return proxy.

Using equation (T5) instead, we show that in this case the measurement error is simply the forecast error or news, i.e.,  $\omega_{i,t} = er_{i,t} - r_{i,t+1} = -\delta_{i,t+1}$ , so that:

$$\begin{aligned} \text{Var}_i(\omega_{i,t}) &= \text{Var}_i(\delta_{i,t+1}), \text{ and} \\ \text{Var}_t(\omega_{i,t}) &= \text{Var}_t(\delta_{i,t+1}). \end{aligned}$$

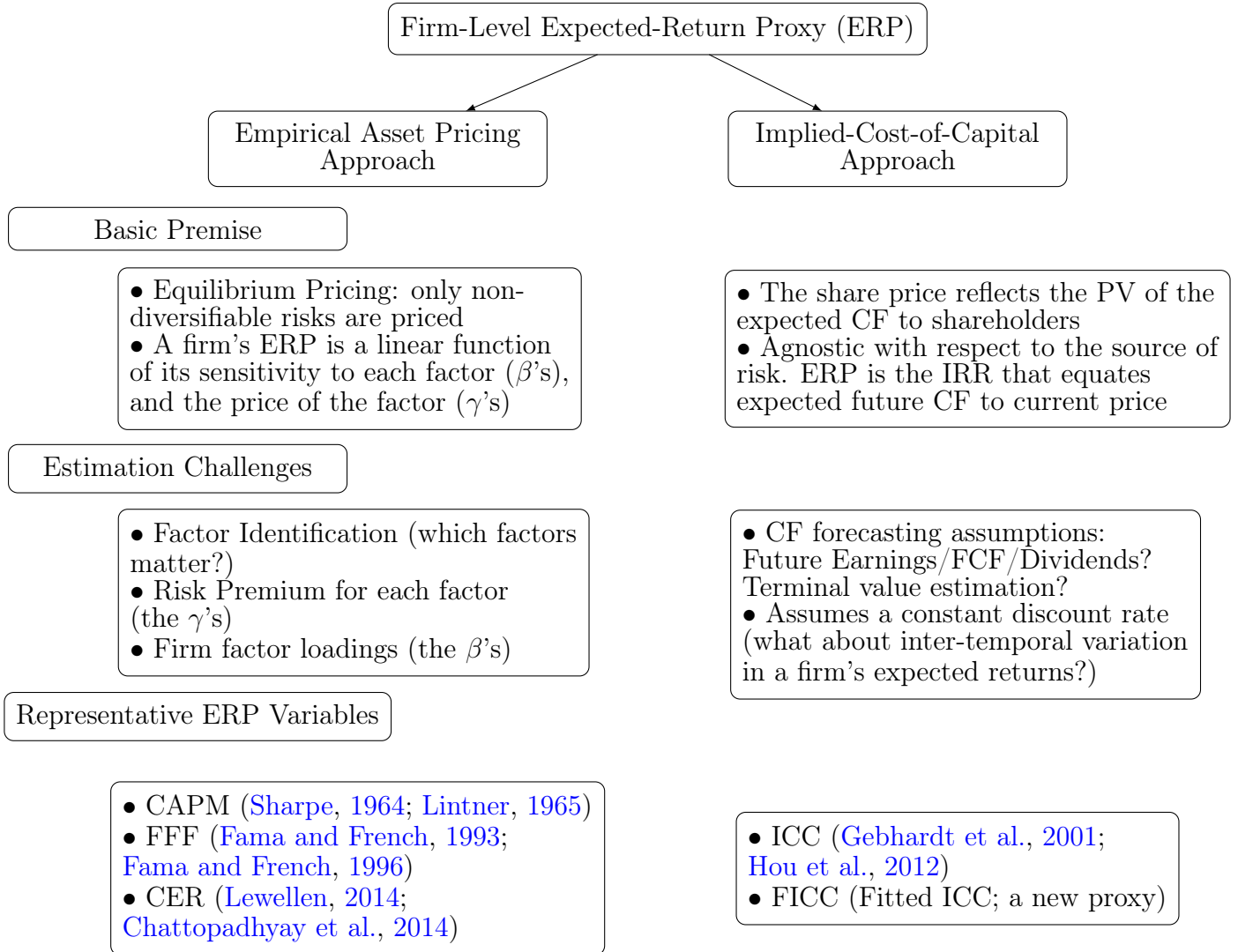
That is, the measurement-error variance of future realized returns is essentially the variance of the news shocks.

