A Market-Based Cost of Capital*

Niels Joachim Gormsen Kilian Huber Theis Ingerslev Jensen

March 2026

Please click here for the latest version

Abstract

How should firms set their cost of capital and does their choice matter for their market value? A classical intuition suggests that firms should base their cost of capital on the opportunity cost in financial markets if they want to maximize their value. This approach is taught to firm managers and used as a standard assumption in models of firm behavior. But the cost of capital is unobserved and no existing method estimates it without bias; no dataset reports an unbiased cost of capital that can be used by practitioners and academics; and no evidence validates the classical intuition that the cost of capital determines firm market values. We introduce a method based on machine learning that estimates firm-level cash flow forecasts and from those a "market-based cost of capital" (MCC). The resulting MCC is a strong and unbiased predictor of future financial asset returns, meaning it indeed measures the opportunity cost of capital. In practice, most firms' perceptions of the cost of capital differ substantially from the MCC. In the data, firms using a perceived cost of capital closer to the MCC achieve higher market values.

*We are grateful for helpful comments from Paul Fontanier, Stefano Giglio, Bryan Kelly, Lasse Pedersen and Alp Simsek; and conference and seminar participants at 2025 HBS Juniors Conference, UCLA Anderson, Yale (accounting), Yale (finance). We acknowledge support from the Danish National Research Foundatino (grant no. DNRF167 and DNRF 193). Gormsen is at The University of Chicago and Copenhagen Business School; www.voices.uchicago.edu/gormsen. Huber is at The University of Chicago; www.kilianhuber.org. Jensen is at Yale University; www.tijensen.com.

1 Introduction

A fundamental principle in economics is that firms should determine the opportunity cost of an investment before undertaking it. The intuition is simple. The opportunity cost of an investment—known as the "cost of capital"—is the expected return on an alternative, external investment that is similar to the firm's investment. Firms should only undertake an investment if it is expected to yield a return above that cost of capital. By following this rule, the firm maximizes its economic profits and thereby its market value.

Many researchers and practitioners take this principle as given. Most academic models assume that firms make investment decisions based on the cost of capital, including neoclassical frameworks like Tobin's Q and Modigliani-Miller, models with financial frictions, and behavioral models with irrational expectations. In the real world, firm executives attempt to incorporate the cost of capital in their decision-making, inspired by business school teaching and consulting firms.

Despite the importance of the cost of capital in academia and practice, fundamental issues remain. First, the cost of capital is not directly observed, since it refers to the hypothetical expected return on a similar external investment. There is no general method that estimates these expected returns out-of-sample without bias (i.e., without predictable errors). By extension, there is no dataset measuring an unbiased cost of capital at the firm level, so it is unclear which values can inform academic models and firm decisions. Finally, there exists no evidence validating the classical intuition that firms improve their market value when they use an appropriate cost of capital.

We contribute in three ways. First, we develop a method that uses machine learning to forecast expected cash flows and then solves for the cost of capital. The method estimates a firm's "market-based cost of capital" (MCC), which captures the expected return in financial markets on the firm's assets and, under the law of one price, the opportunity cost of capital for an investment that is similar to the firm's existing investments.

Second, we apply the method to produce a firm-level dataset of the estimated MCC. The MCC is unbiased, in the sense that it predicts firm-level returns in financial markets without predictable errors, and powerful, in the sense that it captures meaningful variation across firms. The MCC differs strongly from existing biased proxies for the cost of capital, including those based on the Capital Asset Pricing Model (CAPM), the implied cost of capital (ICC), and the simplest measure – interest rates.

Third, we provide evidence on the cost of capital and market values. We find that firm

equity market value is greater when a firm's perceived cost of capital (which it actually uses in its internal decisions) is closer to its MCC. Intuitively, firms with "too low" a perceived cost of capital inadvertently undertake value-destroying investments, which lowers their market value; firms with "too high" a cost of capital inadvertently pass on investments that would have increased the value of the firm. In addition, the equity market responds more positively to a firm announcing new investments when the firm's perceived cost of capital is "too high" relative to the MCC, consistent with these firms investing too little. The findings support the classical intuition that the cost of capital is a crucial input into firms' decisions and that firms using an unbiased cost of capital, like the MCC, achieve higher market values.

We begin the paper by defining the object of interest: a firm's cost of capital, also known as weighted average cost of capital (WACC). It captures the opportunity cost of the firm's capital (i.e., what investors could earn when deploying the firm's capital externally, outside the firm). Firm executives are taught to use this firm-level cost of capital in their investment decisions. In classical frameworks, firms need to know this cost of capital to maximize their market value. It also influences the allocation of capital in essentially all models in economics and finance.

Following the opportunity cost logic, a firm's cost of capital is the expected return on external financial assets that have the same characteristics (e.g., same term and risk) as the firm's existing assets (adjusting for potential tax savings). Under the law of one price, the return on the external assets equals the return on the firm's own assets because they have the same characteristics. Textbook logic therefore suggests that estimating the expected return on the firm's own assets allows us to determine the firm's cost of capital.

The main challenge is that expected returns on the firm's assets are not observed. In particular, equity holders provide two-thirds of the total capital of listed firms, so estimating the expected return that firms have to provide their equity holders—the so-called cost of equity—is key. Firm investments have long-term duration, so the challenge is estimating long-run expected equity returns. Existing estimates, like the CAPM, ICCs based on analyst forecasts, or interest rates, are typically biased, which means the estimates suffer from predictable errors and do not line up with average realized equity returns.

We introduce a machine learning method to forecast firm-level cash flows, which then allows us to derive the firm-level cost of capital. The method iteratively identifies which predictors are associated with higher cash flows at different future horizons. Each cash flow forecast relies only on information available in the given year, so the forecasts are made out-of-sample. In addition, we avoid overfitting by checking the results of each iteration against

an external holdout validation dataset. As soon as the forecast error in the validation data increases, we stop the prediction procedure, ensuring that we do not overfit to random noise in the training data. We build on the LightGBM algorithm (Ke et al. 2017) to develop an efficient method that can be practically implemented on the large underlying dataset.

The new cash flow forecasts are available for roughly 16,000 firms listed in 93 countries. We forecast individual horizons up to ten years ahead and thereafter apply a terminal growth rate assumption based on country-level forecasts. The forecasts begin in 1950 for the U.S. and 1985 globally. We verify that, on average, the cash flow forecasts are strong predictors of future realizations. The forecasts explain around 40% of future variation at two-year horizons and around 15% at ten years. In comparison, existing cash flow forecasts relying on financial analysts or vector autoregressions (VARs) explain around 10% at two years and close to 0% at ten years.

The present value identity implies that we can determine the firm-level cost of equity using the new cash flow forecasts and the current equity market value. We term the resulting estimate the "market-based cost of equity" (MCE) because it is consistent with the current market value. It can also be interpreted as an estimate of the long-run expected equity return of the firm.

Following a standard asset pricing method, we can evaluate whether the MCE is unbiased (i.e., correct on average) by regressing realized equity returns on the MCE. We find coefficients close to 1 on the MCE, consistent with unbiasedness. In fact, across all parts of the distribution of the MCE, the average MCE is close to the average realized return. In comparison, realized equity returns are positively associated with analyst-based ICCs and VARs, but negatively associated with the CAPM. We also predict the cost of equity directly using our machine learning method (rather than forecasting cash flows first), which produces slightly biased and excessively dispersed estimates.

We estimate the firm-level cost of capital by combining the MCE with established estimates of firm leverage, corporate taxes, and the cost of debt. We term the resulting estimate the "market-based cost of capital" (MCC). According to standard teaching and models, firms should base their investment decisions on an unbiased cost of capital, like the MCC, to maximize market value.

We construct a firm-level "cost of capital deviation" by subtracting the MCC from the perceived cost of capital used internally by firms. Data on firm perceptions are collected by Gormsen and Huber (2025b) who find that firms' long-run investment decisions are indeed associated with the observed perceived cost of capital.

A simple model generates two predictions about the relation between firm market value and the cost of capital deviation. First, there should be an inverted U relation, so that firms with a smaller deviation achieve higher market values for a given level of book equity. Intuitively, firms with a small deviation make investment decisions more in line with the effective cost of capital used by the equity market, which is rewarded by the market. In the data, we find exactly this relation. The coefficients suggest that a 1 percentage point decrease in the deviation lowers equity market value by 7 to 9 log points.

Second, the model predicts that the equity market reacts more positively to new investments when the cost of capital deviation is larger. A positive deviation implies that the perceived cost of capital is "too high" relative to the MCC, so the firm invests too little relative to the market's view. We measure when firms announced large new investments using a database of firm-level news. The sample contains unambiguous announcements of concrete projects undertaken by the firm itself. In the data, the equity market returns in 4-day windows around the announcements are significantly more positive for firms with a larger cost of capital deviation, consistent with the model.

Taken together, the empirical results suggest that the cost of capital used by firms is a meaningful determinant of market values. Textbooks widely teach this relation as a practical guideline and standard models use it as a standard assumption. However, there is little evidence for this relation, in part due to the absence of an unbiased estimate of the cost of capital. Using the MCC enables us to present such evidence on the importance of the cost of capital for firm value.

Our paper complements recent work by Hommel et al. (2025) who evaluate how well different estimates of the cost of capital predict firm equity market values. Consistent with our findings, they show that the CAPM performs poorly, whereas a Lasso-based method is relatively more accurate. Previous surveys reveal qualitative evidence about the eclectic methods used by firms to set their perceived cost of capital (e.g., Graham and Harvey 2001; Jacobs and Shivdasani 2012; Mukhlynina and Nyborg 2016; Jagannathan et al. 2016; Graham 2022). Gormsen and Huber (2025b) find that the average firm does not use an unbiased perceived cost of capital, which can lead to capital misallocation, but they do not provide an unbiased cost of capital and do not estimate the relation between firm market values and cost of capital deviations.

2 Conceptual Overview

In this section, we define the cost of capital that we are interested in measuring, before describing our method for doing so.

2.1 The Cost of Capital and Expected Returns

Financial investors provide capital to the firm in the form of debt and equity. The market value of the debt provided to the firm is D and that of equity is E. The firm invests its capital in assets with an expected long-run rate of return r_t^{firm}. In addition, the firm deducts its payments to debt holders from its taxes, so it also generates tax savings of $\text{D} \times \tau \times r_t^{\text{debt}}$, where τ is the tax rate and r_t^{debt} the debt rate. In total, the expected return on capital invested in the firm is:

$$\text{ROR}_t = \frac{(\text{D+E}) \times r_t^{\text{firm}} + \text{D} \times \tau \times r_t^{\text{debt}}}{\text{D+E}} \tag{1}$$

$$= r_t^{\text{firm}} + \frac{\text{D}}{\text{D+E}} \tau r_t^{\text{debt}}. \tag{2}$$

Instead of investing in the firm, debt holders could purchase an alternative, external debt security that has the same characteristics as firms' debt and generates an expected return $\widetilde{r_t^{\text{debt}}}$. Similarly, equity holders could invest in an external equity security that has the same characteristics as the firm's equity and generates an expected return of $\widetilde{r_t^{\text{equity}}}$. If all the firm's investors reallocated their capital to the external securities, their expected return would be:

$$\text{OCC}_t = \frac{\text{D} \times \widetilde{r_t^{\text{debt}}} + \text{E} \times \widetilde{r_t^{\text{equity}}}}{\text{D+E}}, \tag{3}$$

which is therefore the opportunity cost of the capital (OCC) invested in the firm.

Assuming the law of one price holds, investments with the same characteristics (e.g., risk and term) yield the same expected return. Hence, the expected return on the firm's debt equals the return on the external security, $r_t^{\text{debt}} = \widetilde{r_t^{\text{debt}}}$, which is also known as the firm's "cost of debt."[1] Similarly, the expected return on the firm's equity is $r_t^{\text{equity}} = \widetilde{r_t^{\text{equity}}}$, called the firm's "cost of equity."

[1] This formulation follows the simplifying assumption, suggested in many textbooks and by practitioners, that the tax saving is proportional to the expected return on debt, whereas in reality the tax saving may depend on realized interest payments. Relaxing the assumption introduces additional terms, but leaves the basic intuition intact.

The law of one price also implies that the return on the capital invested in the firm (ROR) is identical to the return on investing the capital in external securities (OCC), as both investments have the same characteristics. By equalizing (2) and (3), we can solve for the expected return on the firm's assets under the law of one price:

$$r_t^{\text{firm}} = \frac{\text{D}}{\text{D+E}}(1-\tau)\, r_t^{\text{debt}} + \frac{\text{E}}{\text{D+E}}\, r_t^{\text{equity}} = \text{WACC}. \tag{4}$$

This formulation shows that the firm can use the expected returns on debt and equity to determine the expected return on its assets in the eyes of financial investors. This expected return r_t^{firm} is known as the firm-level after-tax cost of capital or the weighted average cost of capital (WACC). Intuitively, it can also be thought of as the funding cost of the firm's typical investment. The measurement in our paper estimates the WACC at the firm level. In the remainder of the paper, we use the terms "cost of capital" and WACC interchangeably.

2.2 The Cost of Capital, Firm Value, and Real Investment

The cost of capital plays an important role in academia and corporate practice. First, the cost of capital determines firm value. The net present value v_t of a firm equals the sum of expected free cash flows discounted by the matching expected return on the firm's assets. Assuming that the same expected return applies to all future cash flows, the cost of capital is the appropriate discount rate:

$$v_t = \sum_{h=1}^{\infty} \frac{\mathbb{E}_t[x_{t+h}]}{(1+r_t)^h}, \tag{5}$$

where x_t is the expected free cash flow to debt and equity holders and r_t the cost of capital. This formulation assumes that firms are infinitely-lived, following the literature and practical valuations (Koller et al. 2020; Rosenbaum and Pearl 2021).

The cost of capital also influences firm investment decisions. Imagine a firm considering a long-run investment that has expected return r_t^{project} and is otherwise similar to the existing assets. If $r_t^{\text{project}} > WACC$, the expected return of the investment exceeds that of similar external securities. If the firm invests, it can expect to pay the debt rate to its debt holders and to have a greater return left for its equity holders than comparable securities. As a result, the equity market value of the firm increases. In contrast, if $r_t^{\text{project}} < WACC$ and the firm invests, the market value of the firm falls. This logic implies that firms should base

their investment decision on the cost of capital (WACC) if they want to maximize market value. Standard academic models assume that firms apply this rule, including neoclassical frameworks, models with frictions, and behavioral models with irrational agents.[2]

Executives are widely taught to use the cost of capital, based on expected returns as in (4), in their investment decisions. Surveys suggest that the vast majority of listed firms indeed know and regularly update a "perceived cost of capital," their internal estimate of the WACC (e.g., Trahan and Gitman 1995; Graham and Harvey 2001).

In practice, firms use a required return on investment, called discount rate or hurdle, as the minimum expected return that potential investments need to meet. Firms typically set discount rates using the perceived cost of capital as the basis and then adding additional terms, for example, to account for the idiosyncratic risk of potential projects, real option values, inflated return expectations provided by division managers, or financial constraints. Although discount rates often do not exactly equal the perceived cost of capital, existing work shows that the perceived cost of capital is a key input that determines the allocation of capital and returns across firms. For instance, the perceived cost of capital is strongly correlated with discount rates, real investment rates, and realized returns on invested capital across firms (Gormsen and Huber 2025b). Moreover, in the long run, changes in the perceived cost of capital strongly affect discount rates (although the short-run incorporation is weak, Gormsen and Huber 2025a).

2.3 The Measurement Challenge

Taken together, the cost of capital is a key input into academic models and the real decision-making of firms. However, expected returns on securities are fundamentally unobserved. They depend on the risk perceptions and preferences of financial investors. As a result, firms do not know their cost of capital. The CAPM is often used in practice to approximate the cost of capital, but it is well known that the CAPM is biased with respect to expected returns.

In general, there exists no method that estimates an unbiased cost of capital and no readily available firm-level dataset. This absence is especially glaring given that, according to standard teaching, firms need to know their cost of capital to maximize their market value. In the remainder of this paper, we develop an estimation method of the cost of

[2]If the firm's objective is not to maximize current market value, it may prefer an alternative return threshold, for instance, a future expected return (Stein 1996) or a fundamental risk benchmark (Gormsen and Huber 2025b).

capital, present firm-level estimates, and finally evaluate its economic relation to market value.

3 Method to Estimate a Market-Based Cost of Capital

3.1 Overview of the Method

Our aim is to estimate the firm-level cost of capital, given by (4). There exist established estimates for firm leverage, tax rates, and the cost of debt, which we adopt in our implementation in Section 4. Our method focuses on the remaining term: the cost of equity. On average, listed firms are financed with two-thirds equity, so the cost of equity is a key component of the firm-level cost of capital.

Analogous to the firm's overall value in (5), the firm's equity market value p_t is:

$$p_t = \sum_{h=1}^{\infty} \frac{\mathbb{E}_t[\text{fcfe}_{t+h}]}{(1 + r_t^{\text{equity}})^h}, \tag{6}$$

where $\mathbb{E}_t[\text{fcfe}_{t+h}]$ is a forecast for the free cash flow accruing to equity holders in period $t+h$. Firm market value p_t is directly observed, but we do not easily observe cash flow forecasts that capture substantial variation. The key challenge in estimating the cost of equity r_t^{equity} is therefore determining cash flow forecasts.

Our method focuses on the estimation of cash flow forecasts with strong predictive power, which can then be used to derive the cost of equity.

3.2 Forecasting Cash Flows and Net Present Value

Since we cannot estimate infinite-horizon cash flows, we define an explicit forecasting horizon of H periods and a terminal value period. We impose the terminal value condition that all firms' cash flows grow at the same constant rate g_t:

$$p_t = \sum_{h=1}^{10} \frac{\mathbb{E}_t[\text{fcfe}_{t+h}]}{(1 + r_t^{\text{equity}})^h} + \frac{\mathbb{E}_t[\text{fcfe}_{t+10}](1 + g_t)}{r_t^{\text{equity}} - g_t} \frac{1}{(1 + r_t^{\text{equity}})^{10}}. \tag{7}$$

In our estimation, we set $H = 10$. The terminal value condition can be motivated as an economic prior that competition eliminates above-average growth in the long run. Consistent

with this view, the terminal growth rates used in valuation practice do not vary much across firms (Decaire and Guenzel 2023).

We need to forecast cash flows in every period t for all horizons $h \leq 10$. We employ a machine-learning technique based on decision tree learning. This technique essentially splits firms into groups based on potential predictors of cash flows, and then estimates which groups are associated with higher future cash flows. However, traditional tree learning is time-intensive, and would be infeasible for a large number of firms, periods, and horizons. We therefore build on the LightGBM model of Ke et al. (2017), which allows running many simultaneous decision trees efficiently and fast.[3] All forecasts are out-of-sample, meaning that the cash flow forecasts are based on information that was available at time t. We elaborate on our implementation details in Section 4.

