The Prospective Book-to-Market Ratio and Expected Stock Returns

Mengmeng Dong* Kewei Hou† Yan Xu‡ Yuzhao Zhang§

August 27, 2025

We propose a novel stock return predictor, the "*prospective* book-to-market", defined as the present value of expected demeaned book-to-market ratios. A high-minus-low investment strategy based on this ratio generates a significant alpha across various factor models. The return spread is also shown to be non-redundant as an alternative value factor in pricing the cross-section of stock returns. Compared to the conventional value measure, it contains more useful information about future return, but not about future profitability. We also show that the prospective book-to-market can predict industry and market returns.

*School of Business, University of California Riverside; 900 University Avenue, Riverside, CA 92521, USA; phone: +1 951 827 6329; fax: +1 951 827 2672; email: mdong@ucr.edu.

†Fisher College of Business, The Ohio State University, 2100 Neil Avenue, Columbus, OH 43210, USA; phone: +1 614 292 0552; fax: +1 614 292 7062; email: hou.28@osu.edu.

‡Faculty of Business and Economics, University of Hong Kong; Pokfulam Road, Hong Kong; phone: +852 2859 7037; fax: +852 2548 1152; email: yanxuj@hku.hk.

§Rutgers Business School, Rutgers, The State University of New Jersey, Newark, NJ 07102, USA; Phone: +1 973 353 2727; fax: +1 973 353 1006; email: yzhang@business.rutgers.edu

Prospective book-to-market ratio and expected stock returns

We propose a novel stock return predictor, the "*prospective* book-to-market", defined as the present value of expected demeaned book-to-market ratios. A high-minus-low investment strategy based on this ratio generates a significant alpha across various factor models. The return spread is also shown to be non-redundant as an alternative value factor in pricing the cross-section of stock returns. Compared to the conventional value measure, it contains more useful information about future return, but not about future profitability. We also show that the prospective book-to-market can predict industry and market returns.

JEL Classification: G12, F31

Keywords: prospective book-to-market ratio, return predictability, value premium

1 Introduction

This paper proposes a new stock return predictor by decomposing the book-to-market ratio into permanent and transitory components. This decomposition relates the present value of demeaned stock returns to the transitory component of the book-to-market ratio, the present value of the demeaned book-to-market ratio, and the present value of the demeaned return on equity. When the expected return moves with any of these three terms, the future stock return can be predicted when the investor observes new information about the book-to-market ratio. Specifically, we focus on the "*prospective* book-to-market",[1] defined as the expected sum of all future book-to-market ratios around their long-run trend. When this expected sum exceeds the long-run trend, it indicates that the expected return is above its long-run trend or that the market value is temporarily underpriced relative to the book value, with an expectation to rise. Indeed, our findings suggest that the prospective book-to-market ratio is particularly useful in predicting returns, both in the short and long term.

Empirically, we model the prospective book-to-market by assuming a simple autoregressive form of the book-to-market ratio and then estimate its infinite sum while accounting for the historical average. Similar to van Binsbergen and Koijen (2010) who utilize the variation of persistence in state variables to more precisely predict stock returns, the superior predictive power of the prospective book-to-market in our setup depends on both the persistence and discount coefficients of the book-to-market ratio, and its current level relative to the long-run trend.

We present the predictability tests on three portfolio levels: the cross-section of individual firms, industry, and market. In out-of-sample tests, we use only currently available information to avoid look-ahead bias. For all three portfolio levels, we use the historical sample average of the book-to-market ratio as a proxy for the long-run trend. However, the estimation of its persistence and discount coefficients deserves further elaboration. For both market and industry portfolios, we rely on their own time series. For the cross-section of individual stocks, we assign the values of parameters from their respective industries to mitigate noise from a limited time series.

[1] The naming of this term is analogous to Engel (2016), although our setting is equity market instead of currency market.

We find that our prospective book-to-market ratio is a significant predictor of returns at all levels. We observe substantial cross-industry variations in the model parameters, which create additional degrees of heterogeneity and lead to a highly profitable investment strategy as we sort firms into portfolios. To demonstrate the forecasting power of our new predictor, we take long positions in firms whose expected sum of future book-to-market ratios is higher than their historical average and short positions when it is lower after controlling for firm size. This strategy generates significant monthly alphas from various empirical factor models, ranging from 3-, 4-, or 5-factor Fama-French models, q-factor model, or its augmented version. Further adding the momentum factor or the size component of Gerakos and Linnainmaa (2018) still does not weaken our results.

We then conduct time-series spanning tests by regressing the returns of alternative HML factors on our prospective book-to-market factor. HML factor of Fama and French (1992) uses the end-of-December market value, while those of Asness and Frazzini (2013) use the end-of-June or most recent month market value. These versions of HML factors show an insignificant alpha when regressed on our prospective book-to-market factor, but not vice versa. We also contribute to the debate on whether HML is a redundant factor in existing factor models.

Our evidence suggests that prospective book-to-market is useful in developing a better value strategy. The component of the monthly updated book-to-market of Asness and Frazzini (2013) that is orthogonal to our prospective book-to-market can not predict returns anymore in value-weighted portfolio sorts. Compared to the traditional version of book-to-market measured at December, Cohen, Polk, and Vuolteenaho (2003) show that its cross-sectional variation largely reflects differences in future book-to-markets and profitability, instead of return. In contrast, while we find prospective book-to-market is a stronger return predictor, it has little information about future profitability. Thus, our approach could lead to a more precise estimation of the value effect. Finally, prospective book-to-market still predicts returns in 5 years, while variations of conventional book-to-market, i.e., measured in December, in June, or monthly updated, all fail to do so.

In industry-level time-series tests, we demonstrate that the prospective book-to-market ratio predicts returns for the Fama-French 48-industry portfolios. Employing a zero-cost long-short strategy, the prospective book-to-market ratio of the industry is shown to generate a significant annual spread of more than 3.5% and approximately 2% risk-adjusted returns across industries. In contrast, the original book-to-market ratio does not produce similar results, which is consistent with the findings of Asness, Porter, and Stevens (2000).

At the market level, we show that one version of the prospective book-to-market ratio yields an out-of-sample adjusted R-squared of 4.6%. This contrasts with the conclusion of Goyal and Welch (2008) that market returns can not be reliably predicted out-of-sample. Moreover, as shown in Campbell and Thompson (2008), these out-of-sample R-squared values imply substantial economic gains for investors.

Our study makes a significant contribution to the existing literature in at least two ways. Methodologically, we extend the model presented in Engel (2016) by decomposing the book-to-market ratio into transitory and permanent components. Our model derives the transitory component of the book-to-market ratio, implying the return predictability of a multi-period sum of expected book-to-market ratios, which we term the "prospective book-to-market ratio". This innovation is similar to the "prospective interest rate differentials" found in Engel (2016) and Dong, Goto, Hou, Xu, and Zhang (2024), which has been shown to predict currency excess returns.

Our work also contributes to the literature on the value premium. Recent studies have focused on decomposing the book-to-market ratio and examining the return predictive power of each component. For example, Asness, Porter, and Stevens (2000) demonstrate that most of the return predictability of book-to-market comes from the within-industry components. Daniel and Titman (2006), and Fama and French (2008) also study different components of the book-to-market ratio. Gerakos and Linnainmaa (2018) find that a factor built from the variation in the book-to-market that relates to size changes captures the entire value premium. In contrast, we examine the transitory component of the book-to-market ratio as a starting point to enhance the predictive power of the raw book-to-market value.

The paper is organized as follows: Section 2 details our present-value model and presents the decomposition into permanent and transitory components. In section 3, we introduce the data, report the estimation of model parameters, present predictive regressions of expected returns, and compare out-of-sample portfolio performance. Section 4 concludes the paper.

2 Model

We start with the definition of stock return

$$\frac{P_{t+1}}{P_t}\left(1+\frac{D_{t+1}}{P_{t+1}}\right)=R_{t+1}, \tag{1}$$

where P_t, D_t, and R_t denote the stock price, dividend, and returns, respectively. We use the lower-case letters to denote the logarithm of these variables. Let δ_t be the log dividend-price ratio $\delta_t = \log\left(1 + D_t/P_t\right)$, and we take the log on both sides:

$$p_{t+1} - p_t + \delta_{t+1} = r_{t+1}. \tag{2}$$

Taking expectations at time t, iterating forward, and summing up, we have

$$E_t p_{t+j} - p_t + \sum_{j=1}^{k} E_t \delta_{t+j} = \sum_{j=1}^{k} E_t r_{t+j}. \tag{3}$$

Define the long-run trends of r and δ as $\bar{\mu}$, and $\bar{\delta}$, respectively, let $\tau = \bar{\mu} - \bar{\delta}$ we have

$$\lim_{j\to\infty}\left(E_t p_{t+j} - p_t - j\tau\right) + \sum_{j=1}^{\infty} E_t\left(\delta_{t+j} - \bar{\delta}\right) = \sum_{j=1}^{\infty} E_t\left(\mu_{t+j} - \bar{\mu}\right). \tag{4}$$

The modeling methodology follows Engel (2016), which focuses on the sum of deviations of expected future interest rates from their long-run trend. Dong, Goto, Hou, Xu, and Zhang (2024) develop an empirical proxy for this "prospective interest rate differential" and demonstrate that it predicts currency returns beyond the conventional carry trade. According to the Beveridge and

Nelson (1981) decomposition, $\lim_{j\to\infty} E_t p_{t+j} - j\tau$ can be viewed as the permanent component of the stock price p_t^P. Once we eliminate the permanent component, both sides of the equation become stationary.

The above equation is reminiscent of the well-known Campbell and Shiller's approximate identity, which relates log dividend-to-price to the present value of returns and cash flows. It also appears similar to Vuolteenaho's approximate identity, which relates book-to-market to the present value of returns and cash flows, as discussed in Vuolteenaho (2002) and Cohen, Gompers, and Vuolteenaho (2002). However, there are several important conceptual differences. Our decomposition focuses on an unobservable term, namely, the transitory component of the book-to-market ratio, and our equation holds as an identity rather than an approximation. Additionally, the return decomposition by Cohen, Gompers, and Vuolteenaho (2002) and Campbell (1991) is related to the Beveridge and Nelson (1981) decomposition in the time-series literature. This is motivated by the intuition that news about cash flow and expected returns each correspond to shocks to the random-walk and stationary components of the log stock price. In contrast, our decomposition is motivated by the present-value relationship.