This expression highlights a key weakness in the use of realized returns as a proxy for expected returns. If stock prices are subject to large unpredictable shocks (whether to cash-flows or to investor sentiment), then the variance in measurement errors for realized returns will be large, making realized returns poor proxies for expected returns in our framework. This is, of course, the main motivation for alternative ERP estimates in the first place.<sup>17</sup>

---

<sup>17</sup>However, as Wang (2014) notes, realized returns may still be a desirable alternative in regression settings.

## Appendix I. Family Tree of Expected-Return Proxies



## Appendix II. Summary of Expected-Return Proxies (ERP)

This table summarizes of five expected-return proxies (ERPs). For each of the five proxies (listed in the first column), the table provides a short description (in the second column), explains how the factor risk-premium is estimated (in the third column), explains how the factor loadings on risk factors are estimated (in the fourth column), and cites related prior studies (in the last column). A full description of each proxy appears in Section IV.

Variable	Description	Factor risk-premium (Gamma) estimation	Factor-loading (Beta) estimation	Related prior studies
CAPM	One-factor expected-return proxy based on the Capital Asset Pricing Model (CAPM)	Based on the realized risk premium over the preceding 12 months	Based on time-series regressions using realized returns over the preceding 60 months	<a href="#">Sharpe (1964)</a> ; <a href="#">Lintner (1965)</a>
FFF	A four-factor expected-return proxy based on realized returns and each firm's estimated sensitivity to four-factors (MKT, SMB, HML, and UMD)	Based on each factor's realized risk premium over the preceding 12 months	Based on time-series regressions using realized returns over the preceding 60 months	<a href="#">Fama and French (1993)</a> ; <a href="#">Fama and French (1996)</a>
CER	Characteristic-based expected-return proxy whereby a firm's exposure to a factor is simply its scaled rank on that characteristic (Size, BTM, Momentum)	Based on ten-year average Fama-MacBeth coefficients for each factor	Based on current year firm characteristics	<a href="#">Lewellen (2014)</a> ; <a href="#">Chattopadhyay et al. (2014)</a>
ICC	The internal rate of return (IRR) that equates a firm's forecasted cash-flows to its current market price	Not applicable	Not applicable	<a href="#">Gebhardt et al. (2001)</a> ; <a href="#">Hou et al. (2012)</a>
FICC	Characteristic-based expected-return proxy whereby each firm's ICC is "fitted" to a set of firm characteristics	Not applicable	Based on current year firm characteristics	New proxy

## Appendix III. Regression Coefficients from Earnings Forecasts Regressions

This table reports the average regression coefficients and their time-series t-statistics from annual pooled regressions of one-year-ahead through three-year-ahead earnings on a set of variables hypothesized to capture differences in expected earnings across firms. Specifically, for each year  $t$  between 1970 and 2011, we estimate the following pooled cross-sectional regression using the previous ten years (six years minimum) of data:

$$E_{j,t+\tau} = \beta_0 + \beta_1 EV_{j,t} + \beta_2 TA_{j,t} + \beta_3 DIV_{j,t} + \beta_4 DD_{j,t} + \beta_5 E_{j,t} + \beta_6 NEGE_{j,t} + \beta_7 ACC_{j,t} + \epsilon_{j,t+\tau}$$