3.3 Cash Flows Forecasts and the MCC

With the cash flow forecasts in hand, we can solve (7) for the "market-based cost of equity" (MCE), our estimate of r_t^{equity}. This MCE equates the discounted sum of expected cash flows with the observed market value. It is therefore the expected return on the firm's equity in financial markets.

We insert the MCE into the definition of the cost of capital (4). Combined with conventional measures of firm leverage, tax rates, and the cost of debt, we calculate our estimate of the firm-level cost of capital. We term it the "market-based cost of capital" (MCC) since it is consistent with observed market values. Estimating the MCC requires cash flow forecasts with strong predictive power. If we instead used weak forecasts, the resulting cost of capital would be severely biased.

Following standard asset pricing techniques, a tempting alternative approach might be to directly predict the cost of equity, for example, using a factor model. However, this approach is tailored to estimating short-run expected returns. The cost of equity depends on long-run returns that match the horizon of firms' investments, usually at least a decade long. Existing return data do not cover many decades, so this approach is infeasible.

Our method is inspired by seminal work calculating an "implied cost of capital" (e.g., Gebhardt et al. 2001; Claus and Thomas 2001; Easton 2004; Ohlson and Juettner-Nauroth 2005), which relies on a net present value formulation and analyst forecasts to derive the cost

[3]The model is a gradient-boosted decision tree ensemble (GBDT). Such GDBTs perform well on tabular data, outperforming deep learning methods across a wide variety of tabular benchmarks (Shwartz-Ziv and Armon 2022).

of equity. Since analyst forecasts are biased, the implied cost of capital is also biased. This has made it challenging to use it to assess the importance of the cost of capital for market value.

4 Estimating a Market-Based Cost of Capital

4.1 Data

We construct a database for around 16,000 firms listed in 93 countries. We start the samples in 1950 for U.S. stocks and 1985 for non-U.S. stocks to ensure that accounting data are available. Data on market values and stock returns originally come from CRSP for U.S.-listed firms and from Compustat for non-U.S.-listed firms. The accounting data are from Compustat. We rely on the database provided by Jensen et al. (2023) through WRDS and GitHub. Finally, we add analyst forecasts from I/B/E/S, and we always use the median forecasts from the consensus file.

We exclude financial firms (i.e., groups 45–48 in the Fama and French 1997 classification) because free cash flows for financial firms are conceptually different from free cash flows from non-financial firms.[4] We also require firms to have a non-missing and positive accounting value of assets and shares outstanding.

4.2 Estimating the cost of equity (MCE)

Our method requires cash flow forecasts to determine the cost of equity. We forecast real cash flows in U.S. dollars. The forecasts are made in every month t for up to ten years in the future. The terminal growth rate for cash flows is the country-specific forward real GDP growth rate between years 4 and 5 from the IMF World Economic Outlook.[5]

We employ two measures of cash flows to equity holders: free cash flows to equity and residual income. Details on the data construction are in Appendix A.1.

When using the free cash flow to equity, we directly rely on (6). The free cash flow to equity is:

$$\text{fcfe}_t = e_t + (\text{d\&a}_t - \text{capx}_t) - \Delta\text{nwc}_t + \Delta\text{debt}_t, \tag{8}$$

[4]We aim to estimate firms' unlevered free cash flows by separating cash flows due to operations from cash flows due to financing. However, this separation is difficult for financial firms because their business model is all about financing.

[5]We use a forward rate rather than a cumulative rate because the terminal growth rate should reflect steady-state growth, uncontaminated by near-term cyclical effects.

where e_t is net income, $\text{d\&a}_t$ is depreciation and amortization, capx_t is capital expenditures, Δnwc_t is the change in net operating working capital, and Δdebt_t is net debt issuance. We forecast free cash flows to equity in $t+1,\ldots,t+10$, and solve for r_t^{equity} numerically.

When using the residual income measure (Edwards and Bell 1961; Ohlson 1995), we rely on a slightly modified net present value formulation:

$$p_t = b_t + \sum_{h=1}^{10} \frac{\mathbb{E}_t[\text{ri}_{t+h}]}{(1+r_t^{\text{equity}})^h} + \frac{\mathbb{E}_t[\text{ri}_{t+10}](1+g_t)}{r_t^{\text{equity}} - g_t} \frac{1}{(1+r_t^{\text{equity}})^{10}}, \tag{9}$$

where b_t is the firm's book equity and ri_t is residual income:

$$\text{ri}_t = e_t - r_t^{\text{equity}} b_{t-1}. \tag{10}$$

To estimate the cost of equity from this model, we forecast net income in $t+1,\ldots,t+10$, and infer the future book equity as:

$$b_t = b_{t-1} + e_t - d_t, \tag{11}$$

where d_t is the firm's dividends. We estimate dividends by multiplying our earnings forecast with the firm's most recent payout ratio (d_t/e_t).[6] The method for inferring the future book equity is inspired by Gebhardt et al. (2001). We then solve for r_t^{equity} numerically.

Our MCE measure is the average of the two cost of equity estimates (if only one is available, we use that estimate directly). Because the cash flow forecasts are in real terms, the implied discount rate is a real cost of equity. We convert it to nominal by adding the expected five-year forward inflation rate from the IMF World Economic Outlook. Further, we remove estimates where the implied real cost of equity exceeds 10%, because such high values are implausible, and indicate that our model assumptions are invalid for the specific firm.[7] Further details are available in Appendix B.[8]

The explanatory variables in the cash flow prediction models are 153 characteristics from Jensen et al. (2023); 79 characteristics we derive from their data set related to current cash

[6] Our method for inferring the future book equity can result in negative book equity, in which case the residual income from (10) is not well defined. When projected book equity becomes negative, we set it to zero, which implies zero residual income from that point forward.

[7] For example, our valuation models assume an infinite firm lifetime, which is a poor assumption for firms that the market expects to go out of business soon.

[8] The appendix contains additional information about the MCE. Specifically, Figure A10 shows the coverage, and Figures A11 to A13 show how the real, nominal, and excess cost of equity vary over time.

flow levels and changes; 12 characteristics based on analyst forecasts from IBES related to the expected cash flows for the firm over the next three years and the analyst long-term growth rate forecast, as well as the corresponding median within the firm's industry and country, respectively; 45 industry dummies based on the classification from Fama and French (1997), expanded to also include the computer software group, but without firms from the four financial groups; and 93 country dummies. In total, the models are fit on 240 continuous and 138 dummy characteristics. Further details about the input features are in Appendix A.2.[9]

All predictions are out-of-sample. The model is trained on an expanding training sample, with the last 10 years used as the validation sample. We select hyperparameters by evaluating 20 candidate sets, chosen to maximize entropy over a predefined search space spanning five specific hyperparameters, and select the ones with the lowest mean squared error in the validation sample. Further details on the model and hyperparameters are in Appendix A.3.

When the hyperparameters have been found, we re-estimate the model on the combined training and validation data. The first model is trained on a data set ending on December 31$^{\text{st}}$, 1983, and we re-estimate the model parameters and hyperparameters every year. All forecasts are made out-of-sample (OOS) on a monthly basis, with the first forecast made on January 31$^{\text{st}}$, 1984, and the last on December 31$^{\text{st}}$, 2023. Further details on the sample split are in Appendix A.4.

4.3 Estimating the cost of debt

We observe traded bond yields for U.S. firms, so we can calculate the cost of debt as:

$$\mathbb{E}_t[r^{\text{debt}}_{t+1}] = \text{yield}_t - [\text{prob. of default}]_t \times (1 - [\text{recovery rate}]_t). \tag{12}$$

We compute the yield as the weighted average over all the firm's outstanding bonds using bond market value as weights. Data on bond yields are from Dick-Nielsen et al. (2023), based on dealer quotes over 1973-2001 and traded prices over 2002-2023. The probability of default is the average realized default rate of bonds with the same rating over the past 3 years, following Campello et al. (2008). The recovery rate is again rating specific and from

[9]Some of the explanatory variables—especially those that require analyst data—are missing for a large fraction of our sample. Removing observations with missing explanatory variables would thus drastically reduce our sample and tilt it towards relatively large stocks with analyst coverage. Fortunately, LightGBM handles missing values natively by learning how missing values map to the dependent variable. As a result, we keep observations with missing values and delegate their handling to the learning algorithm.

Altman and Kishore (1998). To adjust for differences in bond maturity, we subtract the duration-matched risk-free rate and add the ten-year U.S. treasury yield:

$$r^{\text{debt}} = \mathbb{E}_t[r_{t+1}] - r^f_{\text{duration-matched}} + r^f_{10}. \tag{13}$$

This serves as our cost of debt estimate for firms with traded bonds.

We project the cost of debt for firms with traded bonds on their distance-to-default DtD_t, estimated as in Bharath and Shumway (2008):

$$r^{\text{debt}}_t = a_t + b_t \times \text{DtD}_t + \varepsilon. \tag{14}$$

We then use the estimates from this regression to predict the cost of debt for firms without traded bonds. Figure A14 shows how the cost of debt vary over time.

4.4 Estimating the firm-level cost of capital (MCC)

We calculate the MCC, our market-based estimate of the firm-level cost of capital, using the WACC formula (4). We directly observe the market value of equity and proxy for the market value of debt using the book value of debt.[10] We estimate the tax rate as an equal-weighted average of the firm's effective tax rate and the median effective tax rate of all firms in the same year and country. The firm's effective tax rate is Compustat item `txt` divided by `pi`, capped at zero and one, to control for outliers (e.g., due to carry-forward tax losses). Figure A15 shows how the firm-level cost of capital vary over time.

5 Properties of Cash Flow Forecasts and the MCC

We show that the new cash flow forecasts have relatively strong predictive power for realized cash flows. Accordingly, the predicted cost of equity is an unbiased estimator of the firm-level cost of capital.

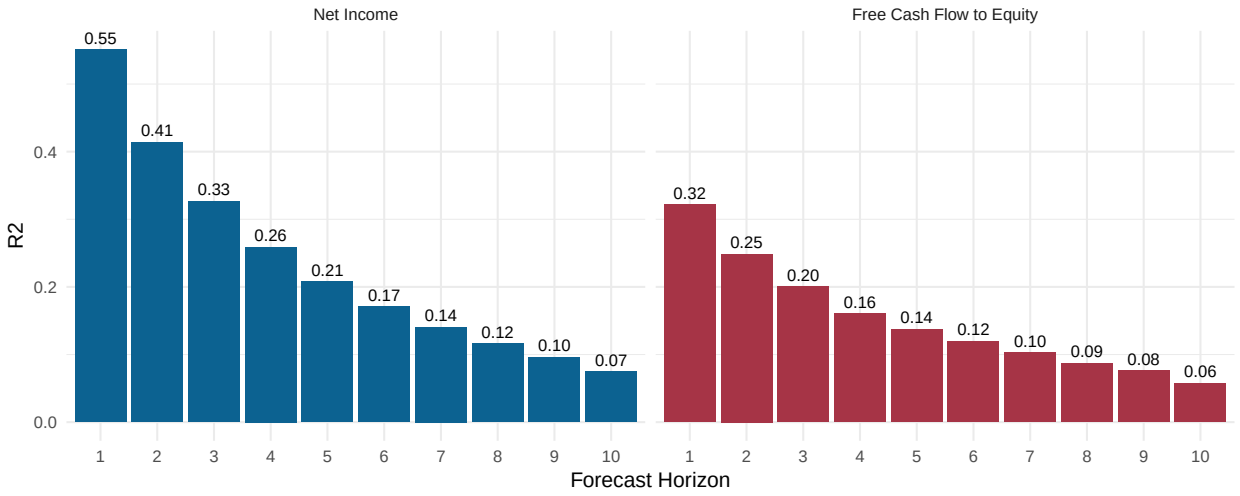
5.1 Cash flow predictability

Figure 1 shows the out-of-sample R^2 of the cash flow prediction models for each of the two cash flow targets and ten forecast horizons. The R^2 is positive across all the outcome

[10] In some textbooks, the leverage measure is the firm's long-run target leverage (Ross et al. 2022). Under that view, one can view our market-based current leverage measure as a proxy for the target leverage.

Figure 1
MCE Cash Flows

This figure shows the out-of-sample R^2 of the cash flow forecasts used by the MCE model. We consider two different cash flow variables, as indicated by the panel header, and ten forecast horizons, as indicated by the x-axis. The results are based on global stocks. The sample is balanced, meaning that stocks must have non-missing realized and predicted cash flow for both variables and all ten horizons. Figure A2 shows the results from the unbalanced sample.



variables used in our prediction model. A positive R^2 is not guaranteed when conducting out-of-sample predictions. We predict short-horizon cash flows more accurately than long-horizon cash flows. Furthermore, net income is more predictable than free cash flows to equity because accounting rules smooth underlying cash flows. For example, a large investment strongly reduces free cash flows to equity in one year, but has weaker effects on net income over several years due to depreciation allowances. Figure A3 shows that the R^2 is relatively stable over time, with a slight upward trend.

We compare our cash flow forecasts to existing benchmarks from the literature. Few papers predict long-run cash flows, so we mainly consider earnings forecasts.[11] The first two benchmark methods we consider come from Claus and Thomas (2001, CT) and Gebhardt et al. (2001, GLS), which are part of the analyst-based ICC literature. Both methods take analyst forecasts over the next two years as input, but differ in how they extrapolate cash flows beyond year two. CT assumes that cash flows grow at analysts' long-term growth

[11] Analysts typically produce earnings forecasts one quarter or one year ahead. For papers trying to forecast short-run cash flows with machine learning, see Hillenbrand and McCarthy (2023), Van Binsbergen et al. (2023), Zhang et al. (2024), and De Silva and Thesmar (2024). Among these, the longest horizon considered is four years ahead, by De Silva and Thesmar (2024).

rate forecasts, while GLS assumes that each firm's return on equity converges linearly to its corresponding industry average over a 12-year period. The third model is a vector autoregression (VAR) method similar in spirit to Vuolteenaho (2002), Campbell et al. (2010), and Cho et al. (2024). Finally, we consider the random walk model, which assumes that all future cash flows are equal to the most recent realization. The random walk (RW) model of earnings is supported by early research in accounting (Ball and Watts 1972), and earnings surprises are often defined with respect to the RW model (see, e.g., Foster et al., 1984 and Martineau, 2022). For a detailed description of the benchmark models, see Appendix C.

Figure 2 shows that the out-of-sample R^2 of the MCC cash flow forecasts exceeds that of all the benchmarks. At a one-year horizon, the difference relative to CT and GLS (which at this horizon are effectively analyst forecasts) is relatively small.[12] However, by year three, both have a negative out-of-sample R^2. The VAR model performs slightly worse than the ICC methods at the one- and two-year horizons, but maintains a positive R^2 at longer horizons. The RW model performs relatively poorly at short horizons but is the best benchmark at the ten-year horizon. Overall, the results suggest that the MCC cash flow forecasts outperform existing methods in the literature—especially at the long horizons that are key for the cost of capital.[13]

5.2 Return predictability

The predictive power of the MCE The market-based cost of equity (MCE) is meant to predict returns over a very long-run horizon. We estimate the association between the MCE and future ten-year realized equity returns:

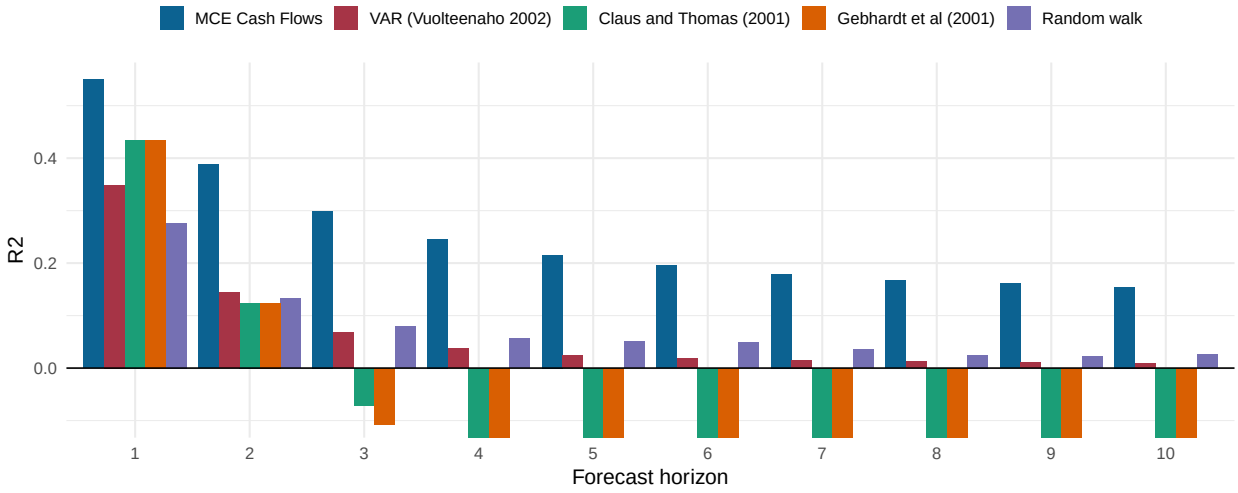
$$r^{\text{realized},i}_{t\to t+120} = \alpha_{c,t} + b \times \left[\left(1+\text{MCE}^i_t\right)^{10} - 1\right] + \varepsilon^i_{t+h}, \tag{15}$$

[12]The dependent variable in Figure 2 is for earnings realized over a constant yearly horizon, but the analyst forecasts used by CT and GLS are for the next fiscal year, which is typically closer than one year. Figure A7 shows that the ML model also beats analysts when we map their forecasts to a constant one-year horizon. It also shows that the analysts' out-of-sample R^2 improves when the forecast is demeaned to remove bias arising from well-known analyst conflicts of interest (Richardson et al. 2004). However, even after demeaning the forecasts and outcome, the ML model remains superior.

[13]Figure A1 show a binned scatter plots of predictions relative to realizations for each of the models. The CT, GLS, and RW models have systematic biases for various parts of the distributions, whereas both our method and VAR appear approximately unbiased. Our method outperforms VAR because it remains unbiased while making more dispersed predictions. For example, at the ten-year horizon, our method predicts earnings to assets of -0.2 for the bottom percentile of firms and 0.3 for the top percentile. By contrast, the comparable spread in VAR predictions is only 0.01 to 0.12.