We can rewrite the permanent-transitory component decomposition as

$$p_t^P - p_t + \sum_{j=1}^{\infty} E_t\left(\delta_{t+j} - \bar{\delta}\right) = \sum_{j=1}^{\infty} E_t\left(\mu_{t+j} - \bar{\mu}\right). \tag{5}$$

Next, we apply the same approach to the log book equity $b_t = \log(B_t)$. Define the log dividend-book equity as $\psi_t = \log(1 + D_t/B_t)$, then we also have

$$b_t^P - b_t + \sum_{j=1}^{\infty} E_t\left(\psi_{t+j} - \bar{\psi}\right) = \sum_{j=1}^{\infty} E_t\left(g_{t+j} - \bar{g}\right). \tag{6}$$

where g_t is return on equity. Now define the log book-to-market ratio as $\theta_t \equiv \log(B_t/P_t) = b_t - p_t$. By subtracting equation (6) from equation (5), we obtain

$$\sum_{j=1}^{\infty} E_t\left(\left(\delta_{t+j} - \bar{\delta}\right) - \left(\psi_{t+j} - \bar{\psi}\right)\right) = \sum_{j=1}^{\infty} E_t\left(\mu_{t+j} - \bar{\mu}\right) - \sum_{j=1}^{\infty} E_t\left(g_{t+j} - \bar{g}\right) - \left(\theta_t - \theta_t^P\right). \tag{7}$$

This equation decomposes the transitory component of the book-to-market ratio into the infinite sum of three terms: the expected demeaned future return, the expected demeaned future return on equity, and a function of expected demeaned log book-to-market ratio (as shown below). Therefore, if the book-to-market ratio is temporarily high, it may indicate that investors expect a persistently above-average future discount rate (i.e., the first term), or persistently below-average future cash flows (i.e., the second term). If there is no time variation in these first two terms, then the current change in the book-to-market ratio will purely reflect the time variation in future book-to-market values (i.e., the third term).

Next, we propose an empirical proxy for $\sum_{j=1}^{\infty} E_t\left(\left(\delta_{t+j}-\bar{\delta}\right)-\left(\psi_{t+j}-\bar{\psi}\right)\right)$. First, we assume the same AR(1) coefficient β such that $\delta_t-\bar{\delta}=\beta\left(\delta_{t-1}-\bar{\delta}\right)+\varepsilon_{\delta,t}$ and $\psi_{t-1}-\bar{\psi}=\beta\left(\psi_{t-1}-\bar{\psi}\right)+\varepsilon_{\psi,t}$. Secondly, we conduct the log-linearization (Campbell and Shiller (1988)) and assume that, as in Vuolteenaho (2002) (omitting a constant):

$$\log\left(1+\frac{B_t}{D_t}\right)-\log\left(1+\frac{P_t}{D_t}\right)\approx\rho\theta_t \tag{8}$$

Where ρ is the discount coefficient of book-to-market. Then

$$\delta_t-\psi_t=\log\left(1+\frac{D_t}{P_t}\right)-\log\left(1+\frac{D_t}{B_t}\right)=\log\left(1+\frac{P_t}{D_t}\right)-\log\left(1+\frac{B_t}{D_t}\right)-p_t+b_t\approx(1-\rho)\theta_t. \tag{9}$$

Then we can rewrite the exact decomposition to an approximation

$$(1-\rho)\frac{\beta\left(\theta_t-\bar{\theta}\right)}{1-\beta}\approx\sum_{j=1}^{\infty}E_t\left(\mu_{t+j}-\bar{\mu}\right)-\sum_{j=1}^{\infty}E_t\left(g_{t+j}-\bar{g}\right)-\left(\theta_t-\theta_t^P\right). \tag{10}$$

This decomposition dictates that $(1-\rho)\frac{\beta\left(\theta_t-\bar{\theta}\right)}{1-\beta}$ should correlate with at least one term of the following: current transitory component of book-to-market, expected return, and expected roe. We label $(1-\rho)\frac{\beta\left(\theta_t-\bar{\theta}\right)}{1-\beta}$ as the "prospective book-to-market," and the main objective of our paper is to test if it strongly correlates with future return. Therefore in the empirical analysis, we do not focus on using complicated time-series model to estimate the transitory component of the book-to-market ratio, $\theta_t - \theta_t^P$.[2]

3 Data and Empirical results

3.1 Cross-section of returns

Our data are from the CRSP-Compustat, comprising all common shares (share codes 10 or 11) traded on the NYSE, Amex, and Nasdaq from 1950 to 2023. Every fiscal year, we calculate book equity as the sum of shareholder equity, balance sheet deferred taxes, and balance sheet investment tax credits, less the value of preferred stock. We set missing deferred taxes and investment tax credits to zero. Preferred stock is valued at its redemption, liquidatio, respectivelying value, in this sequence of availability. Fama-French and devil factors are from Ken French's and AQR's websites separately.

3.1.1 Parameter estimation

Due to the limited number of observations, running regressions firm-by-firm would likely result in noisy estimates. To mitigate potential estimation errors, we estimate the parameters of the Fama-French 48 industries and then assign them to the individual firms. Implicitly, we assume that all firms within the same industry share the same AR(1) coefficient (β), discount coefficient (ρ), and long-run trend ($\bar{\theta}$). Alternatively, groups can be portfolios sorted by other predictive variables; however, their potential correlation with book-to-market might contaminate the construction of prospective book-to-market.

Our OOS period begins from July 1961 to December 2023. We use log book-to-market (θ_t), log annual return from July of the current year to June of the next year (r_t), and log return on

[2] In a companion paper, we propose a slightly different empirical approach that better accommodates the much weaker persistence of return on equity, demonstrating that this framework also allows the prospective ROE to predict returns. We introduce a much larger array of accounting variables to enhance their return predictive power, applicable both to the U.S. equity market and international markets, all within the same framework.

equity (roe_t) from 1951 to 1960 to generate the first estimate of β and ρ, following Vuolteenaho (2002):

$$\theta_t = \alpha_1 + \beta\theta_{t-1} + \varepsilon_{1,t} \tag{11}$$

$$roe_t - r_t + \theta_{t-1} = \alpha_2 + \rho\theta_t + \varepsilon_{2,t} \tag{12}$$

We estimate these two regressions jointly with a seemingly unrelated regression (SUR) approach as clearly, $\varepsilon_{1,t}$ and $\varepsilon_{2,t}$ are strongly correlated. We also estimate an industry value-weighted average as the long-run trend $\bar{\theta}$. For each subsequent year, we expand our estimation window and repeat these procedures. We then assign these estimated parameters to each firm within that industry to compute the prospective book-to-market ratio (π).[3]

Table 1 about here.

Table 1 reports the sample summary statistics for monthly firm-level stock returns, annual book-to-market ratios (bm), and annual prospective book-to-market ratios (π). The statistics are estimated for every cross-section and averaged over time. Compared with bm, π is slightly more volatile (standard deviation of 2.24 vs. 1.77), much less left-skewed (skewness of -0.88 vs. -9.93), and distributes with a much flatter peak (Kurtosis of 32.98 vs. 724.49). Panel B reports time-averaged pairwise correlations between returns, bm, and π. While return weakly positively correlates with both bm and π, the correlation between the latter two is 0.548.

3.1.2 Predicting return

We examine whether the return predictability of the prospective book-to-market, defined as $\pi = (1-\rho)\frac{\beta(\theta_t - \bar{\theta})}{1-\beta}$, can translate into profitable portfolio performances. The conventional book-to-market ratio is well-known for producing the value anomaly across stock returns. However, Fama

[3] We have also employed various regression methods, including firm-specific fixed-effect regression, two-way fixed-effect regression, and random-effect regression (GLS), all of which yield similar results. Using an equal-weighted industry book-to-market ratio as the long-run trend produces comparable outcomes.

and French (2015) find that their HML factor did not produce a significant alpha when regressed against the other four factors, suggesting its redundancy in explaining cross-sectional returns. Asness and Frazzini (2013) demonstrate that the HML factor, when constructed using either the June-end market value or the most recent month market value, could generate significant alphas exceeding those of the original HML. Therefore, our objective in this section is threefold. First, we explore whether π is a useful return predictor. Second, if so, we test whether the return spread of π portfolios, similarly considered as one more version of the value factor, would be explained away by existing factors. Third, we ask if π return spread is more effective in pricing the cross-section of stock returns, compared to the various versions of HML factors.

Following the methodology of Fama and French (1992) and Fama and French (1993), we construct the prospective factor using six value-weighted portfolios formed based on size and the prospective book-to-market ratio. At the end of June each year, stocks are independently sorted into two size-based portfolios determined by the NYSE market-cap median and into three portfolios based on the 30% and 70% NYSE breakpoints for prospective book-to-market ratios. We value-weight these portfolios and update the breakpoints the next year. The return spread of portfolios sorted on π, which we refer to as prospective factor, is derived from the high-minus-low return, calculated as the difference between the average returns of the top two portfolios within the highest 30% of prospective book-to-market ratios and the bottom two portfolios within the lowest 30%.

Table 3 about here.

Panel A of Table 3 provides the summary statistics for the prospective book-to-market ratio factor, including the mean, standard deviation, maximum, minimum, and Sharpe ratio. Over our sample period, the factor achieved a mean return of 34.3 basis points per month, corresponding to an annualized Sharpe ratio of 0.513.

Panel B of Table 3 presents the time-series regression results of the prospective book-to-market factor against various factor asset pricing models. Following the literature, we consider the following models: the q-factor model by Hou, Xue, and Zhang (2015), which includes MKT, ME, IA, and ROE; and the Fama-French models, which include the three-factor (MKT, SMB, HML) model from Fama and French (1993) and the 5-factor (adding RMW and CMA) model from Fama and French (2015), supplemented by a momentum factor (MOM). The t-statistics, adjusted for 6-lag Newey-West robust standard errors, are reported in parentheses. While the q-factors are available from January 1967, all other factors commence from July 1963.

We test whether existing risk factors encompass the proposed prospective book-to-market ratio factor. Our findings show that the α for every regression significantly differs from zero, suggesting that these factors do not span it. Among all models tested, the q 4-factor model yields the largest alpha at 24.3 basis points per month ($t = 2.830$), while the Fama-French three-factor model leaves the smallest alpha with 14.7 basis points per month ($t = 2.451$). Overall, we observe that the prospective factor is weakly negatively correlated with the market (MKT) and the momentum (MOM) factor, negatively correlated with the profitability factors (ROE and RMW), and strongly positively correlated with the size (ME) and investment factors (IA and CMA). The prospective factor also positively correlates with the original HML factor and the size component of the HML factor (HML_s) as in Gerakos and Linnainmaa (2018). Nevertheless, models featuring these two factors cannot fully explain the prospective factor returns.

We next assess whether our proposed prospective factor can encompass various versions of the HML factor. We analyze three HML factors as the dependent variable: the standard version following Fama and French (1992), constructed using the December market cap (HML), and two alternative versions from Asness and Frazzini (2013)-one using the June market cap (HML_6) and the other using the monthly updated market cap (HML_d). Unlike annually constructed HML and HML_6, HML_d is estimated monthly and contains more frequently updated information. Given that HML factors are not part of the q factors, our examination focuses solely on the Fama-French three- and 5-factor models, both with and without the momentum factor. We replace the standard HML factor in the Fama-French factor models with our prospective factor (HML^π) in each regression.