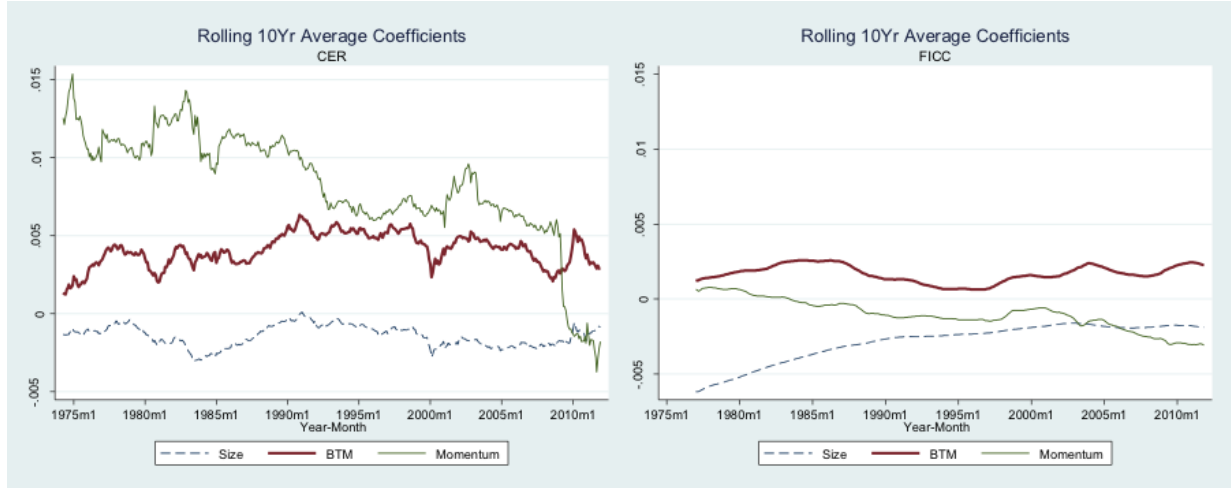
where  $E_{j,t+\tau}$  ( $\tau = 0, 1, 2, \text{ or } 3$ ) denotes the earnings before extraordinary items of firm  $j$  in year  $t + \tau$ , and all explanatory variables are measured to the end of year  $t$ :  $EV_{j,t}$  is the enterprise value of the firm (defined as total assets plus the market value of equity minus the book value of equity),  $TA_{j,t}$  is the total assets,  $DIV_{j,t}$  is the dividend payment,  $DD_{j,t}$  is a dummy variable that equals 0 for dividend payers and 1 for non-payers,  $NEGE_{j,t}$  is a dummy variable that equals 1 for firms with negative earnings (0 otherwise), and  $ACC_{j,t}$  is total accruals scaled by total assets. Total accruals are calculated as the change in current assets plus the change in debt in current liabilities minus the change in cash and short-term investments and minus the change in current liabilities.  $R-Sq$  is the time-series average R-squared from the annual regressions.

Years									
Ahead	Intercept	V	T	DIV	DD	E	NEGE	ACC	R-Sq
1	2.452 (5.32)	0.010 (45.06)	-0.008 (-33.79)	0.330 (38.21)	-2.515 (-3.47)	0.749 (162.23)	1.034 (2.28)	-0.016 (-8.67)	0.849
2	4.268 (6.54)	0.013 (41.67)	-0.010 (-28.46)	0.490 (39.77)	-3.617 (-3.68)	0.669 (98.16)	1.938 (2.62)	-0.020 (-7.95)	0.790
3	15.970 (23.39)	0.000 (7.64)	0.002 (-1.43)	0.663 (46.91)	-10.350 (-9.29)	0.295 (45.43)	-0.418 (-0.67)	-0.008 (-2.95)	0.459



### Figure 1. Ten-Year Rolling Slope Estimates

This figure reports the ten-year rolling-average Fama-MacBeth slope estimates from cross-sectional regressions of expected-return proxies on firms' size, book-to-market (BTM), and return momentum. The left-hand-side panel reports rolling average coefficients using realized returns as the dependent variable; the right-hand-side panel uses the implied-cost-of-capital (ICC) as the dependent variable. A full description of each proxy appears in Section IV.