Figure 2
Cash Flow Forecast Benchmarking

The figure shows the out-of-sample R^2 for net income forecasts at horizons 1 through 10 years for different methods. The MCE cash flows come from a LightGBM machine learning model trained on firm characteristics, explained in detail in Appendix A; the benchmarks are implied cash flow forecasts from the Claus and Thomas (2001) and Gebhardt et al. (2001) analyst-based ICC models; a VAR model similar to the one in Vuolteenaho (2002); as well as a random walk (last observed net income). All forecasts are scaled by total assets and winsorized at the 1st and 99th percentiles. The benchmark models are explain in detail in Appendix C. The y-axis starts at -0.1, even though some methods have lower R^2. The sample is balanced, requiring all methods to have non-missing predictions for each firm-month-horizon observation.



where $r^i_{t \to t+120}$ is the buy-and-hold realized return from t to $t+120$ months (ten years), $\alpha_{c,t}$ is a country-by-time fixed effect, and MCE^i_t is the market-based cost of equity converted to a ten-year horizon. Standard errors Driscoll-Kraay with a 120 month bandwidth, to account for the overlapping observations in the dependent variable.

One challenge with analyzing ten-year returns is that some firms delist, potentially introducing survivorship bias in the sample. Bessembinder (2018), for example, shows that the median time a stock is listed in the CRSP database is 7.5 years and that firms often delist for performance-related reasons. To address this issue, we include all delisted firms in the sample, even if they delist during a ten-year period. We impute returns after delisting using the median return for firms in the same country and month, and continue compounding this imputed return. The imputation approach has no impact on a firm that goes bankrupt and has a delisting return of -100%, but ensures that firms that, say, get acquired continue to accumulate returns.[14]

[14]In unreported results, we find that our results are weaker if we remove stocks that delist from the sample,

Table 1, Panel A, reports regressions of ten-year realized returns on the MCE, controlling for month fixed effects. The point estimates in the U.S. and the global sample are close to 1 and statistically indistinguishable from 1. The estimates are also significantly different from 0. The MCE is an unbiased estimate of expected returns. Panel B reports the average error, that is, the ten-year realized minus predicted return. In the U.S. sample, the average error is a significant 25% (approximately 2% annualized). This result suggests that realized returns in the out-of-sample period exceeded investors' ex ante expectations in the U.S., consistent with Binsbergen et al. (2025). In the global and non-U.S. samples, however, the average errors are closer to zero and statistically insignificant, indicating no systematic errors overall.

	10-year realized return		
	U.S.	Global	Global ex-U.S.
	(i)	(ii)	(iii)
Panel A: Regression estimates			
MCE	0.93***	0.96***	1.01***
	(0.08)	(0.18)	(0.23)
Observations	1,106,464	2,610,489	1,504,025
Panel B: Predictive accuracy ($\alpha = 0, \beta = 1$)			
Avg. error	0.25	0.14	0.06
	(0.12)	(0.10)	(0.15)
OOS R^2_{CAPM} (%)	2.1	3.8	5.3
OOS R^2_{MKT} (%)	0.4	2.1	4.1

The table reports regressions of 10-year realized buy-and-hold returns on the current 10-year cost of equity, estimated separately for U.S., global, and global ex-U.S. samples. All regressions include month fixed effects. The average pricing error is the mean difference between realized returns and the predicted cost of equity. OOS R^2_{CAPM} and R^2_{MKT} report out-of-sample predictive R^2, computed as $1 - \sum(r_i - \hat{r}_i)^2 / \sum(r_i - \hat{r}_{i,\text{bench}})^2$, where the benchmark prediction is either the CAPM-implied cost of equity or the country-level expanding-window average market return. Both benchmarks use only information available at the time of prediction. We report Driscoll-Kraay standard errors using a bandwidth of 120 month in parentheses. Statistical significance is denoted by *** p<0.01, ** p<0.05, * p<0.1.

Table 1: Return predictability

because firms that go bankrupt are relatively more likely to have a low MCC. Further, the results are similar whether we impute using the cross-sectional median or mean.

Table 1, Panel B, also shows the out-of-sample R^2 relative to two benchmark models, defined as:

$$R^2_{\text{bench}} = 1 - \frac{\sum(r_i - \hat{r}_i)^2}{\sum(r_i - \hat{r}_{i,\text{bench}})^2}, \tag{16}$$

where r_i is the realized return, $\hat{r}_i$ is the out-of-sample MCE prediction, and $\hat{r}_{i,\text{bench}}$ is the out-of-sample benchmark prediction. The more standard in-sample R^2 uses the in-sample mean instead of the benchmark prediction, but the in-sample mean is not available in real-time and is therefore an inappropriate benchmark for our out-of-sample predictions. Instead, we consider two alternative cost of equity estimates that could have been used in real time. MKT is the ten-year risk-free rate at the estimation date plus a rolling estimate of the historical ten-year local market excess return based on data through this date. CAPM is the 10-year risk-free rate plus the same rolling estimate of the local market risk premium, multiplied by the stock's market beta, using the last 60 months of data. The out-of-sample R^2s are all positive, suggesting that an investor would achieve higher accuracy using the MCE than either the historical market return or the CAPM.

The results so far show that MCE predicts 10-year buy-and-hold returns. Figure 3 shows that the MCE also predicts forward returns, that is, the returns in a specific month $t + h$, for horizons $h = 1, 2, \ldots, 120$ months. This result suggests that long-run discount rates vary across firms, in contrast to the conclusions from Keloharju et al. (2021).

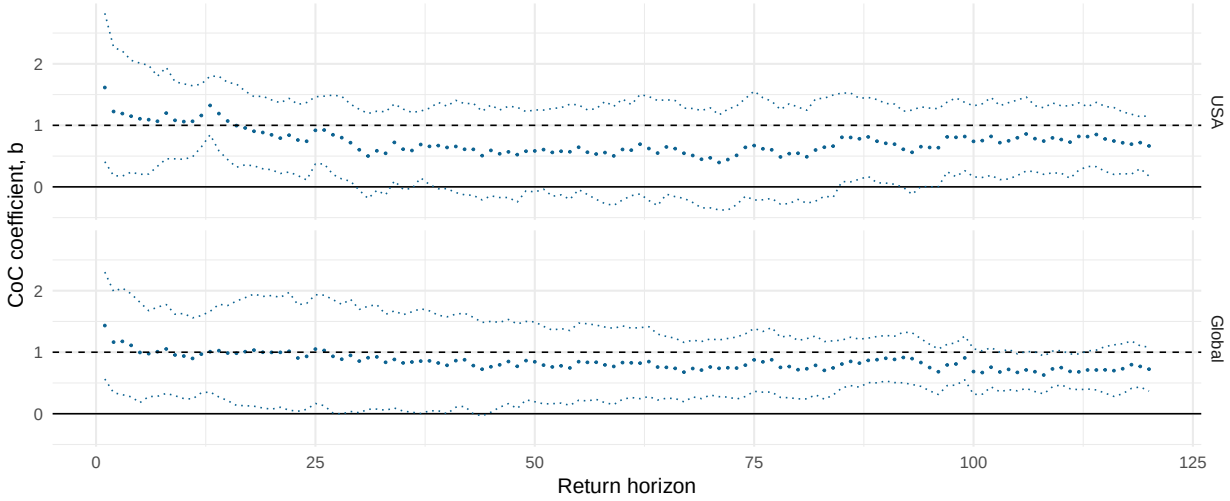
MCE versus benchmarks. We plot a binned scatter plot showing the association between the MCE and future realized ten-year equity returns in the top-left panel of Figure 4. The relation is strong, as all points are close to the 45-degree line.

We also plot the relation between alternative benchmarks for the cost of equity and realized returns (for a detailed description of the benchmark models, see Appendix C). We first consider a method that we call ML-direct (MLR), which uses our prediction model to directly estimate ten-year predicted equity returns, rather than first predicting cash flows and then deriving the MCE. The MLR predicts returns positively, but with a lower slope than the MLE. The main issue is that MLR predicts too extreme a dispersion in returns, which is not subsequently realized. For example, the spread in predicted returns between the top and bottom decile portfolios is 9%, whereas the corresponding spread in realized returns is only 5% (for MCE, the corresponding spreads are 4% for both predicted and realized returns).[15]

[15]The underlying issue is that MLR applies insufficient shrinkage to its out-of-sample forecasts. We examine this in Figure A8, which tracks the optimal ridge penalty λ across training vintages. The test-optimal λ is systematically higher than the validation-optimal λ, indicating that the degree of shrinkage

Figure 3
Forward return predictability

This figure reports the slope coefficient b from regressions of future excess stock returns on the MCE, $r_{t+h}^{\text{realized},i} = \alpha_t + b \times \left[(1 + \text{MCE}_t^i)^{1/12} - 1\right] + \varepsilon_{t+h}^i$, for return horizons $h = 1, 2, \ldots, 120$ months. All regressions include month fixed effects. The top row restricts the sample to U.S. firms; the bottom row uses the global sample. Dotted lines indicate 95% confidence intervals based on Driscoll-Kraay standard errors with a 120-month bandwidth. A coefficient of one (dashed line) implies that a one-percentage-point higher MCE maps one-for-one into higher realized returns.



The other benchmarks we consider have a weaker relationship with future returns than both MCE and MLR. Relatively speaking, the GLS-based implied cost of capital and the VAR model are the next best models, and the CAPM is the worst, having a negative relationship with future returns.

Table 2 reports additional statistics for the regression in (15). The point estimate for the MLE is the only one close to 1. The only other benchmarks with a significantly positive point estimate are MLR, VAR, and GLS. Panel C in that table shows that the relative accuracy of MCE is higher than all other benchmarks considered.

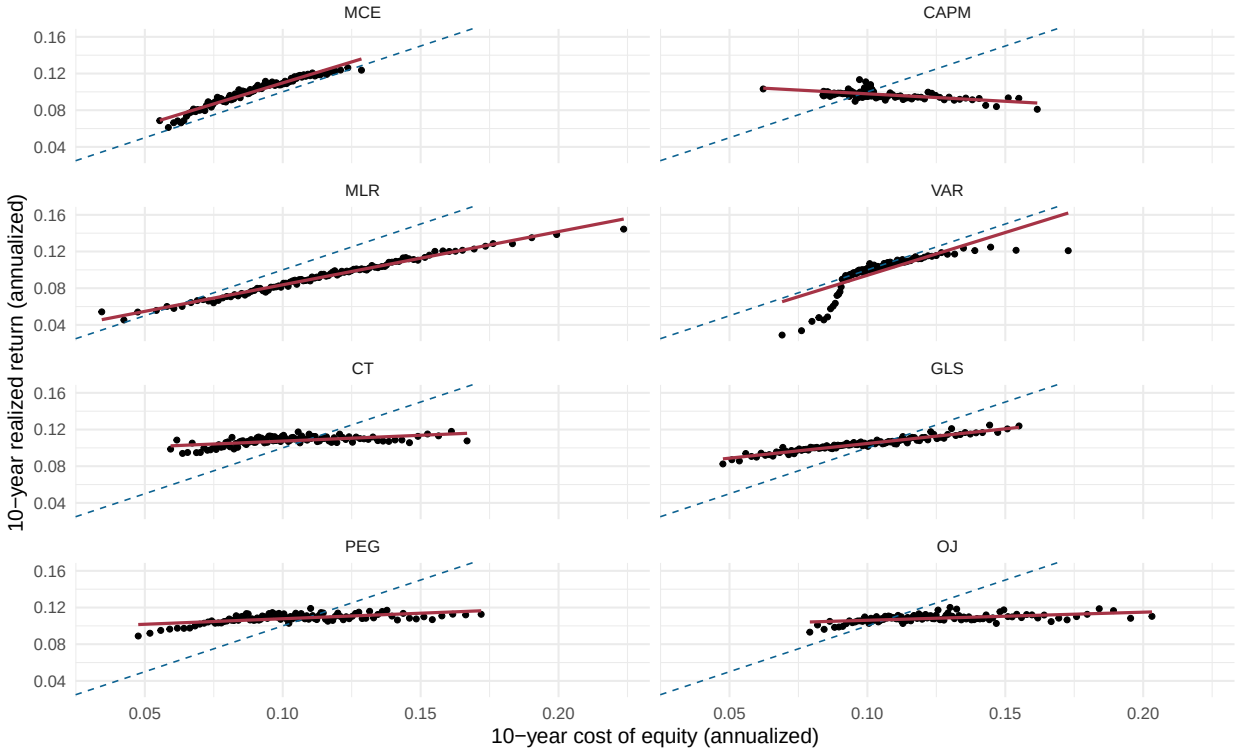
6 The Cost of Capital and Firm Value

Using a simple model, we formalize the classical intuition that firms achieve a higher market value when they use an unbiased cost of capital, like the MCC, in their investment decisions.

chosen ex-ante is consistently too low relative to what would have been optimal ex-post. Similarly, Figure A9 shows that return forecasts from our MCE approach are more stable across time than the corresponding forecasts from MLR.

Figure 4
Return Forecast Benchmarking

This figure shows the relationship between future 10-year realized returns and the current 10-year cost of equity for different estimation methods. MCE is the market-based cost of equity from this paper. CAPM computes the cost of equity as the risk-free rate plus beta times an expanding-window equity risk premium. MLR predicts 10-year cumulative excess returns directly using LightGBM, and then adds the current risk-free rate. VAR uses a firm-level vector autoregression similar to the one used in Vuolteenaho (2002). CT, GLS, PEG, and OJ, are analyst-based ICC estimates following Claus and Thomas (2001), Gebhardt et al. (2001), Easton (2004), Ohlson and Juettner-Nauroth (2005), respectively. Each month, stocks are sorted into 100 portfolios by their estimated cost of equity. For each portfolio, we compute the average cost of equity and the average realized buy-and-hold return over the next 10 years, and then average across months. Both axes are annualized. The solid red line shows the best linear fit, and the dashed blue line shows the 45-degree line. The sample includes all global stocks, 1983–2023.



	10-year realized return							
	MCE	CAPM	MLR	VAR	CT	GLS	PEG	OJ
	(i)	(ii)	(iii)	(iv)	(v)	(vi)	(vii)	(viii)
Panel A: Regression estimates								
Cost of equity	0.96***	-0.13	0.50***	0.66***	0.11	0.38**	0.10	0.07
	(0.18)	(0.13)	(0.16)	(0.21)	(0.13)	(0.17)	(0.10)	(0.10)
Observations	2,610,489	3,724,501	4,872,824	3,463,929	1,003,308	1,877,706	1,000,734	985,775
Panel B: Predictive accuracy ($\alpha = 0, \beta = 1$)								
Avg. error	0.14	0.21	-0.19	-0.04	0.05	-0.00	-0.05	-0.56
	(0.10)	(0.14)	(0.28)	(0.09)	(0.08)	(0.08)	(0.10)	(0.08)
OOS R^2_{CAPM} (%)	3.6	0.0	0.9	2.8	1.4	2.9	1.3	-1.3
OOS R^2_{MKT} (%)	2.1	-1.1	-0.5	1.6	-0.1	1.1	-0.5	-2.6
Panel C: MSE relative to MCE								
ΔMSE	—	0.74	0.94	0.14	0.25	0.12	0.34	0.88
$p^{\text{bs}}_{\text{ind}}$		0.000	0.000	0.000	0.000	0.004	0.000	0.000
$p^{\text{bs}}_{\text{time}}$		0.001	0.004	0.083	0.093	0.109	0.048	0.010

The table reports regressions of 10-year realized buy-and-hold returns on the current 10-year cost of equity for eight estimation methods. MCE is the market-based cost of equity from this paper. CAPM computes the cost of equity as the risk-free rate plus beta times an expanding-window country-specific equity risk premium. MLR predicts 10-year cumulative excess returns directly using LightGBM, and then adds the current risk-free rate. VAR uses a firm-level vector autoregression similar to the one used in Vuolteenaho (2002). CT, GLS, PEG, and OJ, are analyst-based ICC estimates following Claus and Thomas (2001), Gebhardt et al. (2001), Easton (2004), Ohlson and Juettner-Nauroth (2005), respectively. All regressions include month fixed effects. The average pricing error is the mean difference between realized returns and the predicted cost of equity. OOS R^2_{CAPM} and R^2_{MKT} report out-of-sample predictive R^2, computed as $1 - \sum(r_i - \hat{r}_i)^2 / \sum(r_i - \hat{r}_{i,\text{bench}})^2$, where the benchmark prediction is either the CAPM-implied cost of equity or the country-level expanding-window average market return. Both benchmarks use only information available at the time of prediction. Panel C reports $\Delta\text{MSE} = \text{MSE}_{\text{method}} - \text{MSE}_{\text{MCE}}$, computed on the id-month pairs where both MCC and the comparison method have non-missing predictions. Positive values indicate that MCE has lower MSE. $p^{\text{bs}}_{\text{ind}}$ and $p^{\text{bs}}_{\text{time}}$ are one-sided bootstrap p-values for the null that $\Delta\text{MSE} \leq 0$, obtained by resampling Fama-French 49 industries and 10 equal-sized time blocks, respectively, with replacement (10,000 iterations). We report Driscoll-Kraay standard errors using a 120 month window in parentheses. Statistical significance is denoted by *** p<0.01, ** p<0.05, * p<0.1.

Table 2: Return benchmarks – unbalanced – fixed effects

We find an inverted U relationship in the data: firms have greater market value when their perceived cost of capital is closer to the MCC. In addition, firm market value responds more positively to the announcement of a new investment when the perceived cost of capital is relatively larger than the MCC. The findings suggest that equity markets value firms more highly when their investment decisions are more in line with the MCC.

6.1 Model of inverted U

We sketch a simple model to formalize how firms' perceptions about their cost of capital affect firms' capital allocation and equity market value.

We consider an infinite-horizon discrete-time model with periods $t = 0, 1, 2, \ldots$ and a cross-section of firms indexed by i. Each firm i owns capital stock $K_{i,t}$ which depreciates at rate $\gamma \in (0, 1)$ per period. The law of motion for capital is:

$$K_{i,t+1} = (1 - \gamma)K_{i,t} + I_{i,t}, \tag{17}$$

where $I_{i,t}$ is gross investment at date t.

In each period, capital generates stochastic output according to the production function:

$$Y_{i,t}(K_{i,t}) = (\alpha_i + \eta_{i,t})K_{i,t} - \frac{\beta_i}{2}K_{i,t}^2, \tag{18}$$

where $\alpha_i, \beta_i > 0$ are firm-specific productivity parameters and $\eta_{i,t}$ is an i.i.d. productivity shock with $E[\eta_{i,t}] = 0$.

Firm problem. Firm i maximizes the perceived present value of future cash flows using the perceived cost of capital $r_{i,t}^{\text{perc.}}$. The perceived cost of capital may differ from the firm-level cost of capital, r_i^{firm}, that financial investors use to discount firm i's cash flows. We model the deviation as:

$$r_{i,t}^{\text{perc.}} = r_i^{\text{firm}} + \delta_{i,t}. \tag{19}$$

For simplicity, we assume that r_i^{firm} is constant and that the cost of capital deviation, $\delta_{i,t}$, is orthogonal to r_i^{firm} and the fundamentals α_i and β_i.