Table 4 about here.

Table 4 Panel A presents the results when the standard HML factor is regressed on our prospective factor HML^{π}. All models demonstrate that the standard HML is spanned by HML^{π}, evidenced by either an insignificant or a significantly negative alpha, ranging from -0.169% to 0.022% (t-stat from -2.131 to 0.256). Notably, four out of five alphas are negative. Starting with a simple regression, HML is approximately 1.003 times HML^{π}. Adding more factors reveals that SMB barely accounts for the residual returns of the HML_6, while MKT only significantly and negatively explains the residual return in the three-factor model.

Panel B presents the regression results with the HML_6 factor as the dependent variable, quite similar to Panel A. All models show either a close-to-zero or significantly negative alpha, ranging from -0.193 to 0.129 (t-stat from -2.156 to 1.489). Again, four out of five alphas are negative, except when the MOM factor is added to the Fama-French three-factor model. Regarding alpha magnitudes, the results are consistent with Asness and Frazzini (2013) that HML is somewhat inferior to HML_6.

Panel C reports the results for HML_d. Consistent with Asness and Frazzini (2013), HML_d is much more negatively correlated with MOM factor, as shown by the large negative coefficients of MOM (-0.440 and -0.460). These magnitudes are much larger than those in Panels A (-0.031 and -0.053) and B (-0.178 and -0.203). As a result, HML_d displays a significantly positive alpha in the two specifications when MOM is added. Still, alpahs from other three models without MOM are all insignificantly negative, suggesting that HML_{π} can fully span HML_d in the absence of momentum.

In sum, these regression results indicate that our prospective factor cannot be explained by a variety of risk factor models and is the least redundant among all three annually-refreshed HML factors. As time-series regression only provides in-sample evidence that spanned factor cannot contribute more to efficient frontier, in the next section we turn to out-of-sample analysis.

3.1.3 Out-of-sample tangency portfolio weights

Asness, Moskowitz, and Pedersen (2013) claim that value and momentum strategies are highly negatively correlated and should be studied together. They show that a combination of the two portfolios significantly outperforms each separately. Indeed, we find that the correlation of the decile 10 minus the decile 1 return spread between portfolios sorted on momentum (MOM) and HML_d is -0.80 while that between momentum and HML^π is -0.44 in our sample period of 1963:07$\sim$2023:12.[4]

We compare the combination of momentum and various value strategies out-of-sample. Specifically, we examine the Sharpe ratio and weights of these return spreads contributed to the tangency portfolios based on conventional factors. Every period, we use 360 months prior returns to estimate the mean and variance-covariance matrix of the factors to construct an out-of-sample tangency portfolio. As the number of factors is fairly small (from 2 to 8) relative to the length of time series (360 months), these out-of-sample tangency portfolio weights can be estimated rather precisely. We then report the Sharpe ratio and the time-series average of the portfolio weights in Table 5.

Table 5 about here.

Panels A to D examine if each variation of value factor (HML, HML_6, HML_d, or HML^π) can further contribute to an out-of-sample tangency portfolio constructed using Fama-French factors. With only the three classic factors (market, size, and value), all value factors occupy large weights, with HML^π being the largest, 70.7%; HML^π also generates the largest Sharpe ratio of 0.606. After including MOM, this pattern is not affected, as HML^π's weight (53.8%) still far exceeds that of the other three: 38.7% for HML, 36.5% for HML_6, and 36.6% for HML_d. Meanwhile, MOM's weight (26.3%) is the smallest in the presence of HML^π compared to others: 28.4% for HML, 30.9% for HML_6, and 30.8% for HML_d. In Fama-French 5-factor models, all HMLs except HML^π have negative, close to zero weight. This suggests a highly positive correlation between these other value factors and the two newly developed factors RMW and CMA. However, in contrast, HML^π still holds an 11.5% weight. Finally, when we further include MOM, again, HML^π's weight of 16.5% is the largest, also with the largest Sharpe ratio (0.914).

We next focus on the q-factor models in Table 5 from Panel E to H. Panel E serves as a benchmark model, whereas Panels F to H each add a new factor: HML_6, HML_d, or HML^π.

[4] Decile portfolio returns sorted on momentum are obtained via Ken French's website.

Due to the stronger explanatory power of q-factor models on momentum and value, these HML factors generally do not contribute much to the tangency portfolio. Still, combined with the q4-factors (without EG), HML^{π} occupies an 18.4% weight, compared to that of 1.3% for HML_6 and 9.8% for HML_d. The latter two also shrink the efficiency frontier, resulting in smaller Sharpe ratios, possibly because they introduce additional estimation uncertainty, whereas HML^{π} slightly increases the Sharpe ratio from 0.857 in panel E to 0.871 in Panel H. In particular, HML^{π} reduces the weight of ME from 12.5% to 9.3% and that of IA from 39% to 22.9%. These findings indicate that incorporating HML^{π} still yields diversification gains in real-time. Finally, the pattern remains similar when we consider the augmented q-model with EG factor: adding HMLs, in fact, deteriorates the performance of the tangency portfolio, as the Sharpe ratio gradually declines from the benchmark in Panel E. This pattern is more obvious for HML_6 and HML_d though, as the Sharpe ratio is affected the least when HML^{π} is added.

Overall, the evidence from this table suggests that HML^{π} is more important than other value factors, in particular, HML_d while jointly studying value and momentum. HML^{π} is also useful in using q-factors to construct an out-of-sample tangency portfolio. Its role is limited only when the newly developed EG factor is further considered.

3.1.4 Predicting monthly book-to-market

While Asness and Frazzini (2013) propose that monthly-updated book-to-market can predict the cross-section of returns better than the conventional measure using December market value, our previous evidence shows that the return spread sorted on π outperforms that on monthly-updated bm in factor regressions. This section examines if π achieves similarly strong return predictive power through tracking the monthly movement of future book-to-market. Specifically, we ask if π, as a stale value, already contains all the relevant information about future return in the cross-section as a more fresh value of bm does. After all, the argument of Asness and Frazzini (2013) is that the monthly updated book-to-market does not entail a better stand-alone value strategy, rather, it better handles the complex relationship between value and momentum strategies.

Table 6 about here.

Table 6 reports both equal- and value-weighted portfolio returns, sorted individually, by monthly bm, π, and a residual component of bm that is orthogonal to π. Since bm is monthly updated while π remains unchanged for twelve months, we estimate the residual as follows. Specifically, let December be month t, as π_t and bm of month $t+6, t+7, ...t+17$ are used to predict return of month $t+7, t+8, ...t+18$, we use a 5-year rolling window and conduct Fama-MacBeth regression:

$$bm_{t+j} \quad = \quad a_j + b_j \pi_t + \varepsilon_{t+j}; \; \forall j = 6, 7, ...17. \tag{13}$$

The residual component (ε_{t+j}) is calculated as the difference between the realized and predicted value of bm_{t+j} ($j = 6, 7, ...17$). Both bm and π are winsorized monthly at the 1st and 99th percentile before entering regressions.

In Panel A, monthly averages of equal-weighted portfolio decile 10 minus decile one return spread for monthly updated bm, π and residual component (ε) are 1.47% (t=6.36), 1.21% (t=7.49), and 0.71% (t=2.77), respectively. Notably, the decile 10 minus decile one portfolio based on π still achieves the largest Sharpe ratio of 1.02.

We focus more on Panel B as it reports value-weighted portfolio average returns. This time, decile 10 minus decile one return spread for monthly updated bm, π, and residual component (ε) are 0.35% (t=1.13), 0.42% (t=2.20), and 0.00% (t=0.01), respectively. Perhaps it is not surprising that a conventional value strategy, even based on a monthly updated bm, has been unprofitable in recent times. Still, π has the potential to become a better stand-alone value strategy. More importantly, a monthly updated bm does not seem to contain incremental information about future returns other than that embodied in the history of bm which could go back to 17 months ago.

3.1.5 Predicting profitability

Although book-to-market contains information about future returns, as shown in the decomposition of Cohen, Polk, and Vuolteenaho (2003), it is a noisy return predictor as equation (12) ties book-to-market to its future value, along with both return and profitability in the next period together. In fact, Cohen, Polk, and Vuolteenaho (2003) show that most of the cross-sectional variation in book-to-market is due to future book-to-market and profitability differences. In assessing the information content of bm about future return, they further claim that most of a growth stock's atypical valuation is simply due to high expected profitability rather than due to that stock having a low expected return. Similarly, contemplating the disconnect between book-to-market and the value premium, such that some growth firms earn value-like returns, and some value firms earn growth-like returns, Gerakos and Linnainmaa (2018) claim that it should be the component of the book-to-market that is orthogonal to information about future profitability that is more useful for expected returns.

We show that π strongly predicts returns in the cross-section, much more than book-to-market does. Additionally, it also tracks the future movement of book-to-market, as shown in Table 6. Thus, we would expect weaker predictability of profitability.

Table 8 about here.

We conduct FMB regressions and report results in Table 8 to test this conjecture. These regressions use end-of-year $t-1$'s π to forecast year t's book-to-market bm, log of ROE (roe), and return r (compounded from July of year t to June of year $t+1$) separately. Panel A uses π as the single predictor, and results show that there is indeed no predictability of profitability: the slope is 1.076 (t=1.044) in roe regression. In contrast, return predictability is still retained as the slope is 0.069 (t=4.281). Finally, π also predicts next year's book-to-market well: the slope is 0.732 (t=10.521).

In comparison, Panel B uses bm as the single predictor, and predictability of profitability is detected. The slope is -0.036 (t=-3.007) in the roe regression. Also, return predictability is weaker with a slope of 0.056 (t=5.217).

Our hypothesis is thus consistent with a Gerakos and Linnainmaa (2018), who posit that the permanent component of book-to-market is unrelated to discount rate differences. Although the decomposition methodology is different empirically, we confirm that the transitory component of book-to-market is a sharper return predictor but a noisier signal for profitability.

3.1.6 Predicting return in the long horizon

In this section, we examine the long-horizon evidence of return predictability, as per our decomposition (10), π should correlate with at least one term of the following: current transitory component of book-to-market, expected return, and expected *roe*. In other words, in the long run, π would predict either return or *roe* unless these two infinite sums are both expected to be zero.

There is already ample evidence in the recent literature that expected return and *roe* are not constant. In fact, modeling them as AR(1) process has been a popular practice (e.g., van Binsbergen and Koijen (2010)). Given that we have already documented a strong correlation between π and next period return, $Cov(\pi_t, \sum_{j=1}^{\infty} E_t\,(\mu_{t+j} - \bar{\mu}))$ would only equal zero if π significantly predicts lower return in the long horizon, so that short horizon positive return predictability would offset long horizon negative return predictability.