## Figure 2. Efficient Frontier of Firm-Specific Expected-Return Proxies

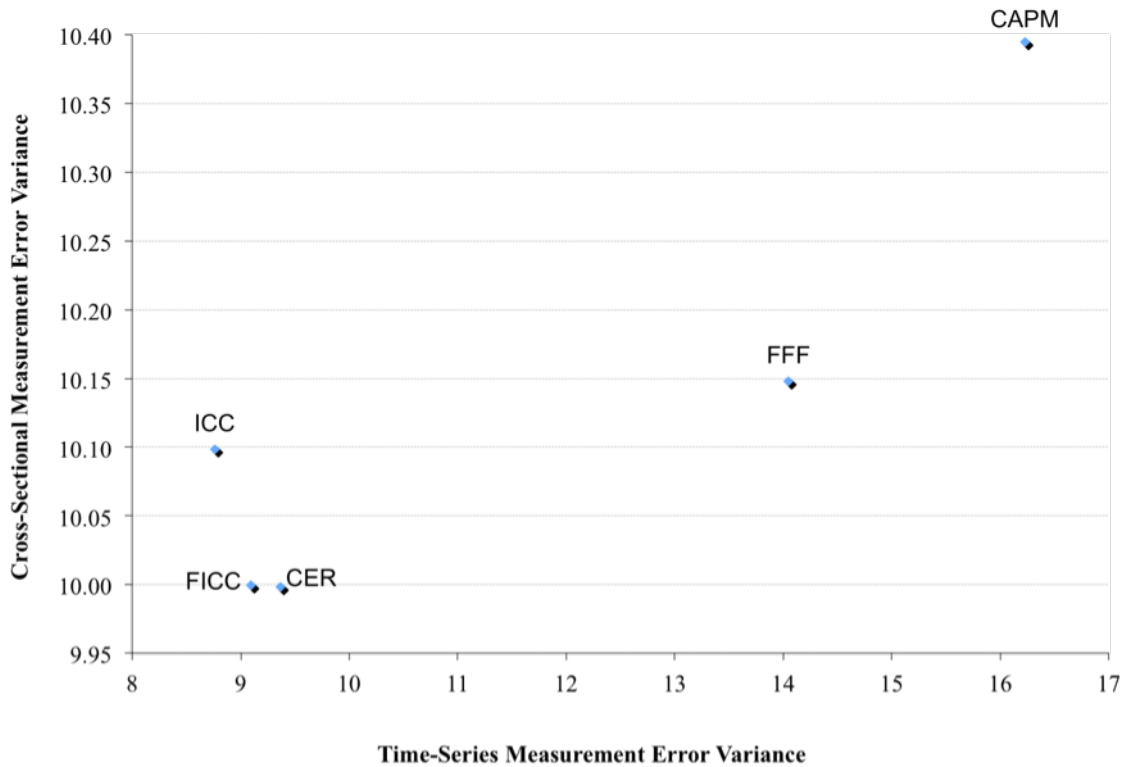
This figure plots the average cross-sectional and time-series scaled measurement-error variances for each of five monthly expected-return proxies. ICC is the implied-cost-of-capital following [Gebhardt et al. \(2001\)](#); CER is a characteristic-based expected-return proxy derived from historical cross-sectional regression of realized returns on a firm's size, book-to-market, and return momentum where historically estimated ten-year average Fama-MacBeth coefficients are applied to current firm characteristics; FICC is analogously defined from a regression of ICC on a firm's size, book-to-market, and return momentum; CAPM is the firm's expected-return derived from the Capital Asset Pricing Model; and FFF is the four-factor model using the market, small-minus-big (SMB), high-minus-low (HML), and up-minus-down (UMD) factors. A full description of each proxy appears in Section IV. Scaled measurement-error variances are calculated for each unique firm/expected-return proxy pair as follows:

$$AvgSVar^{CS} = \frac{1}{T} \sum_t Var_t(\hat{e}r_{i,t}) - 2Cov_t(r_{i,t+1}, \hat{e}r_{i,t})$$

where  $Var_t(\hat{e}r_{i,t})$  is a given ERP's cross-sectional variance at time  $t$  and  $Cov_t(r_{i,t+1}, \hat{e}r_{i,t})$  is the cross-sectional covariance between firms' ERPs at time  $t$  and their realized returns in period  $t + 1$ . Time-series measurement-error variances are defined analogously when summarized over all firms in a given calendar month, spanning 418 months from 1977 through 2011 as follows:

$$AvgSVar^{TS} = \frac{1}{N} \sum_i SVar_i(\omega_{i,t})$$

where  $SVar_i(\omega_{i,t})$  is the scaled measurement-error variance firm  $i$ . Firm-specific scaled measurement-error variances are calculated based on a sample of 12,022 unique firms that meet our data requirements.