Firm i chooses capital to maximize its perceived present value of cash flows:

$$\max_{\{K_{i,t}\}_{t=0}^{\infty}} \mathbb{E}\left[\sum_{t=0}^{\infty} \frac{1}{(1 + r_i^{\text{perc.}})^t} \left[Y_{i,t}(K_{i,t}) - I_{i,t}\right]\right]. \tag{20}$$

Taking expectations over the productivity shocks and noting that firms choose a constant capital stock in the absence of adjustment costs, the firm's first-order condition implies that the optimal choice of capital, K_i^*, is given by:

$$K_i^* = \bar{K}_i - \frac{\delta_i}{\beta_i}, \tag{21}$$

where $\bar{K}_i \equiv (\alpha_i - r_i^{\text{firm}} - \gamma)/\beta_i$ is the capital the firm would choose if it used the correct discount rate ($r_i^{\text{perc.}} = r_i^{\text{firm}}$). In equilibrium, investment equals depreciation: $I_i^* = \gamma K_i^*$.

Market value. Investors price assets using a stochastic discount factor M_t. For firm i, the equity price satisfies:

$$P_i = \mathbb{E}\left[\sum_{t=1}^{\infty} M_t \cdot (Y_{i,t}(K_i^*) - \gamma K_i^*)\right]. \tag{22}$$

We assume that each firm is too small to influence the pricing of risk in the economy, such that the stochastic discount factor is not influenced by the choice of capital of the individual firm. We assume the stochastic discount factor satisfies conditions under which the multi-period discount factor can be represented as $(1+r_i^{\text{firm}})^{-t}$ for a constant firm-specific discount rate r_i^{firm} that reflects both time preference and systematic risk. These assumptions yield the value function,

$$P_i(K) = \sum_{t=1}^{\infty} \frac{1}{(1+r_i^{\text{firm}})^t}[Y_i(K) - \gamma K] = \frac{Y_i(K) - \gamma K}{r_i^{\text{firm}}}. \tag{23}$$

Result 1 (An Inverted U Between Market Value and Cost of Capital Wedges)**.** *The market value of firm i is given by*

$$P_i(K_i^*) = P_i(\bar{K}_i) - \frac{\delta_i^2}{2\beta_i r_i^{firm}}, \tag{24}$$

where $P_i(\bar{K}_i)$ is the market value of the firm when the cost of capital deviation is zero, $\delta_i = r_i^{perc.} - r_i^{firm} = 0$.

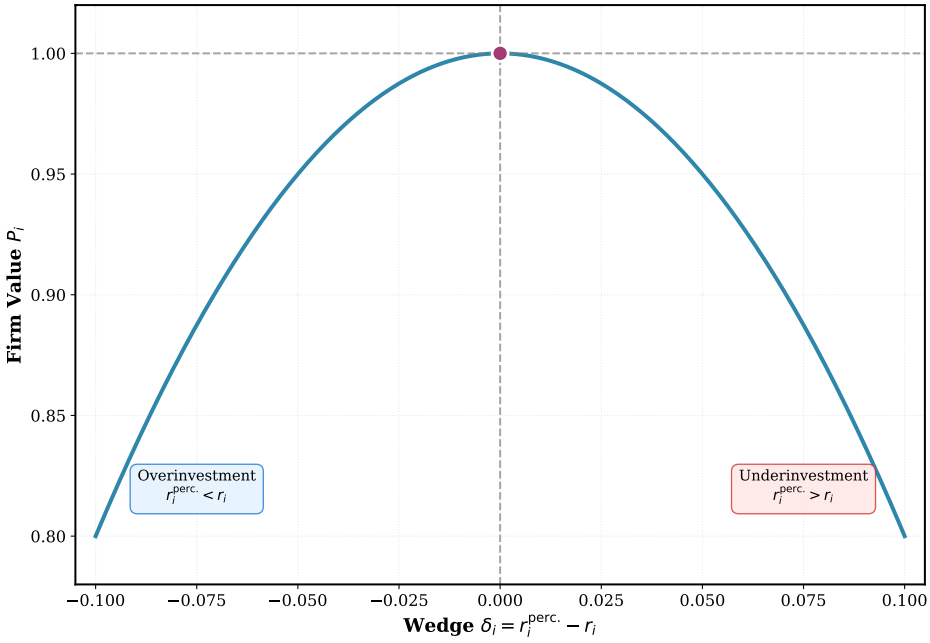
Result 1 implies an inverted U relation: firm market value is maximized when firms use the same cost of capital as financial markets, r_i^{firm}, in their investment decisions. If firms use $r_i^{\text{perc.}}$ below or above r_i^{firm}, market value decreases. The inverted U is illustrated in Figure 5.

Our second result characterizes the relationship between investment and firm value in

Figure 5

Model: Inverted U between firm market value and deviations from MCC

This figure shows the relation between market value and cost of capital deviations in our model. The model is calibrated such that market value is \$1 when the cost of capital deviation is 0.



the cross-section of firms. From equation (6), we have $I_i^* = \gamma K_i^* = \gamma \bar{K}_i - \gamma \delta_i / \beta_i$. Since investment is proportional to capital, the correlation between investment and market value depends on the sign of the deviation.

Result 2 (Investment-value correlation)**.** *In the cross-section of firms, the correlation between investment and value depends on the sign of firms' cost of capital deviation*

$$Corr(I_i, P_i) > 0 \text{ for firms for which } \delta_i > 0$$
$$Corr(I_i, P_i) < 0 \text{ for firms for which } \delta_i < 0$$

Result 2 shows that the comovement between market values and investment depends on the sign of the cost of capital deviation. Firms with a positive deviation underinvest, and higher investment is therefore met with a positive response from investors. Firms with a negative deviation overinvest, and higher investment is therefore met with a negative response from investors.

6.2 Empirical results on inverted U

We investigate the relation between firm equity market value and the cost of capital deviation in the data (result 1 in the model). To measure the deviation, we calculate the difference between a firm's perceived cost of capital and its MCC. The "perceived cost of capital" is the cost of capital that influences the firm's real investment decisions, measured by Gormsen and Huber (2025b). The data are hand-collected from conference calls between firm managers, financial investors, and analysts. They measure explicit statements by managers on the calls about their internal perceptions of their firm-level weighted average cost of capital. In the data, firms with a higher perceived cost of capital have higher returns on invested capital and lower investment rates, consistent with the view that the perceived cost of capital influences capital allocation in the long run. Figure 6 shows a histogram of the cost of capital deviation.

We study which firms use a given amount of book equity more effectively to generate a high market value. We therefore analyze whether firms with similar book equity are valued differently in equity markets depending on their cost of capital deviation. Figure 6 illustrates the average log market value for firms in different bins of the cost of capital deviation. The figure plots point estimates from a regression of log market value on the perceived cost of capital. The regression additionally controls for fixed effects for the firm's book equity decile interacted with country fixed effects, so the point estimates capture differences within firms of similar book equity size. In addition, we control for quarter-by-year-by-country fixed effects, so macroeconomic shocks do not bias the estimates.

Figure 6 reveals an inverted U relation. Firm market value is lower when the cost of capital perceived by the firm (perc. CoC) deviates more from the MCC. This finding suggests that the cost of capital used by firms is strongly associated with market value, consistent with the classical intuition about the opportunity cost of capital and modern models of firm behavior.

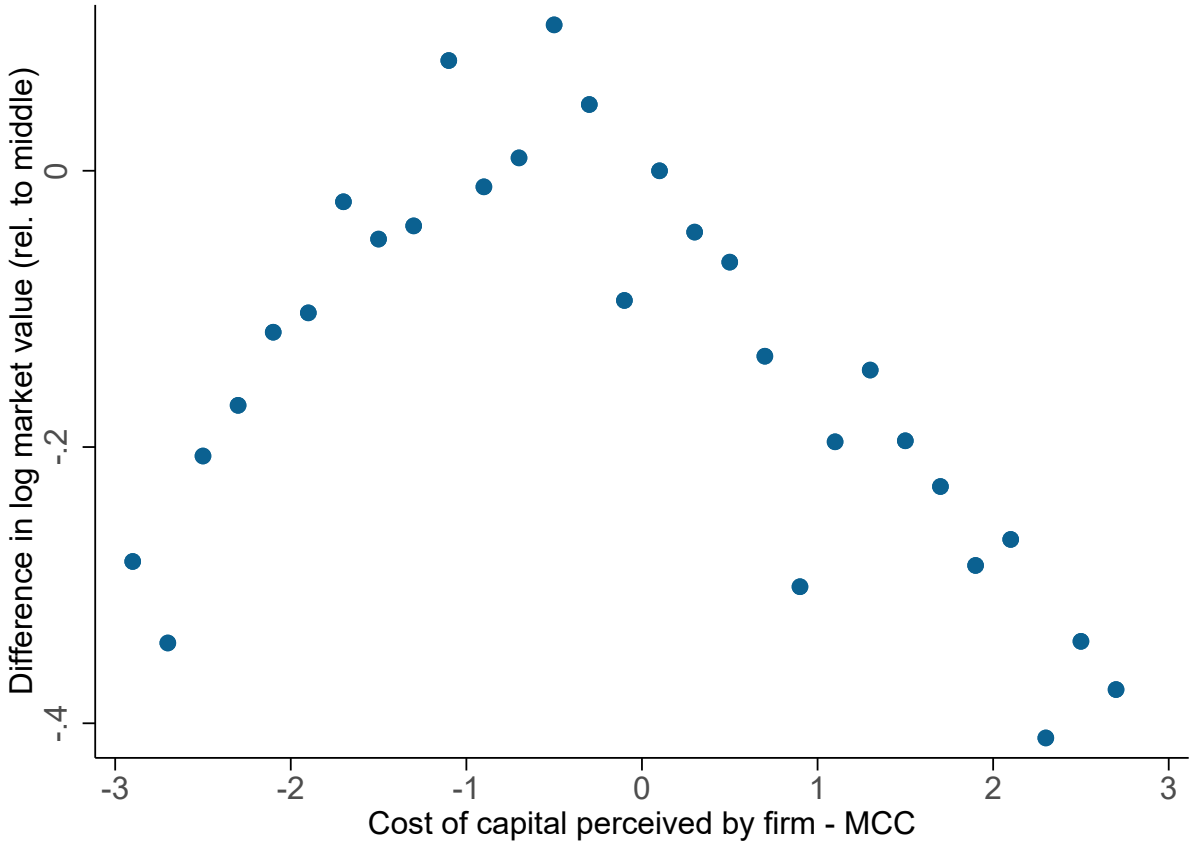
Table 3 suggests that the association between the cost of capital deviation and market value is statistically and economically significant. In columns 1 to 2, the regressor is the absolute value of the deviation. The point estimates are all significant at the 1 percent level and of similar magnitude in the U.S. and the full sample.

In columns 3 to 5, we use the simple subtraction as the regressor. A 1 percentage point increase in the deviation is associated with a 8.7 log point value increase in the range of relatively large negative deviations (below -0.3 percentage points) analyzed in column 3. In contrast, a 1 percentage point increase in the deviation is associated with a 7 log point value decrease in the range of relatively large positive deviations (above 0.3 percentage points)

Figure 6

Inverted U: firm market value and deviations from the MCC

The figure shows the relation between firm market value and the cost of capital deviation (i.e., the difference between perc. CoC and MCC). The perc. CoC is reported by the firm on conference calls and influences the firm's investment decisions. MCC is the market-based cost of capital from this paper. We construct bins for the cost of capital deviation in percentage points. The bins are 0.2 percentage points wide, except for the leftmost bin, which includes all observations below -2.8 (roughly the 1st percentile), and the rightmost bin, which pools all observations above 2.8. The figure plots point estimates from a regression of firm log market value on indicators for the bins. The point estimates give the difference in log market value relative to the middle bin (difference just above 0). The regression controls for fixed effects for book equity deciles-by-country and quarter-by-year-by-country. Book equity deciles are based on the firm's median decile in the distribution of book equity relative to other firms in the same quarter-year. Table 3 reports related regressions.



	(1) U.S.	(2) Global	(3) $\leq$ -0.3	(4) $\geq$ 0.3	(5) $>$ -0.3, $<$ 0.3
| Perc. CoC − MCC |	-0.081***	-0.096***			
	(0.019)	(0.018)			
Perc. CoC − MCC			0.087*	-0.070***	-0.173
			(0.049)	(0.025)	(0.210)
Observations	3,098	5,204	1,509	3,055	640
Within R^2	0.0185	0.0266	0.00924	0.0154	0.00240

The table reports regressions of firm log market value on the deviation between the cost of capital perceived by the firm and the MCC. Perc. CoC is the cost of capital reported by the firm on conference calls that influences the firm's investment decisions. MCC is the market-based cost of capital from this paper. The regressor in columns 1 to 2 is the absolute difference between Perc. CoC and MCC. The regressor in columns 3 to 5 is the simple subtraction. The sample includes only U.S. firms in column 1; all firms in column 2; only firms where the simple subtraction is at most 0.3 percentage points in column 3 (i.e., firms where the perc. CoC is low relative to the MCC); only firms where the simple subtraction is at least 0.3 in column 4; and firms where the simple subtraction is close to 0 in column 5. All specifications control for fixed effects for book equity deciles-by-country and quarter-by-year-by-country. Book equity deciles are based on the firm's decile in the distribution of book equity relative to other firms in the same quarter-year. Standard errors are clustered at the firm level. Statistical significance is denoted by *** p<0.01, ** p<0.05, * p<0.1.

Table 3: Inverted U: firm market value and deviations from the MCC

analyzed in column 4. The magnitude is thus similar on both sides. The point estimate is statistically insignificant when the deviations are close to zero in column 5, albeit statistically noisy. The pattern of results is consistent with the model and the inverted U shape in Figure 5.

Taken together, the empirical results suggest that firms using an unbiased perceived cost of capital, like the MCC, are valued more positively in equity markets.

6.3 Empirical results on investment announcements

We analyze how firm equity market values respond to firms announcing new investment projects (result 2 in the model). If the firm's cost of capital deviation is above zero, the firm's perceived cost of capital is "too high" compared to the MCC and the firm invests relatively little compared to the cost of capital effectively used by the equity market. An announcement of a firm with a positive deviation should therefore lead to a relatively higher market value compared to an announcement of a firm with negative deviation.

Investment announcements are available on the S&P Capital IQ Key Developments

	(1) U.S.	(2) Global	(3) U.S.	(4) Global
(Perc. CoC − MCC) > 0	0.199**	0.199**		
	(0.100)	(0.098)		
Perc. CoC − MCC			0.032*	0.034*
			(0.019)	(0.019)
Observations	414	709	414	709
Within R^2	0.0121	0.0107	0.00472	0.00465

The table analyzes firm's equity market returns in windows around real investment announcements by the firm. Perc. CoC is the cost of capital reported by the firm on conference calls that influences the firm's investment decisions. MCC is the market-based cost of capital from this paper. The regressor in columns 1 to 2 is an indicator for a positive difference between Perc. CoC and MCC. The regressor in columns 3 to 4 is the simple subtraction. The outcome is the abnormal stock return of the firm (firm return minus market return) in a window of -3 to +1 days around an investment announcement reported on the Key Developments Dataset. The outcome is winsorized at the 1^{st} and 99^{th} percentiles. The sample includes only U.S. firms in columns 1 and 3 and all firms in columns 2 and 4. All specifications control for fixed effects for book equity deciles-by-country and quarter-by-year-by-country. Book equity deciles are based on the firm's decile in the distribution of book equity relative to other firms in the same quarter-year. Standard errors are clustered at the firm level. Statistical significance is denoted by *** p<0.01, ** p<0.05, * p<0.1.

Table 4: Stock returns and investment announcements

database, which reports all types of news about firms. We consider all firms for which we observe a perceived cost of capital and the MCC in the same quarter. We prompt Claude AI to extract all news which are unambiguously new announcements of investment projects, not just minor updates on existing projects. We require the investment to be a concrete, specific project undertaken by the firm itself. The investment needs to be the central focus of the news, rather than described only in passing, which avoids classification errors and makes it more likely that the announcement reveals new information. Store openings, site breaking grounds, or product launches are not included, since the investment has already been made in the past in these cases. We also exclude M&A events because they are not typically evaluated by the equity market using the existing firms' MCC. Overall, these restrictions leave 709 firms in our estimation sample.

To analyze the equity market response, we construct abnormal stock returns as the firm's stock return minus the market return in a window of -3 days to 1 day around an investment

announcement. We regress the abnormal stock return on an indicator for a positive cost of capital deviation. The point estimates in columns 1 and 2 of Table 4 suggest that the abnormal return for firms with a positive deviation is 0.2 percentage points higher, both in the U.S. and the global sample. The coefficients are significant at the 5 percent level.

We also find positive and significant coefficients when using the simple subtraction of MCC from perceived cost of capital as regressors in columns 3 and 4. The regressions include the baseline controls of book equity deciles and quarter-by-year fixed effects, interacted with country fixed effects, to ensure that size-specific differences across firms and macroeconomic shocks do not bias the results. In Table 4, we report similar results dropping the size controls in column 2 and using an unwinsorized abnormal stock return as the outcome in column 3.

The findings suggest that the equity market response to new investment announcements depends on the cost of capital deviation. The MCC seems to capture well the cost of capital effectively used by the equity market and therefore how positively the equity market responds to a new announcement, in line with the model. More generally, the announcement results further support the conclusion that the cost of capital used by firms is a key determinant of firm market values.

7 Conclusion

We estimate the firm-level cost of capital using market prices and new machine-learning forecasts of future cash flows. The key challenge is that the textbook cost of capital that firms should use in investment decisions is not directly observed. Existing empirical proxies, including the CAPM and analyst-based implied cost of capital measures, are biased. Our approach combines forecasts of future cash flows with a valuation equation that backs out the discount rate consistent with observed market values. The result is an unbiased market-based cost of capital (MCC).

Empirically, we show that the method performs well along two dimensions. First, the cash flow forecasts exhibit out-of-sample predictive power, especially relative to widely used benchmarks. Second, the market-based cost of equity strongly predicts future long-run realized returns. These findings suggest that the MCC captures variation in expected returns well.

We then address how the cost of capital used by firms matters for firm market values. Using hand-collected data on firms' perceived cost of capital, we find that market value is higher when firms' perceived cost of capital is closer to the MCC, consistent with the classical

view that firms create value by aligning investment decisions with the opportunity cost of capital faced by investors.

References

Altman, E. I. and V. M. Kishore (1998): "Default and returns on high yield bonds: analysis through 1997," *NYU Salomon Center Report, January1995*.

Ball, R. and R. Watts (1972): "Some time series properties of accounting income," *The Journal of Finance*, 27, 663–681.

Bauer, M. D. and G. D. Rudebusch (2020): "Interest Rates under Falling Stars," *American Economic Review*, 110, 1316–54.

Bessembinder, H. (2018): "Do stocks outperform treasury bills?" *Journal of financial economics*, 129, 440–457.

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

Binsbergen, J. v., S. Hua, J. Peeters, and J. Wachter (2025): "Is the United States a lucky survivor? A hierarchical Bayesian approach," *The Journal of Finance*, 80, 2355–2388.