We do not attempt to estimate $\sum_{j=1}^{\infty} E_t\,(\mu_{t+j} - \bar{\mu})$ as it involves terms added up to infinity. Instead, we study each period j individually up to 5 years in the future for demonstration. Thus, $Cov(\pi_t, \sum_{j=1}^{\infty} E_t\,(\mu_{t+j} - \bar{\mu}))$ can be approximated as the sum of 5 covariance terms. Note that each covariance term corresponds to a cross-sectional trading strategy with π_t as the predictive variable. To approximate this covariance, we use the conventional decile portfolio sorts to facilitate comparison with findings in the existing literature. Specifically, we ask if portfolios sorted on past values of π, albeit becoming stale and noisier with time, can still generate a profitable return spread.

Table 9 about here.

Panel A of Table 9 reports equal-weighted decile 10 minus decile one return spread based on *bm* of December of prior year, June of current year, most recent month of current year, and π. The rightmost column shows that all of them can predict returns even with a five-year lag.

Panel B reports value-weighted results. The rightmost column indicates that in the 5th year, only π retains return predictive power, with a monthly average return spread of 0.238% (t=1.756).

In contrast, the average return spread is only 0.081% (t=0.441) for bm of prior December, 0.016% (t=0.086) for June of the current year, and 0.309% (t=1.502) for the most recent month of the current year. Except for the immediate investment horizon for the most recent month, the return predictive power of the other two versions of value decays over the years, as both the average return spread and t-values are monotonically decreasing. In sharp contrast, this pattern is absent for π, as the mean return spread changes from next year's 0.324% (t=1.992) to the third year's 0.171% (t=1.214), then picks up again to the fifth year's 0.238% (t=1.756). Overall, we conclude that π indeed can forecast returns to longer holding periods for at least five years.

3.2 Industry portfolio returns

Lewellen (1999) finds that the book-to-market ratio predicts returns both for the market and for industry portfolios in a time-series setting. In this section, we also investigate whether the prospective book-to-market ratios at the industry level can predict industry returns. Our industry classification, returns, book values, and market values are all from Kenneth French's data library. We focus on the 48 industry portfolios and compute the end-of-year industry book-to-market ratio by dividing the book value at the end of the previous year by the market value at the end of the current year. We then estimate the three parameters for each industry using its return, roe, and bm, to compute its prospective book-to-market π. The sample period for our industry portfolio data spans from 1926 to 2022.

3.2.1 Parameter estimation

Table 10 presents the mean and standard deviation of these three parameter estimates. Overall, $\bar{\theta}$ is negative, while β and ρ are less than one, consistent with their model assumptions. Still, there are large cross-industry differences. The SODA industry has the largest β of 0.950, while the HLTH industry has the smallest of 0.760. Similarly, the OTHER industry has the largest ρ of 0.959, while AERO industry has the smallest of 0.765. These large differences certainly create strong cross-industry variation between bm and π.

Table 10 about here.

Table 11 presents the summary statistics of the excess returns, bm, and π for each industry. For each of these three variables, we report the mean, standard deviation, and AR(1) coefficient. Variations in excess returns, bm, and π across industries are considerable. The largest average excess return is 0.169% in the FUN industry, while the smallest is 0.012% in the OTHER industry. TXTL reports the largest bm at -0.156, whereas DRUGS has the smallest at -1.469. INSUR reports the largest π at 0.047, whereas SODA has the smallest at -17.803. Additionally, consistent with the literature, the AR(1) coefficients for excess returns are predominantly negative across all industries, while those for bm are very large. In comparison, the persistence of π is much milder.

Table 11 about here.

Table 12 further displays the cross-industry average of summary statistics after pooling all industry estimates. Panel A provides the summary statistics of three model parameters, returns, bm, and π. Compared to bm, π exhibits higher volatility (3.082 vs. 0.495) and lower persistence (0.579 vs. 0.834). Panel B shows the correlations among the variables of interest, where the cross-industry correlation between the original and prospective book-to-market ratios is 0.685.

Table 12 about here.

3.2.2 Cross-industry portfolio analysis

We now assess the return predictability of π on industry portfolios, again through the lens of a portfolio analysis. Each year-end, we sort the 48 industries into five quintiles based on bm and π separately. The highest and lowest quintiles each include 10 industry portfolios. We then report the average raw and risk-adjusted returns of these five quintiles, as well as the zero-cost high-minus-low industry portfolios in Table 13.

Table 13 about here.

Industry portfolio excess returns increase monotonically on both bm and π. While bm generates a cross-industry return spread of 0.287% per month, π yields a slightly larger spread of 0.292% per month. The statistical significance of the raw return spread based on bm is stronger than that reported in Lewellen (1999), which could be due to a larger number of industry portfolios in our analysis.[5]

However, the risk-adjusted returns show a much greater difference between these two strategies. With the Fama-French 3-factor model, the bm HML α is -0.035% ($t = -0.41$), while that for π is significantly higher at 0.183% ($t = 2.04$). With the Fama-French 5-factor model, the bm HML α is -0.061% ($t = -0.69$), while that for π is still 0.183% ($t = 2.09$). Finally, when we further add the momentum factor, the bm HML α is -0.030% ($t = -0.34$), while that for π is still 0.167% ($t = 1.84$).

Again, we conclude that π is a more useful return predictor, even when trading industry portfolios.

3.3 Market returns

Finally, we estimate π on the market level and examine its predictive power for market returns. For the aggregate market data, we rely on the dataset from Goyal and Welch (2008), which is available on Amit Goyal's website. The book values are sourced from Value Line and Dow Jones. The annual book-to-market ratio is calculated as the ratio of the book value at the end of the previous year to the market value at the end of the current year, specifically for the Dow Jones Industrial Average. This dataset starts in 1921 and the market returns data from Kenneth French's website begins in 1926, with both datasets ending in 2022.

3.3.1 Parameter estimation

We again use the same three parameters to construct market π. This time, data availability poses an even bigger challenge in our estimation. To facilitate a fair comparison with bm, these parame-

[5]When using 12 Fama-French industry portfolios, we obtain results that are similar to those reported by Lewellen (1999). The results are available upon request.

ters are re-estimated annually using only the data available at the time of estimation. Specifically, we start with the first 10 observations to obtain their initial estimates. Each subsequent year, we add one more observation and re-estimate both parameters to generate the out-of-sample time series of π.[6] All the model parameters are updated annually. We emphasize that each observation is constructed using only the information available then, ensuring there is no look-ahead bias in obtaining the π variable.

Table 14 Panel A again presents the summary statistics of market excess return, bm, three parameters, and π. The mean of the long-run trend is -0.550, slightly higher than the full sample average of -0.737. The average β is 0.790, confirming that bm is a slow-moving random variable. The discount coefficient has a mean of 0.943, with a min of 0.880 and a maximum of 0.978.

Table 14 about here.

Given the extensive study of market returns in prior literature, we focus only on comparing π to bm. Notably, π has a smaller mean in magnitude (-0.164 vs. -0.737 for bm), but both minimum and maximum are more extreme, also with larger kurtosis. This amplified variability is due to several outliers where π becomes exceptionally large as either β or ρ approaches one. Moreover, π is much less persistent than the original variable (0.651 vs. 0.939), which is expected since it removes the persistence inherent in bm. In Panel B, the pairwise correlation between bm and π is 0.769, indicating that the two variables still share a great deal of common information.

3.3.2 Predictive regressions

Different from previous sections, we now examine the predictability of the market risk premium in a time-series setting, with predictive regression results reported in Table 15.

Table 15 about here.

[6] Our main results depend on the choice of the initial year used in the estimation process.

Our results are based on the whole sample period, yet π still makes no use of future information. We find that bm marginally predicts the market risk premium, with a moderate t-statistics of 1.46.[7]. More importantly, when we use π as the predictor, its coefficient is 8.800 and highly significant, with a t-statistic of 2.36 and an adjusted R^2 of 2%. The economic significance is also notable; a one-standard-deviation increase in π predicts a positive excess return of 3.61% (computed as 8.800×0.410) in the following year.

Many known return predictors have failed in the post-oil shock sample (Goyal and Welch (2008)). In response, we examine the performance of predictors from 1975 to 2022 in Panel B. As indicated, bm further loses predictive power for market returns during this period, with a t-statistic of 0.50 and an R^2 of -0.02. In contrast, the point estimate of the slope of π, 0.102, is larger than that in the full sample and is statistically significant, with a t-statistic of 3.17 and an adjusted R^2 of 8%.

To minimize the effect of outliers, we first winsorize the 1933 observation—the largest value of π—and replace it with the next largest value. This adjustment only reduces the t-value to 2.35 from 2.36.

Goyal and Welch (2008) caution that in-sample predictability often fails to translate into out-of-sample (OOS) predictability. We examine the out-of-sample predictive power in Table 16 to address this concern. The metrics we use include adjusted R^2, root-mean-square deviation ($\Delta RMSE$), and a statistic that tests for equal mean-squared error between the unconditional forecast and the conditional forecast (MSE-F), as advocated by Goyal and Welch (2008).[8]

Table 16 about here.

Panel A presents the out-of-sample predictive test results from 1960 to 2022, using the first 15 observations to initialize the regression, also known as the burn-in period.[9] As documented in

[7] Because our sample period covers the 2009 financial crisis and the 2022 Pandemic, this result is generally consistent with, albeit different from, those in Goyal and Welch (2008) and Kothari and Shanken (1997)

[8] We provide the formulas for these measures in the appendix. These statistics are used to examine the relative performance of a predictor against the historical mean.

[9] Because we require an additional 10 observations to begin estimating the prospective book-to-market ratio, the first 25 observations are excluded from evaluating model performance. Our results are not sensitive to the choice of the out-of-sample period.

the existing literature, bm shows negative values across all three metrics we examine. In contrast, the prospective book-to-market ratio, π, exhibits stronger out-of-sample performance. However, all three measures remain slightly negative, suggesting underperformance relative to a simple historical moving average estimate. Keep in mind that we need to run regressions to estimate β and ρ, and the estimation uncertainty may weaken the predictive power of π. To explore this possibility, we construct two alternatives. We assume that ρ is constant and denote this first alternative as π'. Then, we assume that β is constant and denote this second alternative as π''.

We find that indeed, the estimation noise in ρ could be the culprit as all three measures are significantly positive for π'. The adjusted out-of-sample R^2 is 3.4% with a p-value of 0.00, while $\Delta RMSE$ is 0.3% (p=0.00) and $MSEF$ is 2.788 (p=0.00), significantly outperforming the naïve model that uses a simple moving average. In comparison, ignoring the time variation of β appears to weaken our out-of-sample results further.