**Table 1. Monthly Expected-Return Proxies by Year**

This table reports the median of monthly expected-return proxies. ICC is the implied-cost-of-capital following [Gebhardt et al. \(2001\)](#); CER is a characteristic-based expected-return proxy derived from a historical regression of realized returns on a firm's size, book-to-market, and return momentum where historically estimated ten-year average Fama-MacBeth coefficients are applied to current firm characteristics; FICC is analogously defined from a regression of ICC on a firm's size, book-to-market, and return momentum; CAPM is the firm's expected-return derived from the Capital Asset Pricing Model; and FFF is the four-factor model using the market, small-minus-big (SMB), high-minus-low (HML), and up-minus-down (UMD) factors. A full description of each proxy appears in Section IV. We compute a firm-specific expected-return estimate for each stock in our sample based on the stock price and publicly available information at the conclusion of each calendar month, where the expected-return corresponds to the following month. Proxies are treated as missing if they are either below zero or above 100 percent.

Year	Observations	ICC	CER	FICC	CAPM	FFF
1977	31,012	1.55%	0.86%	2.35%	-0.06%	1.11%
1978	30,930	1.46%	0.77%	2.16%	2.87%	2.85%
1979	31,707	1.45%	0.90%	2.07%	3.38%	3.56%
1980	33,339	1.43%	1.35%	1.95%	4.53%	3.52%
1981	33,688	1.28%	1.43%	1.78%	0.56%	1.18%
1982	34,673	1.36%	1.11%	1.86%	0.93%	1.11%
1983	36,503	1.10%	2.19%	1.55%	2.50%	3.03%
1984	36,535	1.10%	1.78%	1.58%	0.54%	-0.25%
1985	38,481	1.04%	1.86%	1.54%	2.11%	1.93%
1986	39,268	0.93%	1.81%	1.43%	1.90%	1.24%
1987	39,076	0.88%	1.67%	1.42%	2.25%	1.61%
1988	40,754	0.95%	1.41%	1.53%	1.08%	1.22%
1989	42,434	0.97%	1.52%	1.45%	1.75%	1.61%
1990	41,607	1.08%	1.15%	1.49%	-0.58%	-0.67%
1991	41,266	1.00%	1.25%	1.40%	2.32%	2.88%
1992	41,372	0.85%	1.39%	1.30%	0.98%	1.58%
1993	43,232	0.76%	1.05%	1.24%	1.31%	2.01%
1994	50,544	0.79%	1.09%	1.27%	1.01%	0.38%
1995	55,627	0.79%	1.10%	1.25%	2.92%	2.56%
1996	58,724	0.77%	1.12%	1.19%	2.09%	2.07%
1997	60,221	0.72%	1.13%	1.13%	2.60%	2.13%
1998	61,487	0.74%	1.16%	1.09%	2.52%	1.51%
1999	59,517	0.79%	1.15%	1.12%	1.55%	0.84%
2000	57,352	0.80%	1.57%	1.09%	-0.45%	0.01%
2001	54,367	0.76%	1.46%	1.09%	0.54%	1.66%
2002	51,465	0.70%	1.28%	1.06%	-0.16%	-0.43%
2003	48,234	0.68%	1.21%	1.04%	2.29%	3.44%
2004	47,406	0.60%	1.40%	0.90%	2.35%	2.22%
2005	46,816	0.63%	1.21%	0.97%	2.02%	1.25%
2006	47,107	0.63%	1.16%	0.96%	2.33%	1.97%
2007	46,956	0.66%	1.03%	0.99%	1.51%	0.64%
2008	46,291	0.82%	0.77%	1.20%	-1.88%	-1.29%
2009	43,379	0.87%	0.62%	1.32%	3.69%	3.55%
2010	43,035	0.75%	0.59%	1.08%	2.40%	2.62%
2011	35,125	0.74%	0.62%	1.09%	0.04%	-0.52%
Mean	44,272	0.93%	1.23%	1.37%	1.59%	1.55%
Median	43,035	0.82%	1.16%	1.27%	1.90%	1.61%
Std	8,919	0.27%	0.37%	0.36%	1.33%	1.27%
Min	30,930	0.60%	0.59%	0.90%	-1.88%	-1.29%
Max	61,487	1.55%	2.19%	2.35%	4.53%	3.56%