Campbell, J. Y., C. Polk, and T. Vuolteenaho (2010): "Growth or glamour? Fundamentals and systematic risk in stock returns," *The Review of Financial Studies*, 23, 305–344.

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

Cho, T., L. Kremens, D. Lee, and C. Polk (2024): "Scale or yield? A present-value identity," *The Review of Financial Studies*, 37, 950–988.

Claus, J. and J. Thomas (2001): "Equity Premia as Low as Three Percent? Evidence from Analysts' Earnings Forecasts for Domestic and International Stock Markets," *Journal of Finance*, 56, 1629–1666.

De Silva, T. and D. Thesmar (2024): "Noise in expectations: Evidence from analyst forecasts," *The Review of Financial Studies*, 37, 1494–1537.

Decaire, P. and M. Guenzel (2023): "What Drives Very Long-Run Cash Flow Growth Expectations?" *Jacobs Levy Equity Management Center for Quantitative Financial Research Paper*.

Dick-Nielsen, J., P. Feldhütter, L. H. Pedersen, and C. Stolborg (2023): "Corporate bond factors: Replication failures and a new framework," *Available at SSRN 4586652*.

Easton, P. D. (2004): "PE Ratios, PEG Ratios, and Estimating the Implied Expected Rate of Return on Equity Capital," *Accounting Review*, 79, 73–95.

Edwards, E. O. and P. W. Bell (1961): *The theory and measurement of business income*, Univ of California Press.

Eskildsen, M., M. Ibert, T. I. Jensen, and L. H. Pedersen (2024): "In search of the true greenium," *Available at SSRN*.

Fama, E. F. and K. R. French (1997): "Industry costs of equity," *Journal of financial economics*, 43, 153–193.

Foster, G., C. Olsen, and T. Shevlin (1984): "Earnings releases, anomalies, and the behavior of security returns," *Accounting Review*, 574–603.

Gebhardt, W. R., C. M. C. Lee, and B. Swaminathan (2001): "Toward an Implied Cost of Capital," *Journal of Accounting Research*, 39, 135–176.

Gormsen, N. J. and K. Huber (2025a): "Corporate Discount Rates," *American Economic Review*, 115, 2001–2049.

——— (2025b): "Firms' Perceived Cost of Capital," .

Graham, J. R. (2022): "Presidential Address: Corporate Finance and Reality," *Journal of Finance*, 77, 1975–2049.

Graham, J. R. and C. R. Harvey (2001): "The Theory and Practice of Corporate Finance: Evidence from the Field," *Journal of Financial Economics*, 60, 187–243.

Hillenbrand, S. and O. McCarthy (2023): "The Optimal Stock Valuation Ratio," *Available at SSRN 4288780*.

Hommel, N., A. Landier, and D. Thesmar (2025): "Corporate Valuation: An Empirical Comparison of Discounting Methods," .

Jacobs, M. T. and A. Shivdasani (2012): "Do You Know Your Cost of Capital?" *Harvard Business Review*, 90, 118–124.

Jagannathan, R., D. A. Matsa, I. Meier, and V. Tarhan (2016): "Why Do Firms Use High Discount Rates?" *Journal of Financial Economics*, 120, 445–463.

Jensen, T. I., B. Kelly, and L. H. Pedersen (2023): "Is there a replication crisis in finance?" *The Journal of Finance*, 78, 2465–2518.

Ke, G., Q. Meng, T. Finley, T. Wang, W. Chen, W. Ma, Q. Ye, and T.-Y. Liu (2017): "Lightgbm: A highly efficient gradient boosting decision tree," *Advances in neural information processing systems*, 30.

Keloharju, M., J. T. Linnainmaa, and P. Nyberg (2021): "Long-Term Discount Rates Do Not Vary Across Firms," *Journal of Financial Economics*, 141, 946–967.

Koller, T., M. Goedhart, and D. Wessels (2020): *Valuation: Measuring and Managing the Value of Companies*, John Wiley & Sons, 7 ed.

Martineau, C. (2022): "Rest in peace post-earnings announcement drift," *Critical Finance Review*, 11, 613–646.

Mohanram, P. and D. Gode (2013): "Removing predictable analyst forecast errors to improve implied cost of equity estimates," *Review of Accounting Studies*, 18, 443–478.

Mukhlynina, L. and K. G. Nyborg (2016): "The Choice of Valuation Techniques in Practice: Education Versus Profession," .

Ohlson, J. A. (1995): "Earnings, book values, and dividends in equity valuation," *Contemporary accounting research*, 11, 661–687.

Ohlson, J. A. and B. E. Juettner-Nauroth (2005): "Expected EPS and EPS growth as determinantsof value," *Review of accounting studies*, 10, 349–365.

Richardson, S., S. H. Teoh, and P. D. Wysocki (2004): "The walk-down to beatable analyst forecasts: The role of equity issuance and insider trading incentives," *Contemporary accounting research*, 21, 885–924.

Rosenbaum, J. and J. Pearl (2021): *Investment banking: valuation, LBOs, M&A, and IPOs*, John Wiley & Sons.

Ross, S., R. Westerfield, J. Jaffe, B. Jordan, and K. Shue (2022): *Corporate Finance*, McGraw-Hill Education, 13th ed.

Shwartz-Ziv, R. and A. Armon (2022): "Tabular data: Deep learning is not all you need," *Information Fusion*, 81, 84–90.

Stein, J. C. (1996): "Rational Capital Budgeting in an Irrational World," *Journal of Business*, 69, 429–455.

Trahan, E. A. and L. J. Gitman (1995): "Bridging the Theory-Practice Gap in Corporate Finance: A Survey of Chief Financial Officers," *Quarterly Review of Economics and Finance*, 35, 73–87.

Van Binsbergen, J. H., X. Han, and A. Lopez-Lira (2023): "Man versus machine learning: The term structure of earnings expectations and conditional biases," *The Review of financial studies*, 36, 2361–2396.

Vuolteenaho, T. (2002): "What drives firm-level stock returns?" *The Journal of Finance*, 57, 233–264.

Zhang, Y., Y. Zhu, and J. T. Linnainmaa (2024): "Man versus Machine Learning Revisited," *Available at SSRN 4899584*.

Supplemental Appendix

Appendix A Machine-Learned Cash Flow Forecasts

This section documents the implementation details of the machine learning models used to forecast firm-level cash flows. We cover the output targets, the explanatory variables, the model specification, hyperparameter tuning procedure, temporal sample splits, and the final refit strategy.

Appendix A.1 Output targets

We forecast two cash flow measures—net income (NI) and free cash flow to equity (FCFE)—at each of 10 annual horizons, yielding 20 target variables per firm-month. The following describes how each target is constructed.

- The valuation equations assume that the first cash flow occurs exactly one year from now, and that cash flows thereafter occur in one-year increments. However, fiscal year ends do not necessarily allow for this one-year constant horizon. We handle this by defining a constant-horizon cash flow over 1 year that interpolates between the adjacent accounting reports using inverse-distance weighting: for a target date falling between two fiscal year-ends, the weight on each observation is proportional to its proximity to the target date. When only one fiscal year observation is available, it is used directly. Equal weighting arises as the special case when the target date is equally far from both fiscal year-ends.

- Fiscal year ends differ across companies, which can create a mismatch when comparing annual earnings across companies. Our constant-horizon approach, described in the previous bullet point, handles this issue.

- The model draws on data from 93 countries, so to address currency mismatches, we convert all cash flows to U.S. dollars using the exchange rate at fiscal period-end. We obtain exchange rates from Compustat.

- The inflation rate varies a lot over our sample. Therefore, we adjust the dependent variable for the realized inflation between the forecast date and the cash flow realiza-

tion, using the `CPIAUCNS` series from FRED. This means our model effectively forecasts real cash flows.

- When forecasting long-run cash flow, we face the problem that not all firms available at the forecast date have realized cash flows over all forecasting horizons. For example, the company may have ben delisted or taken private. One approach would be to fit the model only to observations from firms with non-missing realized cash flows, but this would introduce a survivorship bias.[A1] Instead, we keep all firms that were alive on the forecast date, and impute missing cash flows by carrying forward the last non-missing cash flow with the firm's stock return, if available, and zero otherwise. By using the stock return, we also incorporate delisting returns, which is a flexible way of handling both good and bad delisting reasons.[A2]

- Equity issuances after the forecast date used, say, to acquire another company, could create a mechanical boost in future cash flow (the dependent variable) that would not necessarily accrue to the firm's current owners (because their ownership share would be diluted due to the equity issuance). Therefore, we divide the realized cash flows at all points by the number of shares outstanding, such that the dependent variable is the realized cash flow from owning one share at the forecast date.

- Working with a per-share variable, we need to adjust for stock splits that mechanically change the per share value without any change in the fundamentals (e.g., a two-for-one split would cut the cash flow per share in half, but since owners get two shares for each share they owned, they are fundamentally unaffected). We make this adjustment using the `CFACSHR` variable from CRSP for U.S. stocks, and the `AJEXDI` variable from Compustat for non-U.S. stocks.

- Per share variables are somewhat arbitrary in the sense that the company can choose its number of shares outstanding, and this choice will lead to a difference in scale across companies. To handle this issue, we divide all realized cash flows by the most recent asset per share at the forecast date.

- Cash flow variables are computed from both annual and quarterly Compustat. When both are available for the same firm-date, the annual value is used.

[A1] For example, if firms typically drop out of the sample for performance-related reasons, conditioning on firms that survive would introduce an upward bias in our cash flow forecasts.

[A2] For example, a firm that goes bankrupt will typically see its cash flows approach zero, whereas a firm that gets acquired will not.

- All target variables are winsorized at the 1st and 99th percentiles within each horizon-month combination.

Appendix A.2 Input features

Each prediction model uses 244 continuous and 138 categorical input features, grouped into four categories. Table A1 provides the full list of features by category.

JKP characteristics (148). We use 153 firm characteristics from Jensen et al. (2023), of which 5 are excluded as duplicates with our cash flow features (`gp_at`, `ocf_at`, `dbnetis_at`, `at_turnover`, `eqnetis_at`). These cover a broad range of anomaly categories, like including momentum, value, profitability, investment, and trading frictions. Most variables a kept at their raw values, but three non-stationary features—`market_equity`, `dolvol_126d`, and `age`—are cross-sectionally rank-transformed each month. We refer to Jensen et al. (2023) for detailed definitions.

Cash flow features (84). We construct 21 additional accounting items from Compustat, each divided by total assets to normalize scale: sales, cost of goods sold, gross profit, SG&A (with and without R&D), R&D, EBITDA, depreciation & amortization, EBIT, net income, operating cash flow, free cash flow, free cash flow to equity, unlevered free cash flow, capital expenditures, change in net operating working capital, net debt issuance, dividends, unlevered net income, interest expense, and the marginal tax rate. For each of these 21 variables, we compute 1-, 2-, and 3-year changes, yielding 63 additional features that capture recent trends. The total is $21 + 63 = 84$ cash flow features.

IBES analyst forecasts (12). We use consensus EPS forecasts for fiscal years 1–3 and the long-term growth rate (LTG) from I/B/E/S, each scaled by assets per share. For each of these 4 variables, we also compute the industry median (Fama-French 49) and the market-wide median, yielding 12 features in total. Industry and market medians require at least 10 non-missing observations.

Categorical features. Fama-French 49 industry classification (45 active categories, excluding the 4 financial industries) and country (93 countries) are passed as native LightGBM categoricals rather than one-hot encoded.

Missing values. LightGBM handles missing values natively by learning optimal split directions for missing inputs. We therefore retain all observations regardless of missing explanatory variables.

Table A1: Input Features to Cash Flow Prediction Models

Category	Variables
JKP (148)	age, aliq_at, aliq_mat, ami_126d, at_be, at_gr1, at_me, be_gr1a, be_me, beta_60m, beta_dimson_21d, betabab_1260d, betadown_252d, bev_mev, bidaskhl_21d, capex_abn, capx_gr1, capx_gr2, capx_gr3, cash_at, chcsho_12m, coa_gr1a, col_gr1a, cop_at, cop_atl1, corr_1260d, coskew_21d, cowc_gr1a, debt_gr3, debt_me, dgp_dsale, div12m_me, dolvol_126d, dolvol_var_126d, dsale_dinv, dsale_drec, dsale_dsga, earnings_variability, ebit_bev, ebit_sale, ebitda_mev, emp_gr1, eq_dur, eqnpo_12m, eqnpo_me, eqpo_me, f_score, fcf_me, fnl_gr1a, gp_atl1, inv_gr1, inv_gr1a, iskew_capm_21d, iskew_ff3_21d, iskew_hxz4_21d, ival_me, ivol_capm_21d, ivol_capm_252d, ivol_ff3_21d, ivol_hxz4_21d, kz_index, lnoa_gr1a, lti_gr1a, market_equity, mispricing_mgmt, mispricing_perf, ncoa_gr1a, ncol_gr1a, netdebt_me, netis_at, nfna_gr1a, ni_ar1, ni_be, ni_inc8q, ni_ivol, ni_me, niq_at, niq_at_chg1, niq_be, niq_be_chg1, niq_su, nncoa_gr1a, noa_at, noa_gr1a, o_score, oaccruals_at, oaccruals_ni, ocf_at_chg1, ocf_me, ocfq_saleq_std, op_at, op_atl1, ope_be, ope_bel1, opex_at, pi_nix, ppeinv_gr1a, prc, prc_highprc_252d, qmj, qmj_growth, qmj_prof, qmj_safety, rd5_at, rd_me, rd_sale, resff3_12_1, resff3_6_1, ret_12_1, ret_12_7, ret_1_0, ret_3_1, ret_60_12, ret_6_1, ret_9_1, rmax1_21d, rmax5_21d, rmax5_rvol_21d, rskew_21d, rvol_21d, sale_bev, sale_emp_gr1, sale_gr1, sale_gr3, sale_me, saleq_gr1, saleq_su, seas_11_15an, seas_11_15na, seas_16_20an, seas_16_20na, seas_1_1an, seas_1_1na, seas_2_5an, seas_2_5na, seas_6_10an, seas_6_10na, sti_gr1a, taccruals_at, taccruals_ni, tangibility, tax_gr1a, turnover_126d, turnover_var_126d, z_score, zero_trades_126d, zero_trades_21d, zero_trades_252d
CF levels (21)	sale_at, cogs_at, gp_at, xsga_at, xrd_at, sga_at, ebitda_at, dp_at, ebit_at, ni_at, ocf_at, fcf_at, fcfe_at, ufcf_at, capx_at, dNOWC_at, dvt_at, uni_at, xint_at, dbnetis_at, mtr

Input Features to Cash Flow Prediction Models (continued)

Category	Variables
CF changes (63)	1-, 2-, and 3-year changes of each of the 21 cash flow level variables (e.g., `sale_at_ch1`, `sale_at_ch2`, `sale_at_ch3`)
IBES (12)	`eps_fcst1`, `eps_fcst2`, `eps_fcst3`, `ltg`, `eps_fcst1_ind`, `eps_fcst2_ind`, `eps_fcst3_ind`, `ltg_ind`, `eps_fcst1_mkt`, `eps_fcst2_mkt`, `eps_fcst3_mkt`, `ltg_mkt`
Industry (45)	Fama-French 49 industry classification, excluding the 4 financial industries (banks, insurance, real estate, trading)
Country (93)	Country indicator

This table lists the input features used in the cash flow prediction models, grouped by category. The number in parentheses shows the number of variables. JKP variables are firm characteristics from Jensen et al. (2023). CF levels are accounting items scaled by total assets. CF changes are 1-, 2-, and 3-year changes of the CF level variables. IBES variables are I/B/E/S consensus analyst forecasts at the firm, industry, and market level. Industry and country are native LightGBM categorical features.

Appendix A.3 ML model

Model We use gradient-boosted decision trees implemented in LightGBM (Ke et al., 2017), with boosting type `gbdt`, a regression objective, and RMSE as the evaluation metric. We fit a separate model for each cash flow measure (free cash flow to equity and net income) at each of the ten forecast horizons, for a total of 20 models. Each model pools observations across all countries.

Hyperparameter search space We search over five hyperparameters using a 20-point max-entropy space-filling design. Table A2 summarizes the search space.

Two-stage tuning For each annual training window, we tune hyperparameters in two stages:

1. **Stage 1 (tree structure).** We evaluate all 20 candidate hyperparameter sets with a learning rate of $\eta_1 = 0.3$ and up to 1,000 boosting rounds. Training stops early if the validation RMSE does not improve for 25 consecutive rounds. The hyperparameter set with the lowest validation RMSE is selected.

Parameter	Range	Scale
`feature_fraction`	$[1/P,\ 1]$	Linear
`num_leaves`	$[1,\ 127]$	Linear
`min_data_in_leaf`	$[1,\ 500]$	Linear
`bagging_fraction`	$[0.2,\ 1.0]$	Linear
`lambda_l2`	$[10^{-4},\ 10^{5}]$	Log

This table reports the hyperparameter search space for LightGBM. Each model is tuned over a 20-point max-entropy space-filling design. P denotes the total number of features. `num_leaves` $= 1$ to 127 corresponds to tree depths of 1 to 7 via $2^d - 1$. Two additional parameters are held fixed: `lambda_l1` $= 0$ (no L1 regularization) and `max_depth` $= -1$ (no explicit depth limit, as tree complexity is controlled through `num_leaves`).

Table A2: Hyperparameter Search Space

2. **Stage 2 (iteration count).** We take the winning hyperparameter set from Stage 1 and reduce the learning rate to $\eta_2 = 0.05$. The model is trained for up to 10,000 boosting rounds with the same early stopping rule (25 rounds). This stage determines the optimal number of boosting iterations at the lower learning rate.

The two-stage design balances computational cost with model quality: the coarse first stage efficiently identifies good tree structures, while the fine second stage precisely calibrates the ensemble size.

Appendix A.4 Train, validation, and test splits

We create a temporal train-validation-test sample split, where the train and validation data is used to find hyperparameters, and the test sample is used to make out-of-sample forecasts.

- **Test set.** A one-year window of observations for which we generate forecasts (e.g., January–December 1984 for the first split).
- **Validation set.** All observations whose target realization period ends before the test set begins, going back 10 years.
- **Training set.** All remaining observations whose target realization period ends before the validation set begins.

Once the hyperparameters have been selected, we refit the model on the combined training and validation data using the optimal tree parameters from the first stage hyperparameter

tuning, the learning rate of 0.05, and the optimal number of rounds from the second stage hyperparameter tuning.

The model parameters are updated every year, using an expanding training window. Specifically, every year, we add one year to the training set, while the validation and test set start and end dates are increased by one year. For the first split, all pre-test data ends on December 31, 1983, and splits are re-estimated annually through December 31, 2023. This yields 40 annual re-estimations per cash flow measure and horizon.