Previous literature also indicates that commonly used predictors perform even worse in the modern sample starting from 1975, following the Oil Crisis. To investigate whether our proposed predictor suffers from this reduced predictive power in the later period, we conduct the out-of-sample test using the first 45 observations to initialize the estimate and report the results in Panel B. We find that bm's predictive power further deteriorated during this period. Again, π' still achieves a significant R^2 at 4.6% with a p-value of 0.1. Collectively, the results in Tables 15 and 16 underscore the robust predictive power of our prospective ratio, both in-sample and out-of-sample, across different periods.

4 Conclusion

We model the transitory component of the book-to-market ratio as the sum of the present value of three demeaned terms: stock return, return on equity, and prospective book-to-market. We introduce an empirical proxy for the last term as a novel predictor of stock returns. This new variable requires estimates of the long-run trend, persistence, and discount coefficients of the book-to-market ratio and exhibits greater variations compared to conventional value measures.

Our empirical tests, which encompass the cross-section of individual firms, industry portfolios, and the stock market, reveal that the prospective book-to-market ratio can significantly predict risk-adjusted returns. Prospective book-to-market proves to be useful for a better stand-alone value strategy, and the return spread based on it has the potential for an alternative value factor in asset pricing models.

References

Asness, C. S. and A. Frazzini (2013). The devil in hml's details. *Journal of Portfolio Management 39*(4), 50–69.

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

Asness, C. S., R. B. Porter, and R. L. Stevens (2000). Predicting stock returns using industry-relative firm characteristics. Available at SSRN: http://ssrn.com/abstract=213872.

Beveridge, S. and C. R. Nelson (1981). A new approach to decomposition of economic time series into permanent and transitory components with particular attention to measurement of the business cycle. *Journal of Monetary Economics 7*(2), 151 – 174.

Campbell, J. Y. (1991). A variance decomposition for stock returns. *Economic Journal 101*(405), 157–179.

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

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

Cohen, R. B., P. A. Gompers, and T. Vuolteenaho (2002). Who underreacts to cash-flow news? evidence from trading between individuals and institutions. *Journal of Financial Economics 66*(2-3), 409–462.

Cohen, R. B., C. Polk, and T. Vuolteenaho (2003). The value spread. *The Journal of Finance 58*(2), 609–641.

Daniel, K. and S. Titman (2006). Market reactions to tangible and intangible information. *The Journal of Finance 61*(4), pp. 1605–1643.

Dong, M., S. Goto, K. Hou, Y. Xu, and Y. Zhang (2024). Beyond carry: The prospective interest rate differential and currency excess returns. Fisher School working paper.

Engel, C. (2016). Exchange rates, interest rates, and the risk premium. *American Economic Review 106*, 436–474.

Fama, E. F. and K. R. French (1992). The cross-section of expected stock returns. *The Journal of Finance 47*(2), 427–465.

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

Fama, E. F. and K. R. French (2008). Average returns, b/m, and share issues. *The Journal of Finance 63*(6), 2971–2995.

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

Gerakos, J. and J. T. Linnainmaa (2018). Decomposing value. *The Review of Financial Studies 31*(5), 1825–1854.

Goyal, A. and I. Welch (2008). A comprehensive look at the empirical performance of equity premium prediction. *The Review of Financial Studies 21*(4), pp. 1455–1508.

Hou, K., C. Xue, and L. Zhang (2015). Digesting anomalies: An investment approach. *Review of Financial Studies 28*(3), 650–705.

Kothari, S. and J. Shanken (1997). Book-to-market, dividend yield, and expected market returns: A time-series analysis. *Journal of Financial Economics 44*(2), 169 – 203.

Lewellen, J. (1999). The time-series relations among expected return, risk, and book-to-market. *Journal of Financial Economics 54*(1), 5 – 43.

van Binsbergen, J. H. and R. Koijen (2010). Predictive regressions: A present-value approach. *The Journal of Finance 65*(4), 1439–1471.

Vuolteenaho, T. (2002). What drives firm-level stock returns? *The Journal of Finance 57*(1), pp. 233–264.

Table 1: Summary statistics of return and predictors: cross section of firms

Panel A: Summary statistics

	r	bm	π
Mean	1.17	0.79	0.27
StdDev	15.42	1.77	2.24
Skew	3.45	-9.93	-0.88
Kurt	76.63	724.49	32.98
Min	-76.65	-71.74	-17.06
Median	0.20	0.67	0.19
Max	270.68	21.46	35.34

Panel B: Correlations

	r	bm	π
r	1.000	0.015	0.019
bm	0.015	1.000	0.548
π	0.019	0.548	1.000

This table reports the summary statistics from July 1962 to December 2023 on sample monthly stock returns r, the annually constructed book-to-market bm, and the prospective log book-to-market ratios π. Panel A reports the time-series average of their cross-sectional summary statistics. Panel B reports the monthly average of their cross-sectional correlations. Returns are expressed in percentage points.

Table 2: Asset pricing tests on the prospective book-to-market factor

Panel A: Factor summary statistics (%)							
	N	Mean	Std	NW t-stat	Min	Max	Sharpe
HML^{π}	726	0.343	2.3143	3.5148	-11.848	12.596	0.513
MKT	726	0.568	4.4973	3.3739	-23.240	16.100	0.438
SMB	726	0.215	3.0328	1.8152	-15.320	18.280	0.245
HML	726	0.292	2.9947	2.2295	-13.870	12.750	0.337
RMW	726	0.283	2.2254	3.2212	-18.650	13.070	0.441
CMA	726	0.273	2.0769	3.0616	-7.220	9.070	0.455
MOM	726	0.600	4.2128	3.7431	-34.300	18.200	0.494
HML_6	726	0.280	3.2876	2.0601	-13.099	16.799	0.295
HML_d	726	0.290	3.5446	2.0279	-17.991	26.864	0.283
HML_s	702	0.311	2.9413	2.5263	-14.825	15.851	0.366

This table reports the factor summary statistics. Every June, stocks are independently assigned to two portfolios by median NYSE market equity and three portfolios by 30th and 70th percentile NYSE π breakpoints. HML^{π} is the difference in the average value-weighted return of the two portfolios with the highest 30% π and those with the lowest 30% π. Returns are expressed in percentage points. The sample period is from July 1963 (January 1967 for the q-factors) to December 2023.

Table 3: Asset pricing tests on the prospective book-to-market factor

Panel B: Time series regression, no UMD factor						
MKT	-0.010	0.003	0.016	-0.012	-0.011	-0.007
	(-0.551)	(0.151)	(0.954)	(-0.717)	(-0.826)	(-0.482)
SMB	0.126***	0.096***	0.099***	0.113***	0.013	0.026
	(5.018)	(4.124)	(4.263)	(4.429)	(0.588)	(1.251)
HML	0.597***	0.528***				0.190***
	(22.536)	(15.555)				(2.963)
RMW		-0.134***	-0.125***	-0.052	-0.105***	-0.147***
		(-3.473)	(-3.530)	(-1.121)	(-3.038)	(-4.604)
CMA		0.170***	0.227***	0.397***	0.179***	0.017
		(3.412)	(4.765)	(6.130)	(3.936)	(0.358)
HML_6			0.452***			0.153*
			(13.664)			(1.888)
HML_d				0.323***		-0.018
				(8.177)		(-0.355)
HML_s					0.561***	0.373***
					(15.494)	(9.421)
α	0.147***	0.158***	0.159***	0.138**	0.163***	0.182***
	(2.724)	(2.922)	(2.792)	(2.149)	(3.157)	(3.899)
Obs	726	726	726	726	702	702
R^2	0.627	0.657	0.642	0.591	0.688	0.751
Panel B: Time series regression, with UMD factor						
MKT	-0.017	-0.004	0.020	-0.002	-0.016	-0.006
	(-1.003)	(-0.228)	(1.156)	(-0.118)	(-1.184)	(-0.431)
SMB	0.125***	0.097***	0.098***	0.103***	0.016	0.026
	(4.826)	(4.185)	(4.208)	(4.130)	(0.743)	(1.244)
HML	0.585***	0.508***				0.180***
	(20.129)	(14.107)				(2.604)
UMD	-0.038	-0.041*	0.024	0.135***	-0.032*	0.010
	(-1.590)	(-1.860)	(1.372)	(4.789)	(-1.764)	(0.387)
RMW		-0.127***	-0.131***	-0.070	-0.100***	-0.147***
		(-3.303)	(-3.734)	(-1.553)	(-2.967)	(-4.588)
CMA		0.185***	0.209***	0.282***	0.191***	0.014
		(3.765)	(4.334)	(4.777)	(4.274)	(0.283)
HML_6			0.472***			0.154*
			(12.718)			(1.908)
HML_d				0.461***		-0.003
				(11.147)		(-0.052)
HML_s					0.544***	0.372***
					(14.354)	(9.528)
α	0.178***	0.186***	0.143**	0.049	0.185***	0.175***
	(2.973)	(3.104)	(2.401)	(0.739)	(3.327)	(3.462)
Obs	726	726	726	726	702	702
R^2	0.631	0.662	0.644	0.618	0.691	0.751

This table reports the performance of the prospective book-to-market ratio factor (HML^{π}) against standard risk factor models. Every June, stocks are independently assigned to two portfolios by median NYSE market equity and three portfolios by 30th and 70th percentile NYSE π breakpoints. HML^{π} is the difference in the average value-weighted return of the two portfolios with the highest 30% π and those with the lowest 30% π. Panel A (B) reports the time series regressions excluding (including) UMD as an independent variable. We consider the following risk factors: the q-factors (MKT, ME, IA, ROE, and EG) as in Hou, Xue, and Zhang (2015), the Fama-French five factors (MKT, SMB, HML, RMW, and CMA) as in Fama and French (2015), the momentum factor (UMD), and the size-component of the value factor (HML_s) as in Gerakos and Linnainmaa (2018). Returns are expressed in percentage points. t-statistics (in parentheses) are based on Newey-West robust standard errors. Statistical significance at the 1%, 5%, and 10% levels is indicated by ***, **, and *, respectively. The sample period is from July 1963 (January 1967 for the q-factors) to December 2023.