**Table 2. Correlation between Expected-Return Proxies**

This table reports the average monthly Spearman correlations among the five expected-return proxies. ICC is the implied-cost-of-capital following [Gebhardt et al. \(2001\)](#); CER is a characteristic-based–return proxy derived from a historical regression of realized returns on a firm’s size, book-to-market, and return momentum where historically estimated ten-year average Fama-MacBeth coefficients are applied to current firm characteristics; FICC is analogously defined from a regression of ICC on a firm’s size, book-to-market, and return momentum; CAPM is the firm’s expected-return derived from the Capital Asset Pricing Model; and FFF is the four-factor model using the market, small-minus-big (SMB), high-minus-low (HML), and up-minus-down (UMD) factors. A full description of each proxy appears in Section IV. P-values are shown in parentheses corresponding to one-tailed tests under the alternative hypothesis that expected-return proxies should be positively correlated.

	ICC	CER	FICC	CAPM	FFF
ICC		0.499 (0.00)	0.648 (0.00)	-0.172 (0.00)	-0.117 (0.00)
CER	0.499 (0.00)		0.660 (0.00)	-0.101 (0.00)	-0.060 (0.00)
FICC	0.648 (0.00)	0.660 (0.00)		-0.204 (0.00)	-0.117 (0.00)
CAPM	-0.172 (0.00)	-0.101 (0.00)	-0.204 (0.00)		0.603 (0.00)
FFF	-0.117 (0.00)	-0.060 (0.00)	-0.117 (0.00)	0.603 (0.00)	

**Table 3. Cross-Sectional Scaled Measurement-Error Variances**

This table presents descriptive statistics for the cross-sectional variances of scaled measurement errors (multiplied by 100) of five expected-return proxies. Scaled measurement-error variances are calculated for each unique firm-expected/return proxy pair using Equation (9) as follows:

$$AvgSVar^{CS} = \frac{1}{T} \sum_t Var_t(\hat{e}r_{i,t}) - 2Cov_t(r_{i,t+1}, \hat{e}r_{i,t}).$$

where  $Var_t(\hat{e}r_{i,t})$  is a given ERP's cross-sectional variance at time  $t$  and  $Cov_t(r_{i,t+1}, \hat{e}r_{i,t})$  is the cross-sectional covariance between firms' ERPs at time  $t$  and their realized returns in period  $t + 1$ . Panel A reports summary statistics for the error variance from each model, using data from a sample of 418 calendar months during our 1977–2011 sample period. Table values in this panel represent descriptive statistics for the error variance from each expected-return proxy computed across these 418 months. ICC is the implied-cost-of-capital following Gebhardt et al. (2001); CER is a characteristic-based expected-return proxy derived from a historical regression of realized returns on a firm's size, book-to-market, and return momentum where historically estimated ten-year average Fama-MacBeth coefficients are applied to current firm characteristics; FICC is analogously defined from a regression of ICC on a firm's size, book-to-market, and return momentum; CAPM is the firm's expected-return derived from the Capital Asset Pricing Model; and FFF is the four-factor model using the market, small-minus-big (SMB), high-minus-low (HML), and up-minus-down (UMD) factors. A full description of each proxy proxy in Section IV. Panel B reports t-statistics based on Newey-West-adjusted standard errors corresponding to the pair-wise comparisons of average cross-sectional scaled measurement-error variances within the sample of 418 months used in Panel A. Table values are negative (positive) when the expected-return proxy displayed in the leftmost column has a larger (smaller) scaled measurement-error variance than the expected-return proxy displayed in the topmost row. \*, \*\*, and \*\*\* indicate two-tailed significance at the 10%, 5%, and 1% levels respectively.