Appendix B The Market-Based Cost of Equity

This section documents the implementation choices involved in computing the market-based cost of equity (MCE) from the machine-learned cash flow forecasts, using the valuation frameworks in (6) and (9).

Terminal growth rate. The terminal growth rate for cash flows is the country-specific 5-year forward real GDP growth rate (that is, the expected real growth betwen year $t+4$ and $t+5$) from the IMF World Economic Outlook. We use a forward rate rather than a cumulative rate because the terminal value should reflect steady-state growth. For example, during the Global Financial Crisis, cumulative forecasts over any horizon were depressed by the immediate recession, whereas the forward rate between years 4 and 5 was much less affected, better reflecting long-run growth prospects. The IMF series begins in April 1990; for earlier dates, we use the April 1990 forecast for each country.

Forecast smoothing. Before computing the MCE, the 10-horizon raw per-share forecasts for each firm-month are smoothed by projecting them onto a quadratic basis $[1, h, h^2]$ via OLS, where $h = 1, \ldots, 10$ is the forecast horizon.

Half-year discounting. Cash flows are discounted at $(1+r)^{h-0.5}$ rather than $(1+r)^h$, reflecting the assumption that cash flows arrive at the midpoint of each period rather than at the end.

Root-finding. The cost of equity is found via vectorized bisection with a convergence tolerance of 10^{-8} and a maximum of 100 iterations. The search interval is $[\max(-0.5,\ g + 0.001),\ 0.5]$, where g is the terminal growth rate. This ensures that the discount rate strictly exceeds the terminal growth rate, as required for the terminal value to be meaningful.

Validity screening. Cost of equity estimates are set to missing when any of the following conditions hold: (a) the net present value increases with the discount rate; (b) no root exists because the NPV has the same sign at both bounds; or (c) the implied real cost of equity exceeds 10%

Real to nominal conversion. Because the cash flow forecasts are in real terms and the terminal growth rate is a real GDP growth rate, the root-finding procedure yields a real cost of equity. We convert to a nominal cost of equity by adding the expected U.S. inflation rate. Expected inflation is the 5-year forward CPI inflation rate from the IMF World Economic Outlook, matching the source used for the terminal growth rate. Since all cash flows are denominated in U.S. dollars, we use the U.S. forecast for all firms. For dates before the IMF series begins (April 1990), we use the trend inflation rate from Bauer and Rudebusch (2020).

Combining NI and CF estimates. The combined cost of equity first demeans each model type (residual income and free cash flow to equity) within each country-type-month group to remove level differences across models, averages the demeaned values, and then adds back the overall country-month mean. When only one model produces a valid estimate for a given firm-month, that estimate is used directly.

Non-US start date. Non-U.S. firms before January 1996 are excluded from the MCE sample. The non-U.S. accounting data starts in 1985, so this screen ensures that the cash flow models have at least ten years of non-U.S. data in their training data.

Appendix C Returns and Cash Flow Benchmarks

In this section we describe the methods that we use as benchmarks for our cash flow and returns forecasts.

Appendix C.1 Analyst-based ICC methods

We implement four analyst-based implied cost of capital methods following Eskildsen et al. (2024): Ohlson and Juettner-Nauroth (2005, OJ) and Easton (2004, PEG), which are based on the dividend discount model, and Gebhardt et al. (2001, GLS) and Claus and Thomas (2001, CT), which are based on the residual income model.

All four methods use analyst forecasts of earnings per share (EPS) from the I/B/E/S consensus file (median forecast). We denote the one- and two-year EPS forecasts as $\hat{\mathrm{E}}_t[e_{t+1}]$ and $\hat{\mathrm{E}}_t[e_{t+2}]$, the long-term growth forecast as LTG, the book equity per share as b_t, and

the stock price as p_t. When the two-year EPS forecast is unavailable, we set $\hat{\mathrm{E}}_t[e_{t+2}] = \hat{\mathrm{E}}_t[e_{t+1}] \times 1.15$. We estimate the payout ratio as dividends divided by earnings from the last fiscal year. For firms with negative earnings, we follow Mohanram and Gode (2013) and set the payout ratio to 6% of total assets.

OJ. The Ohlson and Juettner-Nauroth (2005) method provides a closed-form solution for the implied cost of capital:

$$\mathrm{ICC}^{OJ} = A + \sqrt{A^2 + \frac{\hat{\mathrm{E}}_t[e_{t+1}]}{p_t} \times (\mathrm{STG} - \lambda)}, \tag{A1}$$

where

$$A = \frac{1}{2}\left(\lambda + \frac{\hat{\mathrm{E}}_t[d_{t+1}]}{p_t}\right) \quad \text{and} \quad \mathrm{STG} = \max\left[\left(\frac{\hat{\mathrm{E}}_t[e_{t+2}] - \hat{\mathrm{E}}_t[e_{t+1}]}{\hat{\mathrm{E}}_t[e_{t+1}]} \times \mathrm{LTG}\right)^{\frac{1}{2}}, \mathrm{LTG}\right]. \tag{A2}$$

Here, $\hat{\mathrm{E}}_t[d_{t+1}]$ is next year's expected dividend, computed as the payout ratio times $\hat{\mathrm{E}}_t[e_{t+1}]$. The parameter λ captures expected long-run economic growth and is set to the yield on a ten-year US treasury bond minus 3%.[A3]

PEG. The Easton (2004) method simplifies the OJ formula by setting $\lambda = 0$ and ignoring dividends:

$$\mathrm{ICC}^{PEG} = \sqrt{\frac{\hat{\mathrm{E}}_t[e_{t+1}]}{p_t} \times \mathrm{STG}}, \tag{A3}$$

with STG computed as in (A1).

Residual income framework. The GLS and CT methods share a common framework based on the residual income model:

$$p_t = b_t + \sum_{h=1}^{H}\left(\frac{\hat{\mathrm{E}}_t\left[(\mathrm{ROE}_{t+h} - r)b_{t+h-1}\right]}{(1+r)^h}\right) + \mathrm{TV}, \tag{A4}$$

where $\mathrm{ROE}_{t+h} = \hat{\mathrm{E}}_t[e_{t+h}]/b_{t+h-1}$ and book value evolves via clean surplus accounting: $b_t = b_{t-1} + e_t - d_t$, with $d_t = e_t \times$ payout-ratio. The two methods differ in the forecast horizon H, the earnings forecasting procedure, and the terminal value specification.

GLS. The Gebhardt et al. (2001) method forecasts residual income over 12 years, and value the terminal residual income as a constant perpetuity. The implied cost of capital

[A3] Ohlson and Juettner-Nauroth (2005) define STG as the two-year growth in earnings, $(\hat{\mathrm{E}}_t[e_{t+2}] - \hat{\mathrm{E}}_t[e_{t+1}])/\hat{\mathrm{E}}_t[e_{t+1}]$. To obtain a more stable estimate, we follow Mohanram and Gode (2013) and compute STG as the geometric mean of short- and long-term growth.

ICC^{GLS} solves:

$$p_t = b_t + \sum_{h=1}^{11}\left(\frac{\hat{\text{E}}_t\left[(\text{ROE}_{t+h} - \text{ICC}^{GLS})b_{t+h-1}\right]}{(1+\text{ICC}^{GLS})^h}\right) + \frac{\hat{\text{E}}_t\left[(\text{ROE}_{t+12} - \text{ICC}^{GLS})b_{t+11}\right]}{\text{ICC}^{GLS}(1+\text{ICC}^{GLS})^{11}}. \quad \text{(A5)}$$

For years 1–2, EPS forecasts come directly from I/B/E/S. For years 3–12, ROE converges linearly from the analyst-implied ROE_{t+2} to the industry median ROE. We compute the industry median ROE using the 49 industries from Fama and French (1997), measured in US dollars across all global firms with valid data.[A4]

CT. The Claus and Thomas (2001) method uses $H = 5$ explicit forecast periods and a growing perpetuity as terminal value. The implied cost of capital ICC^{CT} solves:

$$p_t = b_t + \sum_{h=1}^{5}\left(\frac{\hat{\text{E}}_t\left[(\text{ROE}_{t+h} - \text{ICC}^{CT})b_{t+h-1}\right]}{(1+\text{ICC}^{CT})^h}\right) + \frac{\hat{\text{E}}_t\left[(\text{ROE}_{t+5} - \text{ICC}^{CT})b_{t+4}\right](1+g)}{(\text{ICC}^{CT} - g)(1+\text{ICC}^{CT})^5}, \quad \text{(A6)}$$

where g is the terminal growth rate, set equal to the yield on a ten-year US treasury bond minus 3%. For years 1–2, EPS forecasts come from I/B/E/S. For years 3–5, EPS is grown from the second-year forecast using the LTG forecast from I/B/E/S.

Appendix C.2 The Market and the CAPM

The Market and CAPM models are used as benchmarks for our return forecasts. They make predictions over a ten-year horizon, matching the horizon in our main return benchmark analysis.

MKT. The market benchmark (which we use to compute out-of-sample R^2) is a country-level prediction based on the expanding-window average of realized equity risk premia. For each country, we compute rolling 10-year cumulative value-weighted market returns and subtract the 10-year risk-free rate compounded from the start of each window to obtain a 10-year equity risk premium (ERP):

$$\text{ERP}_{c,t} = \prod_{s=1}^{120}(1 + r^{mkt}_{c,t-120+s}) - (1 + r^{f,10}_{t-119})^{10}, \quad \text{(A7)}$$

where $r^{mkt}_{c,s}$ is the monthly value-weighted market return for country c and $r^{f,10}_{t-119}$ is the annualized 10-year risk-free rate at the start of the rolling window.

We then take an expanding-window average $\overline{\text{ERP}}_{c,t}$ of the realized 10-year ERP for

[A4] Most ICC papers estimate industry ROE using US firms only. Since our sample includes non-US firms, we use a global industry median for consistency.

each country, capped to the interval $[0, 3]$. Countries with insufficient history use the US expanding-window average as a fill-in. The market benchmark prediction for 10-year cumulative returns is:

$$\hat{R}_{c,t}^{\text{MKT}} = (1 + r_t^{f,10})^{10} - 1 + \overline{\text{ERP}}_{c,t}. \tag{A8}$$

CAPM. The CAPM benchmark is a firm-level prediction that scales the country-level equity risk premium by each firm's market beta. We estimate β_i as the 60-month rolling CAPM beta from firm-level returns, capped to $[0, 3]$. The CAPM prediction for 10-year cumulative returns is:

$$\hat{R}_{i,t}^{\text{CAPM}} = (1 + r_t^{f,10})^{10} - 1 + \beta_i \times \overline{\text{ERP}}_{c,t}, \tag{A9}$$

where $\overline{\text{ERP}}_{c,t}$ is the same country-level expanding-window average used in the market benchmark.

Appendix C.3 Machine-learned returns (MLR)

A key return benchmark is a machine learning model that predicts returns directly, using a machine learning setup that closely follows our approach for forecasting cash flows (Appendix A), with the following differences:

1. **Dependent variable.** The target is the 10-year cumulative excess return, defined as the cumulative stock return from year t to $t + 10$ minus the compounded 10-year risk-free rate. We add the risk-free rate back to the model's predictions to obtain cumulative return forecasts.

2. **Single model.** We fit one model covering the full 10-year horizon, rather than 20 separate models (two cash flow measures $\times$ ten horizons). Each cash flow model targets a one-year window $[t+h-1,\ t+h]$; the return model targets the full $[t,\ t+10]$ window.

3. **Learning rates.** Using the same learning rate as for the cash flow forecasts gave poor results. We therefore lowered the learning rate to $\eta_1 = 0.01$ (vs. 0.3) for the first hyperparameter search stage and $\eta_2 = 0.001$ (vs. 0.05) for the second. The hyperparameter search space is otherwise identical.

4. **Validation buffer.** Since the return target spans ten years, we add an extra 10-year buffer between the validation and test sets. Without this buffer, validation observations near the test boundary would have targets that overlap the test period, contaminating

the out-of-sample evaluation. The one-year cash flow targets do not require such a buffer.

5. **Data processing.** For non-US firms, returns come from Compustat and some illiquid stocks have clear data errors. We therefore winsorize returns at the 0.1%/99.9% level before training.

Appendix C.4 Vector autoregression (VAR)

We also consider VAR benchmark for both cash flows and returns. The VAR benchmark is inspired by Vuolteenaho (2002), but operates on levels rather than logs. The state vector is

$$z_t = \begin{bmatrix} r_t \\ \text{bm}_t \\ \text{roe}_t \end{bmatrix}, \tag{A10}$$

where r_t is the trailing 12-month excess return (cumulative 12-month return minus the correpsonding risk-free rate), bm_t is the book-to-market ratio, and roe_t is the return on equity.

The VAR(1) specification is

$$z_t = A + \Gamma z_{t-1} + u_t, \tag{A11}$$

estimated as three separate weighted least-squares regressions. Each observation is weighted by $1/N_t$, where N_t is the number of firms in month t, so that each cross-section receives equal weight (following the approach in Vuolteenaho, 2002). Before estimation, bm_t and roe_t are winsorized at the 1st and 99th percentiles each month. However, this approach gave poor and unstable results, so we further trimmed the *estimation* sample at the 2.5th and 97.5th percentiles. The model is re-estimated on an expanding window using the same December end dates as the ML models, ensuring comparability.

For each firm-month with observed z_t, we iterate the VAR forward to produce forecasts at horizons $h = 1, \ldots, 10$:

$$\hat{z}_{t+h} = A + \Gamma \hat{z}_{t+h-1}. \tag{A12}$$

The VAR produces two benchmark outputs:

Return forecasts. The VAR produces one-year-ahead excess return forecasts, but benchmarking requires expected 10-year cumulative returns. We approximate the 10-year cumulative return by summing the predicted annual excess returns and adding the compounded

10-year risk-free rate:

$$\hat{R}^{\text{VAR}}_{i,t} = \sum_{h=1}^{10} \hat{r}_{i,t+h} + \left[(1 + r_t^{f,10})^{10} - 1\right]. \tag{A13}$$

Cash flow forecasts. We convert the ROE forecasts into dollar earnings using a clean-surplus identity. Let b_0 denote book equity at the forecast origin and po the payout ratio at the origin, estimated as described in Appendix C for the analyst-based methods. For each horizon $h = 1, \ldots, 10$:

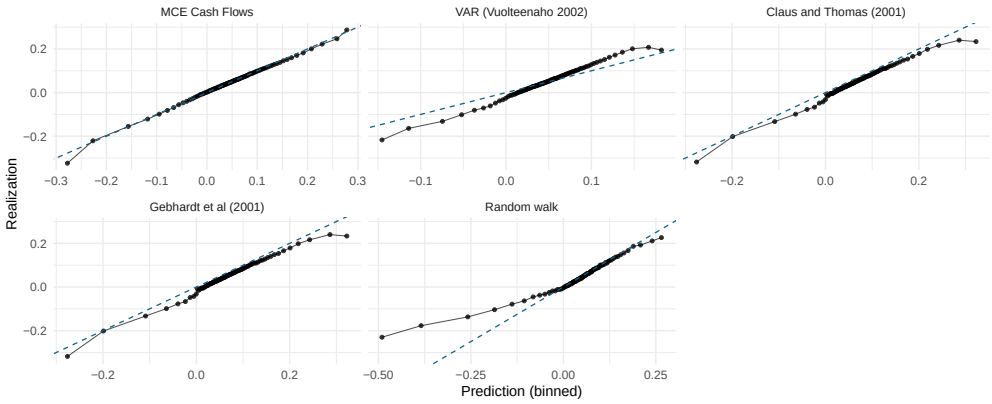
$$\hat{e}_h = \widehat{\text{roe}}_h \times b_{h-1}, \tag{A14}$$

$$b_h = b_{h-1} + \hat{e}_h \times (1 - \text{po}). \tag{A15}$$

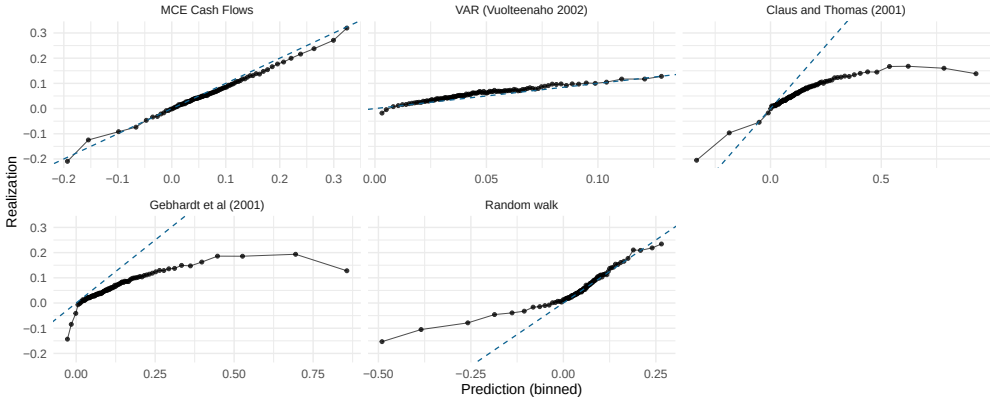
Appendix D Additional Figures and Tables

Figure A1
Cash flow predictability

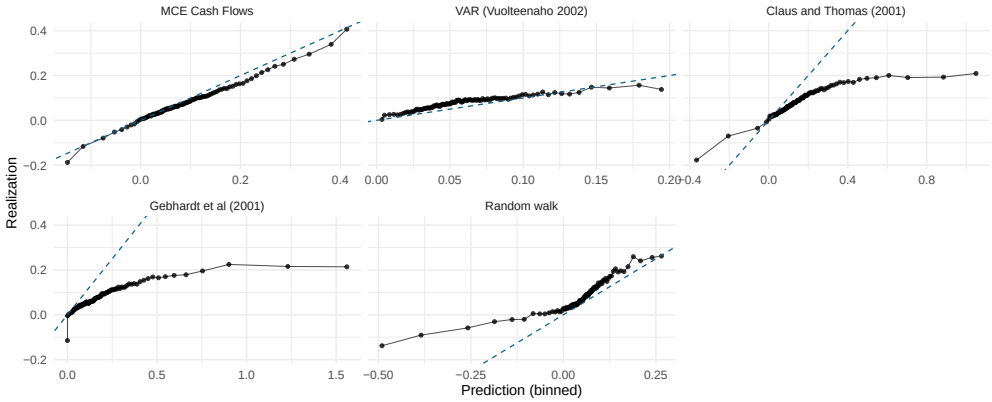
This figure shows binned scatter plots of realized earnings divided by assets at t on cash flow predictions at t, for forecast horizons $h = 1, 5, 10$ years.



Panel A. $h = 1$



Panel B. $h = 5$



Panel C. $h = 10$

Figure A2
Cash flow forecasts (unbalanced)

This figure shows the out-of-sample R^2 of the cash flow prediction models. We consider three different cash flow variables, as indicated by the panel header, and ten forecast horizons, as indicated by the x-axis. The sample is unbalanced, meaning that we use all data available and that the samples differ across variables and horizons. Figure 1 shows the results from the balanced sample.