Table 4: Asset pricing with the prospective book-to-market factor

	Panel A: Regressing HML on HML^{π}			Panel B: Regressing HML_6 on HML^{π}		
HML^{π}	1.013***	0.777***	0.745***	1.082***	0.851***	0.730***
	(22.613)	(16.335)	(17.317)	(18.525)	(14.473)	(11.997)
MKT	-0.048*	0.019	0.010	-0.073**	-0.006	-0.039
	(-1.659)	(0.632)	(0.363)	(-1.999)	(-0.174)	(-1.332)
SMB	-0.107	-0.024	-0.020	-0.114	-0.027	-0.013
	(-1.563)	(-0.511)	(-0.452)	(-1.623)	(-0.622)	(-0.376)
RMW		0.210***	0.214***		0.224***	0.237***
		(2.699)	(2.988)		(2.744)	(4.474)
CMA		0.474***	0.489***		0.468***	0.524***
		(10.101)	(9.958)		(7.421)	(10.237)
MOM			-0.053*			-0.203***
			(-1.898)			(-5.646)
α	-0.005	-0.169**	-0.127	-0.025	-0.193**	-0.033
	(-0.059)	(-2.131)	(-1.632)	(-0.294)	(-2.156)	(-0.392)
Obs	726	726	726	726	726	726
	Panel C: Regressing HML_d on HML^{π}			**Panel D: Regressing HML_s on HML^{π}**		
HML^{π}	1.086***	0.919***	0.643***	1.007***	0.818***	0.783***
	(11.892)	(8.225)	(10.542)	(24.060)	(13.527)	(14.126)
MKT	0.040	0.082*	0.007	-0.012	0.040*	0.030
	(0.985)	(1.913)	(0.290)	(-0.440)	(1.677)	(1.337)
SMB	-0.083	-0.041	-0.009	0.050	0.112***	0.116***
	(-0.816)	(-0.575)	(-0.188)	(1.154)	(2.778)	(2.947)
RMW		0.075	0.106		0.146***	0.149***
		(0.481)	(1.382)		(2.869)	(3.017)
CMA		0.325***	0.452***		0.376***	0.393***
		(3.951)	(9.821)		(6.888)	(7.143)
MOM			-0.460***			-0.058***
			(-12.880)			(-2.865)
α	-0.087	-0.172	0.190**	-0.049	-0.172***	-0.125*
	(-0.902)	(-1.316)	(2.457)	(-0.711)	(-2.712)	(-1.928)
Obs	726	726	726	702	702	702

This table tests whether the prospective book-to-market factor (HML^{π}) spans the three versions of HML, each constructed using the market cap at December end (Fama and French (1992)), June end (HML_6, Asness and Frazzini (2013)), and previous month end (HML_d, Asness and Frazzini (2013). Returns are expressed in percentage points. t-statistics (in parentheses) are based on Newey-West robust standard errors. Statistical significance at the 1%, 5%, and 10% levels is indicated by ***, **, and *, respectively. The sample period is from July 1963 to December 2023 (726 months).

Table 5: Out-of-sample mean-variance frontier

Panel A: Fama-French models: original HML										
SR	MKT	SMB	HML	RMW	CMA	MOM				
0.501	0.390	0.098	0.512							
0.844	0.185	0.110	-0.045	0.339	0.410					
0.881	0.176	0.096	0.012	0.287	0.327	0.102				
Panel B: Fama-French models: HML_6										
SR	MKT	SMB	HML_6	RMW	CMA	MOM				
0.468	0.444	0.102	0.454							
0.810	0.182	0.106	-0.055	0.337	0.430					
0.807	0.177	0.094	0.039	0.281	0.207	0.110				
Panel C: Fama-French models: HML_d										
SR	MKT	SMB	HML_d	RMW	CMA	MOM				
0.463	0.510	0.105	0.385							
0.812	0.190	0.109	-0.023	0.331	0.393					
0.855	0.165	0.085	0.162	0.249	0.167	0.172				
Panel D: Fama-French models: HML_s										
SR	MKT	SMB	HML_d	RMW	CMA	MOM				
0.522	0.428	-0.018	0.590							
0.848	0.191	0.124	-0.062	0.325	0.422					
0.885	0.181	0.095	0.003	0.276	0.339	0.105				
Panel E: Fama-French models: HML^{π}										
0.606	0.320	-0.027	0.707							
0.826	0.182	0.098	0.115	0.321	0.284					
0.914	0.169	0.072	0.165	0.269	0.211	0.114				
Panel F: Fama-French models: all HML factors										
SR	MKT	SMB	HML	RMW	CMA	MOM	HML_6	HML_d	HML_s	HML^{π}
0.412	0.326	-0.030	0.301				-0.107	-0.184	-0.111	0.805
0.867	0.163	0.083	-0.031	0.342	0.378		-0.164	0.041	-0.166	0.352
0.956	0.131	0.053	-0.222	0.261	0.221	0.199	-0.136	0.343	-0.112	0.261

This table reports the Sharpe ratios (SR) and time-series averaged portfolio weights based on the Fama-French factors and q-factors. Every month, we use 360 months of prior returns to estimate the mean and variance-covariance matrix of the factors to construct an out-of-sample tangency portfolio. Panels A to D report the portfolio weights based on Fama-French factors (July 1963 to December 2023), and Panels E to H report those based on q-factors (January 1967 to December 2023).

Table 6: Predicting monthly bm

		Decile										
		1	2	3	4	5	6	7	8	9	10	D10–D1
Panel A: Equal-weighted returns												
monthly bm	$\bar{r}$	0.56	0.73	0.81	0.89	0.98	1.09	1.17	1.33	1.58	2.03	1.47
	t	1.67	2.59	3.16	3.70	4.26	4.73	4.93	5.54	5.79	5.39	6.36
	SR	0.22	0.35	0.42	0.49	0.57	0.63	0.66	0.74	0.77	0.72	0.85
π	$\bar{r}$	0.44	0.75	0.91	1.05	1.12	1.20	1.34	1.37	1.48	1.64	1.21
	t	1.57	2.92	3.68	4.23	4.45	4.75	5.32	5.30	5.37	5.53	7.49
	SR	0.21	0.39	0.49	0.56	0.59	0.63	0.71	0.71	0.72	0.74	1.00
ε	$\bar{r}$	1.19	0.99	0.89	0.85	0.92	0.98	1.04	1.13	1.41	1.90	0.71
	t	4.01	3.64	3.52	3.52	3.91	4.21	4.49	4.74	5.12	4.90	2.77
	SR	0.53	0.49	0.47	0.47	0.52	0.56	0.60	0.63	0.68	0.65	0.37
Panel B: Value-weighted returns												
monthly bm	$\bar{r}$	0.99	0.85	0.87	0.96	0.90	1.02	0.93	1.13	1.23	1.34	0.35
	t	4.00	4.49	4.84	5.41	5.23	5.81	4.84	5.47	4.93	3.97	1.13
	SR	0.53	0.60	0.64	0.72	0.70	0.77	0.64	0.73	0.66	0.53	0.15
π	$\bar{r}$	0.74	0.92	0.86	0.92	1.07	1.03	1.10	1.06	1.08	1.16	0.42
	t	3.52	4.57	4.71	5.13	6.16	5.15	6.08	5.35	4.98	5.16	2.20
	SR	0.47	0.61	0.63	0.68	0.82	0.69	0.81	0.71	0.66	0.69	0.29
ε	$\bar{r}$	1.14	0.98	0.88	0.84	0.94	0.88	0.95	0.95	1.07	1.14	0.00
	t	5.30	4.92	4.54	4.40	5.36	5.03	5.14	4.62	4.65	3.48	0.01
	SR	0.71	0.65	0.61	0.59	0.71	0.67	0.68	0.61	0.62	0.46	0.00

This table reports the average monthly equal and value-weighted portfolio returns sorted on monthly-updated bm, π, and residual ε (the difference between realized and predicted value of monthly-updated bm). Specifically, let December be month t, as π_t and bm of month $t+6, t+7, ...t+17$ are used to predict return from month $t+7, t+8, ...t+18$, we use a five-year rolling window and conduct Fama-MacBeth regression:

$$bm_{t+j} \quad = \quad a_j + b_j \pi_t + \varepsilon_{t+j};\ \forall j = 6, 7, ...17.$$

Both monthly bm and π are winsorized at the 1st and 99th percentiles using Newey-West robust standard errors, and annualized Sharpe ratios (SR). The return prediction sample period is July 1967 to December 2023 (678 months).

Table 7: Predicting profitability

Panel A: predicting with π

	bm_{t+1}	roe_{t+1}	r_{t+1}
π_t	0.677***	-0.026	0.069***
	(10.641)	(-1.313)	(4.281)
Intercept	0.760***	0.026	0.142***
	(17.502)	(1.011)	(5.955)
Obs	211,589	211,488	211,712
R^2	0.305	0.045	0.013

Panel B: predicting with bm

	bm_{t+1}	roe_{t+1}	r_{t+1}
bm_t	0.859***	-0.036***	0.056***
	(28.429)	(-3.007)	(5.217)
Intercept	0.167***	0.056*	0.109***
	(8.998)	(1.845)	(4.744)
Obs	220,646	220,537	220,779
R^2	0.611	0.056	0.015

This table reports the Fama-MacBeth regression results using year $t-1$'s π or book-to-market (bm) to forecast year t's book-to-market, log of Return on equity (roe), and annual return r compounded from July of year t to June of year $t+1$). The sample period is July 1962 to December 2023.

Table 8: Predicting profitability

	π	$\pi + \theta_t$ (baseline)	$bm >= 0.01$	$bm >= 0.01$ + winsor
Panel A: roe_ff1, t+1				
r_{t+1}	0.281**	0.525***	0.511***	0.395***
	(2.349)	(3.697)	(3.561)	(5.366)
roe_{t+1}	-0.993***	-1.770***	-1.699***	-1.444***
	(-2.812)	(-3.136)	(-2.971)	(-2.952)
α	0.364**	-0.073	-0.069	-0.181
	(2.112)	(-0.293)	(-0.279)	(-1.069)
Obs	204651	204645	204439	204439
R^2	0.061	0.074	0.067	0.069
Panel B: roe_gaap, t+1				
	π	$\pi + \theta_t$ (baseline)	$bm >= 0.01$	$bm >= 0.01$ + winsor
r_{t+1}	0.283**	0.527***	0.527***	0.406***
	(2.567)	(4.143)	(4.143)	(6.359)
roe_{t+1}	-1.018***	-1.799***	-1.799***	-1.543***
	(-3.414)	(-3.849)	(-3.849)	(-3.771)
α	0.363**	-0.077	-0.077	-0.181
	(2.583)	(-0.417)	(-0.417)	(-1.519)
Obs	203043	203043	203043	203043
R^2	0.057	0.072	0.072	0.072
Panel C: roe_gaap, t+1 t+10				
	π	$\pi + \theta_t$ (baseline)	$bm >= 0.01$	$bm >= 0.01$ + winsor
r_{t+1}	0.364***	0.673***	0.673***	0.579***
	(3.922)	(5.918)	(5.918)	(7.526)
$roe_{t+1\ t+10}$	-0.467***	-0.891***	-0.891***	-0.812***
	(-4.438)	(-5.793)	(-5.793)	(-6.129)
α	0.424***	0.104	0.104	0.052
	(3.739)	(0.589)	(0.589)	(0.333)
Obs	86330	86330	86330	86330
R^2	0.144	0.186	0.186	0.205