<b>Panel A. Cross-Sectional Measurement-Error Variance</b>					
	Mean	P25	Median	P75	STD
ICC	10.099	9.960	10.073	10.408	0.101
CER	9.998	9.938	9.998	10.059	0.021
FICC	10.000	9.928	10.001	10.057	0.022
CAPM	10.148	9.772	10.085	11.114	0.249
FFF	10.395	9.742	10.215	13.524	0.672

<b>Panel B. t-Statistics of Differences in Variances</b>					
	ICC	CER	FICC	CAPM	FFF
ICC		-0.100 (-5.58)	-0.099 (-5.48)	0.049 (2.45)	0.296 (4.44)
CER	0.100 (5.58)		0.001 (1.32)	0.150 (9.71)	0.396 (6.17)
FICC	0.099 (5.48)	-0.001 (-1.32)		0.148 (9.57)	0.395 (6.13)
CAPM	-0.049 (-2.45)	-0.150 (-9.71)	-0.148 (-9.57)		0.247 (4.56)
FFF	-0.296 (-4.44)	-0.396 (-6.17)	-0.395 (-6.13)	-0.247 (-4.56)	

**Table 4. Time-Series Scaled Measurement-Error Variances**

This table presents descriptive statistics for the time-series variances of scaled measurement errors (multiplied by 100) of five expected-return proxies. Scaled measurement-error variances are calculated for each unique firm-expected-return proxy pair using Equation (5) as follows:

$$AvgSVar^{TS} = \frac{1}{N} \sum_i SVar_i(\omega_{i,t})$$

where  $SVar_i(\omega_{i,t})$  is the scaled measurement-error variance of firm  $i$ . Panel A reports summary statistics for the error variance from each model, using a sample of 12,022 unique firms with a minimum of 20 (not necessarily consecutive) months of data during our 1977-2011 sample period. Table values in this panel represent descriptive statistics for the error variance from each expected-return proxy computed across these 12,022 firms. ICC is the implied-cost-of-capital following [Gebhardt et al. \(2001\)](#); CER is a characteristic-based expected-return proxy derived from a historical regression of realized returns on a firm's size, book-to-market, and return momentum where historically estimated ten-year average Fama-MacBeth coefficients are applied to current firm characteristics; FICC is analogously defined from a regression of ICC on a firm's size, book-to-market, and return momentum; CAPM is the firm's expected-return derived from the Capital Asset Pricing Model; and FFF is the four-factor model using the market, small-minus-big (SMB), high-minus-low (HML), and up-minus-down (UMD) factors. A full description of each proxy appears in Section IV. Panel B reports t-statistics corresponding to the pair-wise comparisons of average firm-specific scaled measurement-error variances within the sample of 12,022 firms used in Panel A. Table values are negative (positive) when the expected-return proxy displayed in the leftmost column has a larger (smaller) scaled measurement-error variance than the expected-return proxy displayed in the topmost row. \*, \*\*, and \*\*\* indicate two-tailed significance at the 10%, 5%, and 1% levels, respectively.

<b>Panel A. Time-Series Measurement-Error Variance</b>					
	Mean	P25	Median	P75	STD
ICC	8.759	8.704	9.600	10.102	0.634
CER	9.368	8.766	9.649	10.139	0.043
FICC	9.095	8.810	9.506	9.847	0.023
CAPM	14.048	7.226	12.387	19.709	0.699
FFF	16.226	3.363	12.828	25.203	1.270

<b>Panel B. t-Statistics of Differences in Variances</b>					
	ICC	CER	FICC	CAPM	FFF
ICC		0.609 (3.00)	0.336 (1.66)	5.288 (13.72)	7.467 (11.25)
CER	-0.609 (-3.00)		-0.273 (-7.35)	4.680 (13.69)	6.858 (11.04)
FICC	-0.336 (-1.66)	0.273 (7.35)		4.953 (14.27)	7.131 (11.43)
CAPM	-5.288 (-13.72)	-4.680 (-13.69)	-4.953 (-14.27)		2.178 (4.17)
FFF	-7.467 (-11.25)	-6.858 (-11.04)	-7.131 (-11.43)	-2.178 (-4.17)	