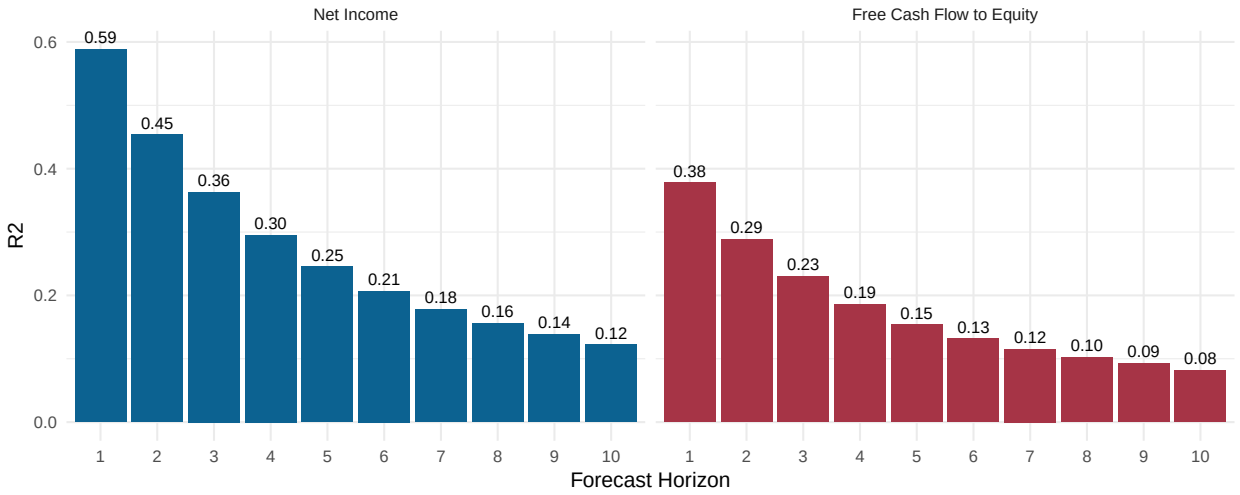


Figure A3
Cash flow forecasts over time (balanced)

This figure shows the out-of-sample R^2 of the cash flow prediction models every month. The panel header shows the cash flow variable, and the x-axis is the cash flow realization date. The sample is balanced. Figure A4 shows the results from the unbalanced sample.

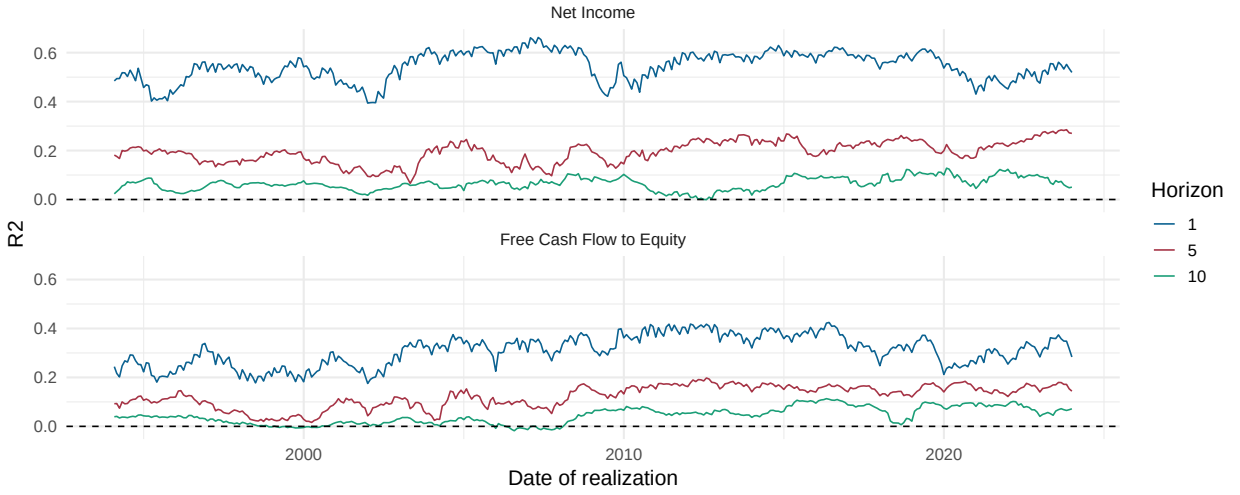


Figure A4
Cash flow forecasts over time (unbalanced)

This figure shows the out-of-sample R^2 of the cash flow prediction models every month. The panel header shows the cash flow variable, and the x-axis is the cash flow realization date. The sample is unbalanced. Figure A3 shows the results from the balanced sample.

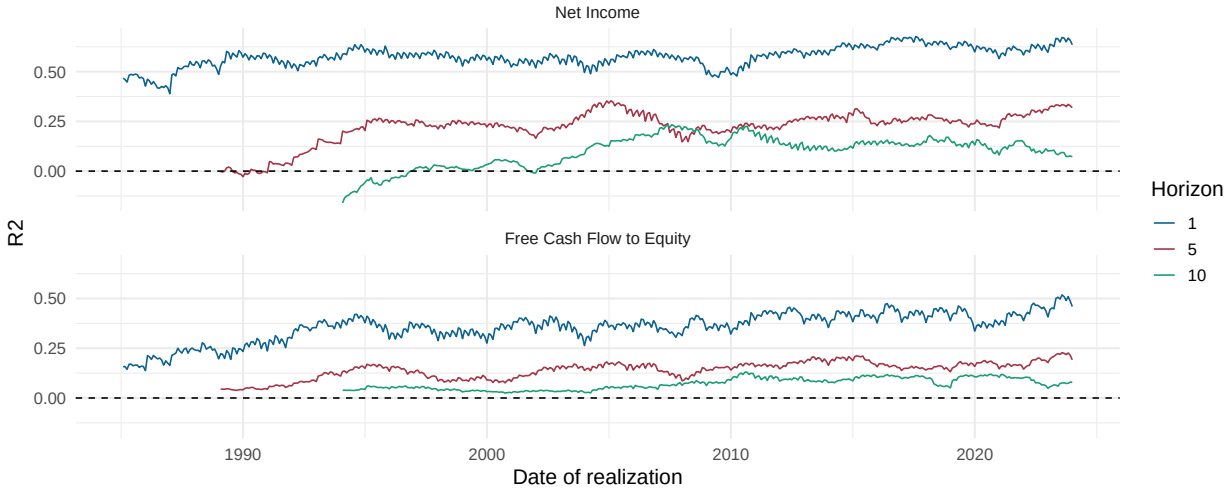


Figure A5
Long-run returns and the cost of equity

This figure shows the relationship between future 10-year realized returns and the current 10-year cost of equity. Specifically, each month we sort stocks into 100 portfolios by their cost of equity, and then compute the average 10-year cost of equity and the average realized buy-and-hold return over the next 10 years. We then average these two numbers for each portfolio over the full sample, 1983–2023. For ease of interpretation, the cost of equity and realized returns have been annualized. The solid red line shows the best linear fit, and the dotted blue line shows the 45-degree line. The sample is **U.S. stocks.**

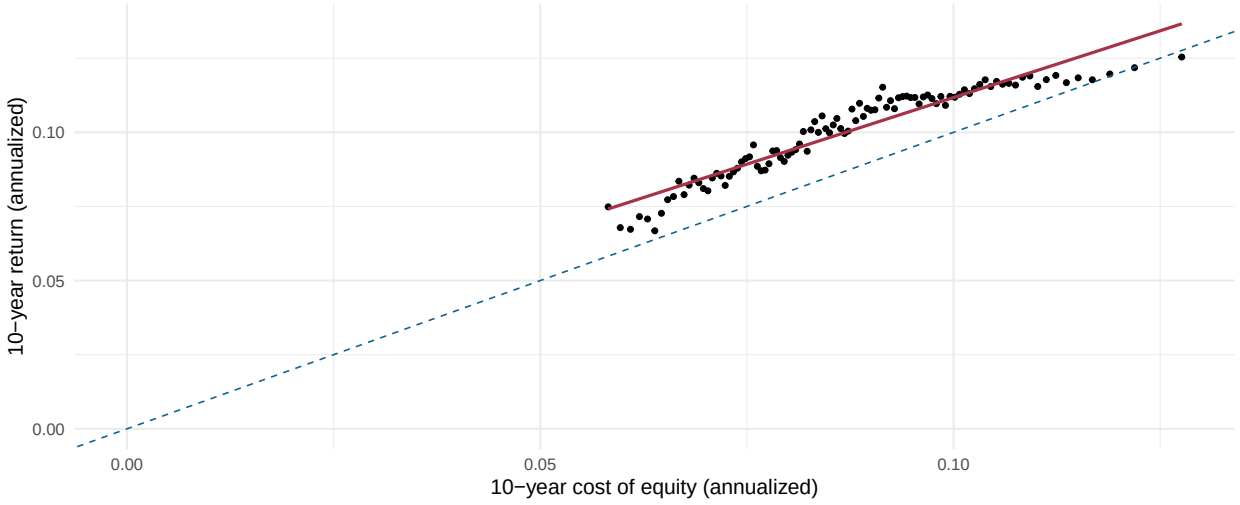


Figure A6
Long-run returns and the cost of equity

This figure shows the relationship between future 10-year realized returns and the current 10-year cost of equity. Specifically, each month we sort stocks into 100 portfolios by their cost of equity, and then compute the average 10-year cost of equity and the average realized buy-and-hold return over the next 10 years. We then average these two numbers for each portfolio over the full sample, 1983–2023. For ease of interpretation, the cost of equity and realized returns have been annualized. The solid red line shows the best linear fit, and the dotted blue line shows the 45-degree line. The sample consists of **non-U.S. stocks** listed in developed or emerging markets, as classified by MSCI.

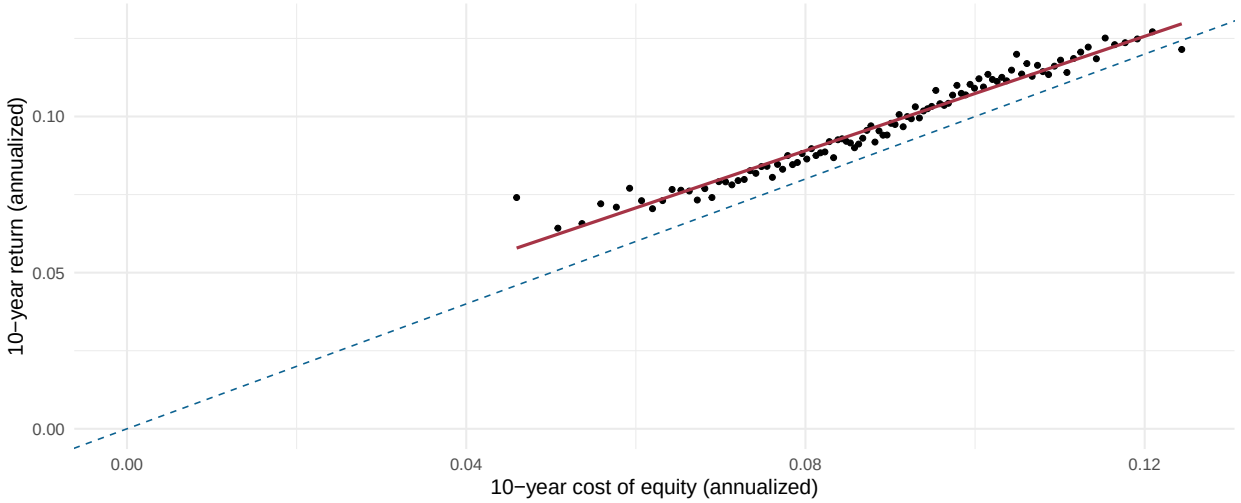


Figure A7
ML forecasts versus analysts

This figure compares the out-of-sample R^2 of one-year-ahead net income forecasts from the machine learning model, IBES analyst consensus forecasts, and a random walk (last observed net income). To ensure a comparable evaluation, the sample is restricted to analyst forecasts where the fiscal year end is exactly one year from the forecasting time, matching the constant-horizon property of the ML predictions. All forecasts and realizations are scaled by total assets and winsorized at the 1st and 99th percentiles. Firm-months where IBES and Compustat actual EPS disagree are excluded, and the sample requires non-missing predictions from all three methods. "Raw forecasts" computes R^2 as $1 - \text{MSE}/\text{Var}(y)$. "De-meaned forecasts" first subtracts the cross-sectional mean from both predictions and realizations before computing R^2, in an attempt to remove the well-known analyst level bias.

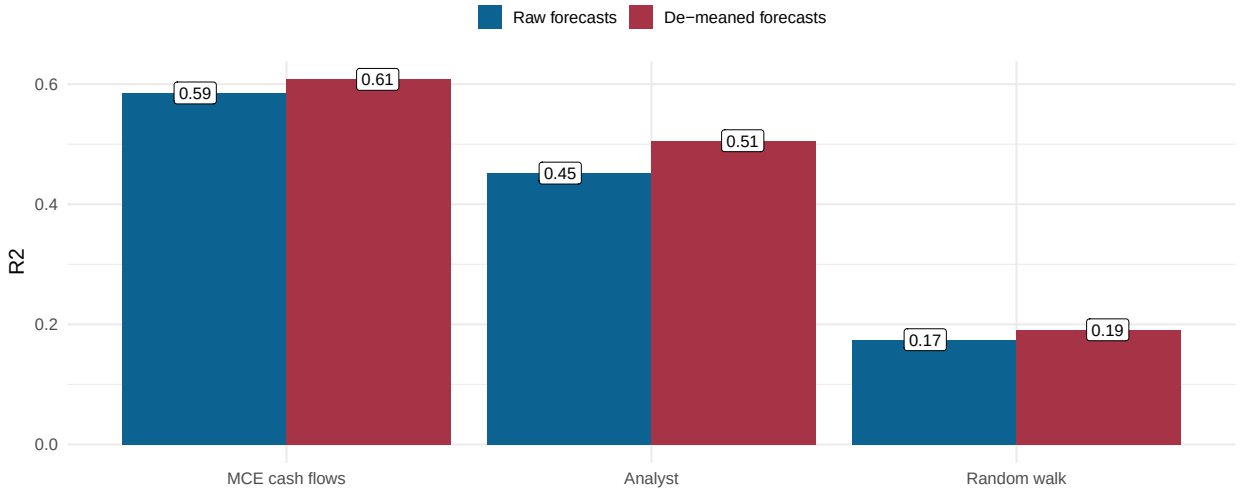


Figure A8
Optimal shrinkage analysis for return prediction

This figure shows the optimal ridge regression penalty parameter λ for a return prediction model that uses the same features as MLR but replaces LightGBM with ridge regression to isolate the role of shrinkage. For each training end date, we fit a ridge regression with λ selected from a grid of 1,000 values spaced evenly on a log scale from 10^0 to 10^6. The validation-optimal λ minimizes MSE on the validation set, while the test-optimal λ minimizes MSE on the out-of-sample test set. Features are cross-sectionally standardized (percentile-ranked, demeaned, and scaled to unit standard deviation) each month.

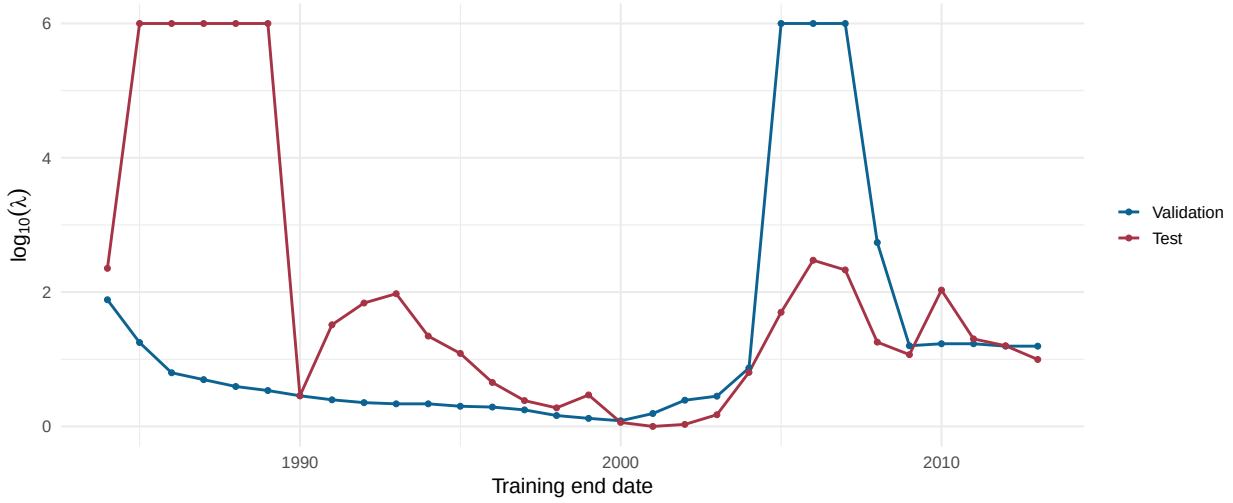


Figure A9
Model stability: MCE versus MLR

This figure compares the stability of MCE and MLR predictions across model training vintages. For each method, we generate predictions using every training vintage (models trained with data ending in 1984, 1985, . . . , 2023) applied to the same set of U.S. stocks. We then compute pairwise correlations between all vintage pairs, yielding one correlation per pair. The histograms show the distribution of these pairwise correlations, with the dashed line indicating the median. For MCE, the predictions are implied costs of equity derived from cash flow forecasts via discounted cash flow valuation. For MLR, the predictions are annualized 10-year return forecasts from LightGBM. Higher correlations indicate greater stability across training vintages.

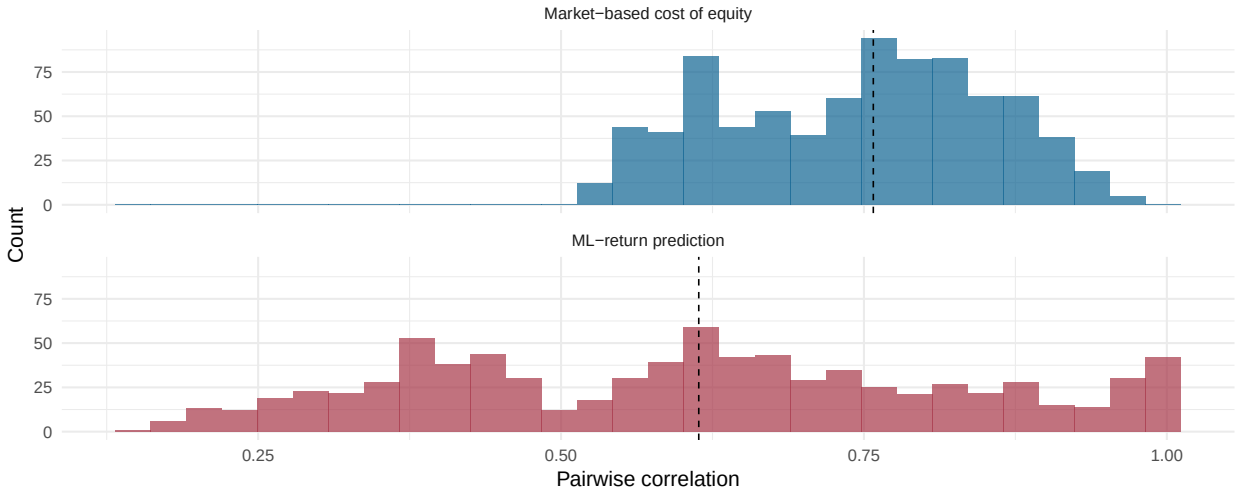


Figure A10
MCE Coverage

This figure shows the number of stocks with a non-missing MCE estimate, split by where the stock is listed. We only include firms that are from developed and emerging market countries, as classifed by MSCI.