Indentity $\pi_t \approx \sum_{j=1}^{\infty} E_t\left(r_{t+j} - \bar{\mu}\right) - \sum_{j=1}^{\infty} E_t\left(roe_{t+j} - \bar{g}\right) - \left(\theta_t - \theta_t^P\right)$ (Eq.(10)) decomposes π_t into three components: the infinite sum of future log returns, the infinite sum of future log ROEs, and the transitory component of the contemporaneous log book-to-market. We take $j = 1$ and run the following cross-sectional regression (firm subscript omitted):

$$\pi_t + \theta_t = \alpha + b_1 r_{t+1} - b_2 roe_{t+1} + \epsilon.$$

If π_t can predict r_{t+1} as in $r_{t+1} = \alpha_\pi + \beta_\pi \pi_t + \epsilon_\pi$, then

$$\begin{aligned} \beta_{\pi,t} &= \frac{Cov\left(r_{t+1}, \pi_t\right)}{Var\left(\pi_t\right)} = \frac{Cov\left(r_{t+1}, b_1 r_{t+1} - b_2 roe_{t+1} - \theta_t + \alpha + \epsilon\right)}{Var\left(\pi_t\right)} \\ &= \frac{b_1 Var\left(r_{t+1}\right)}{Var\left(\pi_t\right)} - \frac{b_2 Cov\left(r_{t+1}, roe_{t+1}\right)}{Var\left(\pi_t\right)} - \frac{Cov\left(r_{t+1}, \theta_t\right)}{Var\left(\pi_t\right)} + \frac{Cov\left(r_{t+1}, \alpha + \epsilon\right)}{Var\left(\pi_t\right)} \\ &= \frac{Var\left(r_{t+1}\right)}{Var\left(\pi_t\right)} b_1 - \frac{Cov\left(r_{t+1}, roe_{t+1}\right)}{Var\left(\pi_t\right)} b_2 - \frac{Var\left(\theta_t\right)}{Var\left(\pi_t\right)} \beta_\theta + \frac{Cov\left(r_{t+1}, \alpha + \epsilon\right)}{Var\left(\pi_t\right)} \\ &= \beta_{1,t} + \beta_{2,t} + \beta_{3,t} + \beta_{4,t} \end{aligned}$$

where $\beta_\theta = \frac{Cov(r_{t+1},\theta_t)}{Var(\theta_t)}$ is the regression slope from $r_{t+1} = \alpha_\theta + \beta_\theta \theta_t + \epsilon_\theta$. The time-series average

Table 9: Predicting returns in the long hoirzon

		1	2	3	4	5
		Panel A. Equal-weighted returns				
bm	$\bar{r}$	1.027	0.751	0.567	0.416	0.313
	t	5.998	4.484	3.527	2.674	2.071
	SR	0.792	0.592	0.466	0.353	0.273
bm_6	$\bar{r}$	1.037	0.843	0.589	0.440	0.362
	t	6.212	5.144	3.544	2.744	2.412
	SR	0.820	0.679	0.468	0.362	0.318
bm_d	$\bar{r}$	1.443	1.308	0.919	0.565	0.568
	t	6.312	7.001	5.181	3.384	3.698
	SR	0.833	0.924	0.684	0.447	0.488
π	$\bar{r}$	0.975	0.660	0.466	0.441	0.364
	t	7.113	5.015	3.898	3.593	3.139
	SR	0.939	0.662	0.514	0.474	0.414
		Panel B. Value-weighted returns				
bm	$\bar{r}$	0.388	0.303	0.225	0.086	0.081
	t	1.977	1.659	1.259	0.493	0.441
	SR	0.261	0.219	0.166	0.065	0.058
bm_6	$\bar{r}$	0.305	0.297	0.178	0.113	0.016
	t	1.507	1.575	0.958	0.617	0.086
	SR	0.199	0.208	0.126	0.081	0.011
bm_d	$\bar{r}$	0.333	0.544	0.455	0.362	0.309
	t	1.097	2.356	2.082	1.773	1.502
	SR	0.145	0.311	0.275	0.234	0.198
π	$\bar{r}$	0.324	0.233	0.171	0.301	0.238
	t	1.992	1.573	1.214	2.087	1.756
	SR	0.263	0.208	0.160	0.275	0.232

This table reports the high-minus-low decile returns for four variables: three book-to-market ratios each constructed using December (bm, Fama and French (1992)), June (bm_6, Asness and Frazzini (2013)), and most recent month (bm_d, Asness and Frazzini (2013)), respectively, and the prospective book-to-market (π). Conventional decile portfolios are held from July to next June, and sorting variable are values of each variable in the past 1, 2, ...5 years. We report average return spread $\bar{r}$ between Decile 10 and 1, t-statistics calculated using Newey-West robust standard errors, and annualized Sharpe ratios (SR). Sample period is from July 1966 to December 2023 (690 months).

Table 10: Parameter Estimates for 48 Industries

Code	Industry	$\bar{\theta}$ Mean	$\bar{\theta}$ Std	β Mean	β Std	ρ Mean	ρ Std
1	AGRIC	-0.087	0.118	0.881	0.036	0.890	0.087
2	FOOD	-0.730	0.139	0.904	0.019	0.912	0.012
3	SODA	-1.379	0.120	0.950	0.015	0.909	0.014
4	BEER	-0.423	0.131	0.909	0.024	0.956	0.019
5	SMOKE	-0.724	0.421	0.938	0.026	0.896	0.041
6	TOYS	-0.834	0.331	0.847	0.032	0.860	0.020
7	FUN	-0.822	0.153	0.784	0.038	0.827	0.044
8	BOOKS	-0.693	0.242	0.858	0.019	0.868	0.051
9	HSHLD	-1.176	0.356	0.889	0.022	0.909	0.014
10	CLTHS	-0.468	0.226	0.843	0.036	0.817	0.043
11	HLTH	-0.812	0.220	0.762	0.051	0.927	0.059
12	MEDEQ	-1.363	0.126	0.808	0.042	0.921	0.026
13	DRUGS	-1.415	0.078	0.806	0.058	0.961	0.030
14	CHEM	-0.731	0.194	0.889	0.021	0.913	0.010
15	RUBBR	-0.598	0.167	0.820	0.020	0.863	0.043
16	TXTLS	0.011	0.193	0.811	0.026	0.806	0.035
17	BLDMT	-0.623	0.135	0.844	0.031	0.864	0.017
18	CNSTR	-0.343	0.135	0.803	0.028	0.863	0.038
19	STEEL	-0.034	0.112	0.859	0.015	0.879	0.030
20	FABPR	-0.261	0.159	0.820	0.021	0.874	0.044
21	MACH	-0.799	0.149	0.855	0.020	0.873	0.018
22	ELCEQ	-0.856	0.132	0.799	0.042	0.807	0.030
23	AUTOS	-0.477	0.146	0.850	0.018	0.849	0.033
24	AERO	-0.405	0.191	0.767	0.034	0.765	0.034
25	SHIPS	-0.167	0.285	0.776	0.033	0.781	0.064
26	GUNS	-0.413	0.220	0.836	0.077	0.911	0.086
27	GOLD	-0.861	0.123	0.830	0.043	0.889	0.040
28	MINES	-0.347	0.167	0.847	0.012	0.888	0.048
29	COAL	-0.343	0.255	0.817	0.079	0.803	0.121
30	OIL	-0.230	0.075	0.811	0.034	0.853	0.029
31	UTIL	-0.239	0.246	0.904	0.028	0.890	0.019
32	TELCM	-0.527	0.193	0.849	0.017	0.933	0.069
33	PERSV	-0.964	0.261	0.816	0.048	0.818	0.101
34	BUSSV	-1.441	0.399	0.810	0.017	0.905	0.033
35	COMPS	-1.357	0.197	0.787	0.039	0.907	0.022
36	CHIPS	-1.090	0.235	0.778	0.024	0.837	0.030
37	LABEQ	-0.818	0.141	0.822	0.020	0.830	0.045
38	PAPER	-0.415	0.127	0.898	0.010	0.904	0.018
39	BOXES	-0.304	0.145	0.828	0.038	0.871	0.036
40	TRANS	-0.127	0.192	0.841	0.017	0.896	0.029
41	WHLSL	-0.640	0.170	0.785	0.034	0.857	0.038
42	RTAIL	-0.942	0.203	0.879	0.017	0.920	0.016
43	MEALS	-0.972	0.188	0.811	0.036	0.870	0.040
44	BANKS	-0.286	0.139	0.841	0.017	0.874	0.028
45	INSUR	-0.315	0.120	0.846	0.046	0.885	0.059
46	RLEST	-0.457	0.273	0.856	0.025	0.943	0.051
47	FIN	-0.439	0.176	0.849	0.020	0.902	0.009
48	OTHER	-1.299	0.328	0.829	0.030	0.959	0.057
Average		-0.738	0.482	0.834	0.045	0.889	0.052

This table reports the mean and standard deviation of parameter estimates ($\bar{\theta}$, β, and ρ) for Fama-French 48 industries and the whole sample (last row). The sample period is 1960~2023.