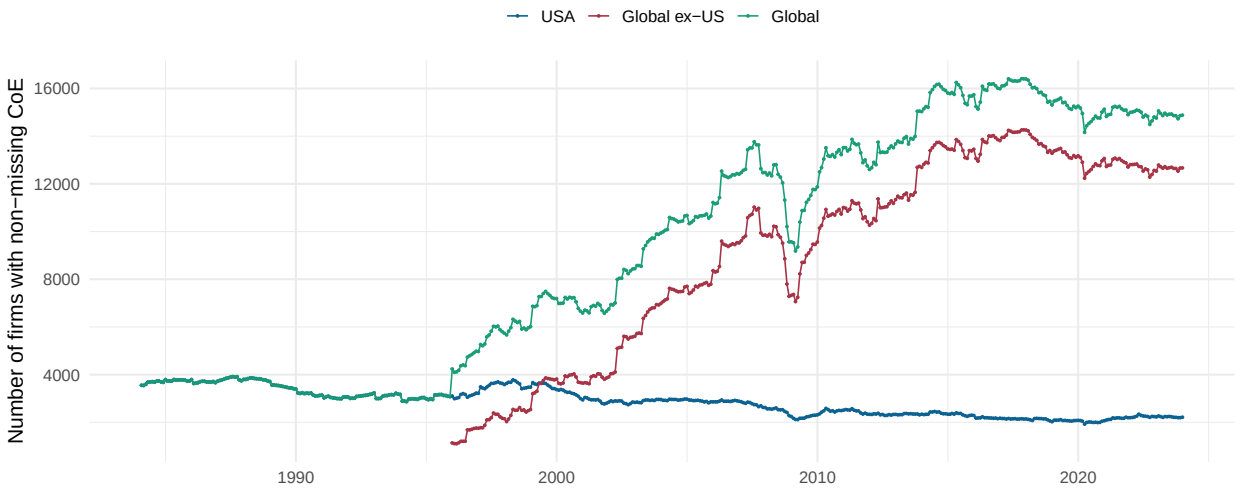


Figure A11
Time variation in the market's real MCE

This figure shows the value-weighted real market-based cost of equity (MCE).

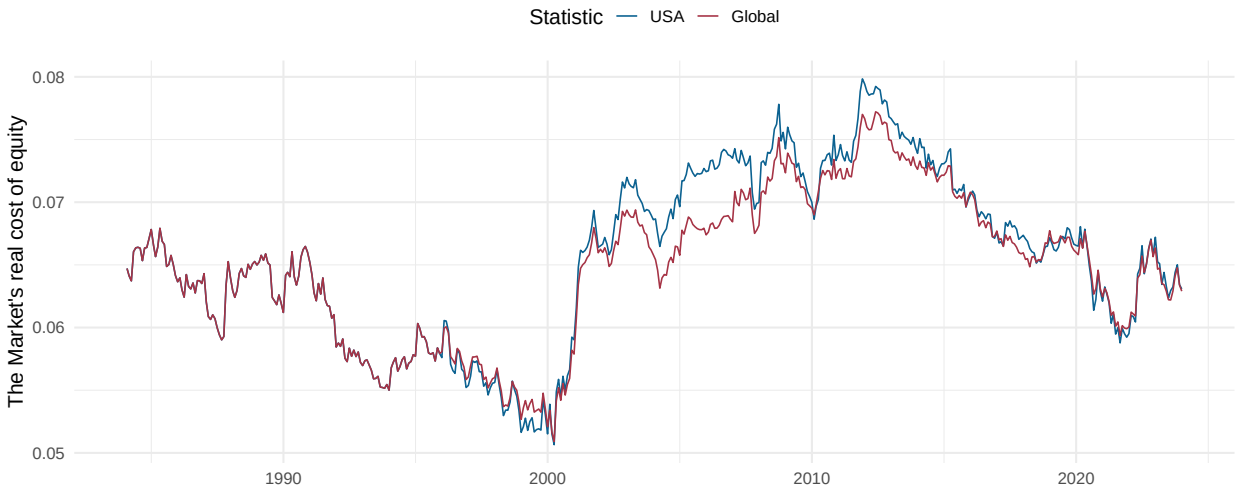


Figure A12
Time variation in the market's nominal MCE

This figure shows the value-weighted nominal market-based cost of equity (MCE).

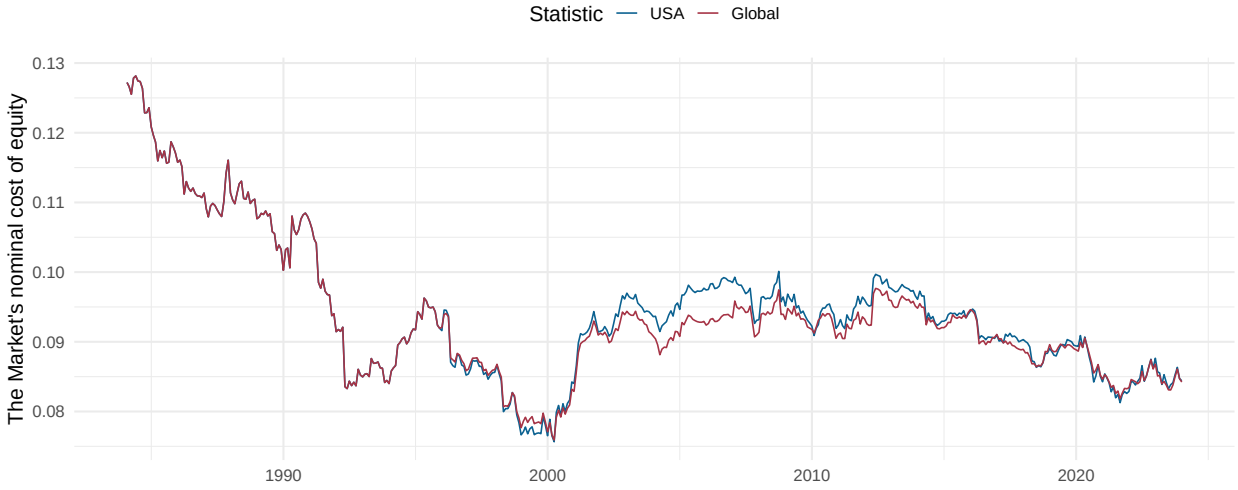


Figure A13
Time variation in the market's excess MCE

This figure shows the value-weighted nominal market-based cost of equity (MCE) minus the ten-year U.S. treasury yield.

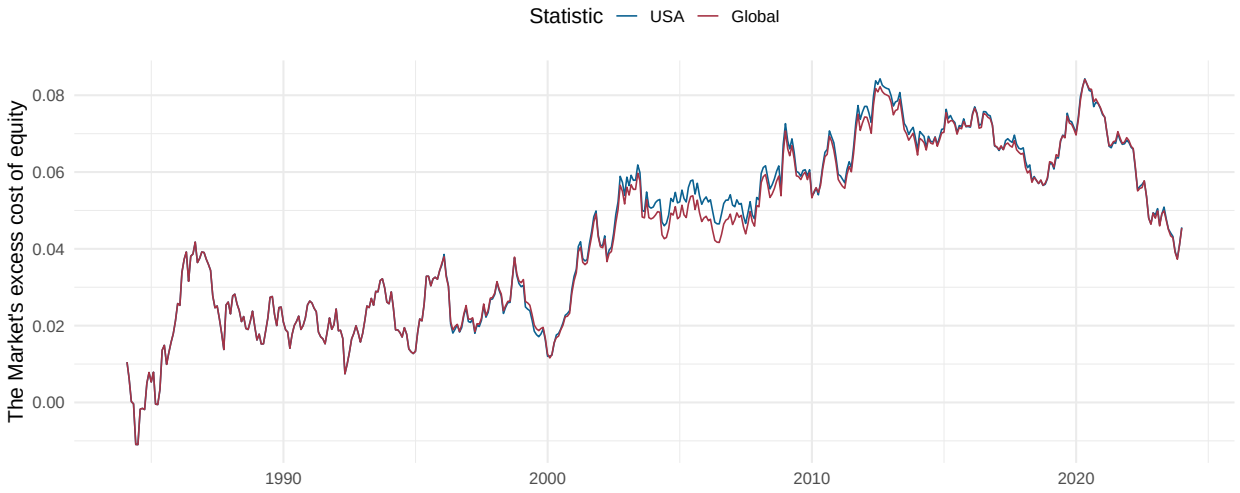


Figure A14
Time variation in the market's nominal MCD

This figure shows the value-weighted nominal market-based cost of debt (MCD).

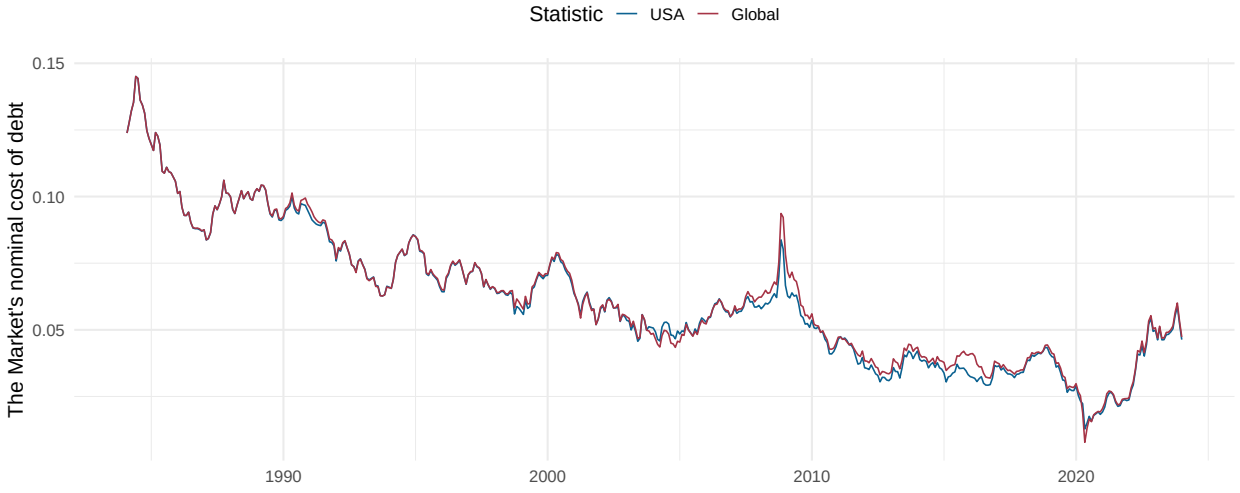


Figure A15
Time variation in the market's nominal MCC

This figure shows the value-weighted nominal market-based cost of capital (MCC).

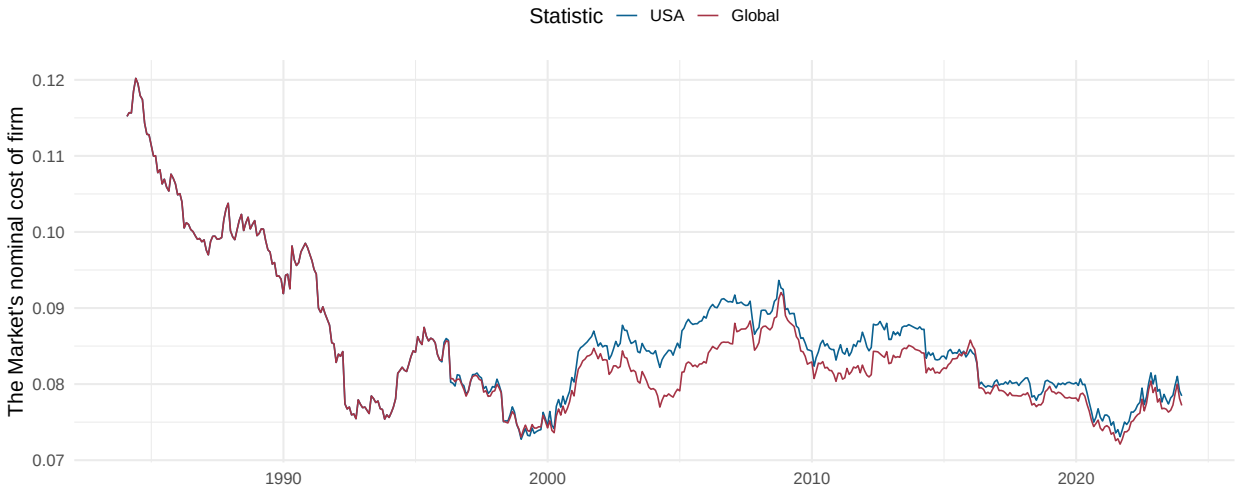


Figure A16
Buy-and-hold return predictability

This figure reports the slope coefficient b from regressions of future buy-and-hold stock returns on the MCE, $r^{\text{realized},i}_{t\to t+h} = \alpha_t + b \times \left[(1+\text{MCE}^i_t)^{h/12} - 1\right] + \varepsilon^i_{t+h}$, for return horizons $h = 1, 2, \ldots, 120$ months. All regressions include month fixed effects. The top row restricts the sample to U.S. firms; the bottom row uses the global sample. Dotted lines indicate 95% confidence intervals based on Driscoll-Kraay standard errors with a 120-month bandwidth. A coefficient of one (dashed line) implies that a one-percentage-point higher MCE maps one-for-one into higher realized returns.

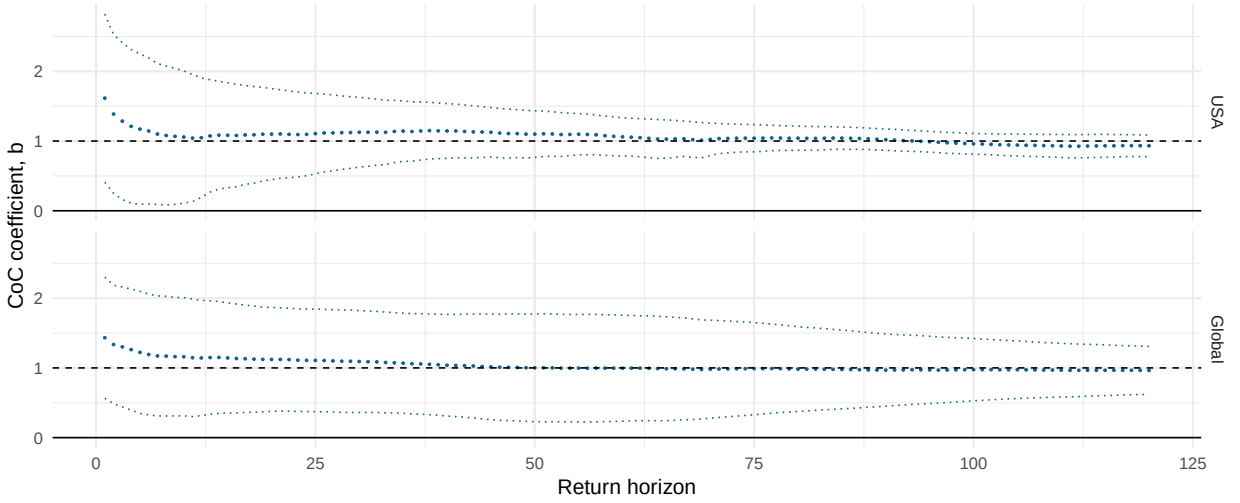
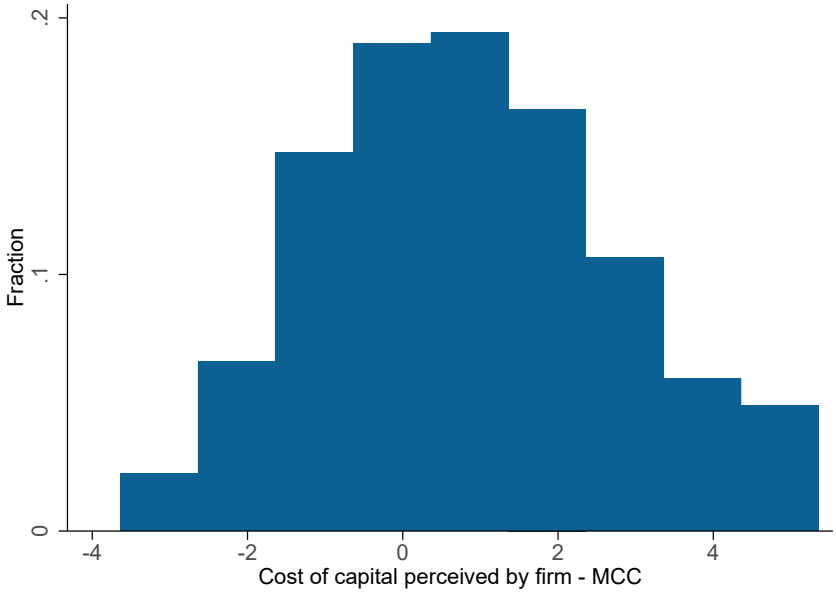


Figure A17
Histogram of deviations from the MCC

The figure shows a histogram of the cost of capital deviation (i.e., the difference between perc. CoC and MCC in percentage points). The perc. CoC is reported by the firm on conference calls and influences the firm's investment decisions. MCC is the market-based cost of capital from this paper. The rightmost bin combines observations above 5 percentage points.



	(1) Global	(2) Global	(3) U.S.
Global			
(Perc. CoC − MCC) > 0	0.199**	0.142*	0.170**
	(0.098)	(0.084)	(0.084)
Observations	709	709	709
Within R^2	0.0107	0.00680	0.00983

Column 1 replicates column 2 of Table 4. Column 2 controls for only quarter-by-year-by-country fixed effects, not book equity deciles-by-country. Column 3 uses the unwinsorized stock return as the outcome. Standard errors are clustered at the firm level. Statistical significance is denoted by *** p<0.01, ** p<0.05, * p<0.1.

Table A3: Robustness: stock returns and investment announcements