Table 11: 48 Industry Portfolios: return, bm, and π

	r_e			bm			π		
Industry	*Mean*	*Std*	*AR*(1)	*Mean*	*Std*	*AR*(1)	*Mean*	*Std*	*AR*(1)
Agric	0.086	0.257	-0.013	-0.711	0.545	0.826	-0.227	0.430	0.742
Food	0.099	0.179	0.112	-0.805	0.403	0.874	-0.329	0.434	0.662
Soda	0.098	0.234	0.085	-1.276	0.592	0.832	-17.803	124.458	-0.002
Beer	0.097	0.189	0.167	-1.139	0.870	0.971	-0.455	0.432	0.724
Smoke	0.149	0.235	-0.071	-1.177	0.808	0.884	-2.001	2.640	0.806
Toys	0.097	0.353	-0.215	-0.899	0.489	0.766	-0.716	0.830	0.715
Fun	0.169	0.320	-0.147	-0.818	0.526	0.736	-0.527	0.485	0.490
Books	0.128	0.286	-0.111	-0.813	0.376	0.806	-0.509	0.650	0.652
Hshld	0.116	0.181	-0.066	-1.267	0.511	0.920	-0.431	0.495	0.702
Clths	0.137	0.290	-0.246	-0.717	0.591	0.912	-0.518	0.573	0.513
Hlth	0.139	0.362	-0.149	-0.739	0.508	0.700	-0.324	0.807	0.248
MedEq	0.144	0.208	-0.039	-1.188	0.380	0.822	-0.376	1.023	0.410
Drugs	0.132	0.189	0.212	-1.469	0.402	0.908	-0.247	0.318	0.774
Chems	0.110	0.207	-0.148	-0.765	0.419	0.876	-0.002	0.278	0.739
Rubbr	0.138	0.241	-0.176	-0.695	0.541	0.923	-0.468	0.365	0.771
Txtls	0.118	0.291	-0.110	-0.156	0.509	0.806	-0.289	0.316	0.724
BldMt	0.120	0.220	-0.222	-0.627	0.439	0.868	-0.135	0.270	0.712
Cnstr	0.133	0.296	-0.228	-0.400	0.320	0.566	-0.207	0.248	0.484
Steel	0.090	0.283	-0.177	-0.177	0.492	0.825	-0.283	0.306	0.676
FabPr	0.102	0.312	-0.208	-0.396	0.361	0.599	-0.240	0.245	0.480
Mach	0.121	0.228	-0.339	-0.697	0.475	0.904	-0.333	0.280	0.646
ElcEq	0.131	0.244	-0.182	-0.921	0.422	0.804	-0.208	0.318	0.656
Autos	0.120	0.308	-0.078	-0.462	0.555	0.818	-0.085	0.436	0.799
Aero	0.138	0.287	-0.138	-0.702	0.612	0.906	-0.216	0.404	0.842
Ships	0.107	0.256	-0.090	-0.425	0.460	0.851	-0.243	0.261	0.469
Guns	0.135	0.266	-0.201	-0.850	0.970	0.913	-0.264	0.383	0.658
Gold	0.091	0.404	-0.226	-0.805	0.383	0.780	-0.408	0.320	0.831
Mines	0.125	0.292	-0.172	-0.621	0.456	0.803	-0.258	1.055	0.277
Coal	0.167	0.491	-0.009	-0.448	0.516	0.678	-0.619	0.753	0.836
Oil	0.109	0.223	-0.167	-0.460	0.366	0.874	-0.191	0.533	0.139
Util	0.079	0.157	-0.225	-0.309	0.409	0.925	-0.052	0.135	0.855
Telcm	0.084	0.200	0.110	-0.477	0.465	0.899	-0.145	0.279	0.834
PerSv	0.079	0.288	-0.146	-0.855	0.383	0.744	-0.426	0.417	0.394
BusSv	0.113	0.254	-0.125	-1.098	0.588	0.892	-0.768	0.501	0.445
Comps	0.111	0.286	-0.023	-1.197	0.505	0.790	-0.118	0.400	0.709
Chips	0.121	0.306	-0.053	-0.961	0.542	0.891	-0.197	0.326	0.779
LabEq	0.115	0.269	-0.227	-1.024	0.378	0.694	-0.093	0.359	0.378
Paper	0.070	0.192	-0.239	-0.689	0.566	0.947	-0.182	0.308	0.863
Boxes	0.077	0.201	-0.225	-0.804	0.376	0.825	-0.229	0.275	0.698
Trans	0.080	0.244	-0.314	-0.285	0.691	0.955	-0.653	0.383	0.718
Whlsl	0.089	0.247	-0.163	-0.668	0.372	0.835	-0.482	0.497	0.318
Rtail	0.091	0.224	-0.120	-0.949	0.517	0.908	-0.346	0.589	0.657
Meals	0.088	0.250	-0.176	-1.015	0.541	0.875	-0.678	1.044	0.012
Banks	0.049	0.240	-0.129	-0.351	0.383	0.850	0.004	0.177	0.609
Insur	0.076	0.207	-0.112	-0.266	0.318	0.855	0.047	0.536	0.290
RlEst	0.066	0.332	-0.104	-0.477	0.527	0.813	-0.572	0.703	0.408
Fin	0.089	0.244	-0.095	-0.378	0.401	0.813	-0.215	0.288	0.620
Other	0.012	0.262	-0.091	-0.693	0.490	0.771	-0.184	0.395	0.035

This table reports the summary statistics of 48 industry portfolios' excess returns, log book-to-market ratios, and the prospective book-to-market estimates from 1960 to 2023.

Table 12: Summary statistics of return and predictors: industries

Panel A: Summary Statistics

	$\bar{\theta}$	β	ρ	r_e	bm	π
$Mean$	-0.293	0.843	0.871	0.107	-0.732	-0.713
Std	0.123	0.042	0.053	0.261	0.495	3.082
$Skew$	-0.132	0.358	-1.094	0.500	-0.004	-1.422
$Kurt$	-0.814	2.449	3.924	1.017	-0.334	8.240
min	-0.508	0.742	0.675	-0.457	-1.838	-22.611
$Median$	-0.271	0.839	0.884	0.092	-0.730	-0.298
max	-0.074	0.942	0.944	0.926	0.326	0.523
$AR(1)$	0.997	0.753	0.704	-0.121	0.834	0.579

Panel B: Correlations

	r_e	bm	π
r_e	1.000	-0.016	-0.001
bm	-0.016	1.000	0.685
π	-0.001	0.685	1.000

This table reports the Fama-French 48-industry summary statistics from 1960 to 2023. Panel A reports the historical log book-to-market average ($\bar{\theta}$), AR(1) coefficient estimates (β and ρ), excess returns (r_e), full-sample log book-to-market (bm), and the prospective log book-to-market ratios (π) of 48 industry portfolios. Panel B reports the linear correlations of r_e, bm, and π. All statistics are first computed at the industry level and summarized across industries.

Table 13: Cross-industry portfolio analysis

	r		FF3		FF5		FF5+MOM	
	bm	π	bm	π	bm	π	bm	π
Low	0.537	0.534	0.026	-0.074	0.306	0.176	0.357	0.233
	(2.87)	(2.67)	(0.36)	(-0.98)	(4.49)	(2.50)	(5.14)	(3.36)
2	0.551	0.550	-0.095	-0.142	0.176	0.127	0.241	0.183
	(2.74)	(2.72)	(-1.45)	(-2.37)	(2.62)	(2.14)	(3.68)	(3.13)
3	0.627	0.646	-0.083	-0.050	0.149	0.237	0.210	0.292
	(3.19)	(3.30)	(-1.22)	(-0.88)	(2.22)	(3.88)	(3.07)	(4.40)
4	0.725	0.705	0.013	0.018	0.289	0.266	0.304	0.334
	(3.83)	(3.79)	(0.20)	(0.27)	(4.23)	(3.75)	(4.50)	(4.63)
High	0.824	0.826	-0.009	0.110	0.245	0.359	0.327	0.401
	(4.01)	(4.43)	(-0.15)	(1.68)	(3.67)	(5.38)	(4.68)	(5.66)
HML	0.287	0.292	-0.035	0.183	-0.061	0.183	-0.030	0.167
	(2.38)	(2.93)	(-0.41)	(2.04)	(-0.69)	(2.09)	(-0.34)	(1.84)

This table presents the cross-industry predictability of bm and π sorted portfolios. Panel A reports the mean excess returns of quintile portfolios and the return spreads. Panels B, C, and D report the corresponding three-factor, five-factor, and six-factor (five factors plus MOM) αs. αs are estimated using monthly time-series regressions with Newey-West robust standard errors. The sample period is July 1963 to December 2023.

Table 14: Summary statistics of return and predictors: market

Panel A: Summary statistics

	r_e	bm	$\bar{\theta}$	β	ρ	π
$Mean$	8.346	-0.737	-0.550	0.790	0.943	-0.164
Std	19.967	0.543	0.095	0.094	0.015	0.410
$Skew$	-0.359	-0.425	-0.827	-0.105	-0.280	-4.287
$Kurt$	-0.069	-0.639	-0.322	0.756	3.192	25.025
min	-46.494	-2.031	-0.842	0.422	0.880	-3.044
$Median$	10.612	-0.643	-0.510	0.756	0.943	-0.022
max	52.879	0.366	-0.434	0.959	0.978	0.439
$AR(1)$	-0.022	0.939	0.956	0.758	0.689	0.651
Obs	94	94	94	94	94	94

Panel B: Correlations

	r_e	bm	π
r_e	1.000	-0.172	0.003
bm	-0.172	1.000	0.769
π	0.003	0.769	1.000

This table reports the summary statistics for annual market-level data from 1960 to 2023. Panel A reports the market excess returns in percentage points (r_e), the market log book-to-market (bm), the parameter estimates (the moving average $\bar{\theta}$ and the AR(1) coefficients β and ρ), and the prospective book-to-market (π). Panel B shows the linear correlations of market excess returns, market-level log book-to-market, and the prospective log book-to-market.

Table 15: Market Return: predictive regressions

	Full sample		Post oil shock		1933 winsorized	
	bm	π	bm	π	bm	π
b	5.635	8.800	2.322	10.206	5.328	8.771
	(1.46)	(2.36)	(0.50)	(3.17)	(1.38)	(2.35)
a	12.842	10.143	11.057	12.039	12.617	10.124
	(3.79)	(4.90)	(2.70)	(5.46)	(3.72)	(4.90)
Adj R^2	(0.01)	(0.02)	(-0.02)	(0.08)	(0.01)	(0.02)

This table reports predictive regression results of $r_{e,t+1} - r_{f,t+1} = a + bx_t + \varepsilon_{t+1}$, where $r_{e,t+1}$ is the annual CRSP value-weighted market return and $r_{f,t+1}$ is the annual risk-free rate, both in percentage points. Panel A reports the full sample results. Panel B reports the results from 1975 to 2023 (post oil shock period). Panel C reports the results with the independent variables winsorized at the 5th percentile of the time-series.

Table 16: Market return: out-of-sample Tests

Panel A: Full sample (1960 to 2022)

	bm	π	π'	π''
Adj R^2	-0.085	-0.008	0.034	-0.027
	(0.80)	(0.10)	(0.10)	(0.20)
$\Delta RMSE$	-0.737	-0.068	0.306	-0.237
	(0.80)	(0.10)	(0.10)	(0.20)
$MSEF$	-6.159	-0.599	2.788	-2.067
	(0.80)	(0.10)	(0.10)	(0.20)

Panel B: Post oil-shock sample (1975 to 2022)

	bm	π	π'	π''
Adj R^2	-0.242	-0.026	0.046	-0.120
	(1.00)	(0.60)	(0.00)	(1.00)
$\Delta RMSE$	-2.000	-0.229	0.409	-1.019
	(1.00)	(0.60)	(0.00)	(1.00)
$MSEF$	-10.525	-1.387	2.618	-5.790
	(1.00)	(0.60)	(0.00)	(1.00)

This table reports the results of out-of-sample (OOS) tests of the predictive regressions using the standard book-to-market (bm) and three versions of prospective book-to-market ratios: one using simultaneously estimated ρ and β as in Vuolteenaho (2002) (π), one using self-estimated β and treating ρ as constant (π'), and one using simultaneously estimated ρ and β but treating β as constant (π''). Panel A reports the out-of-sample results of the full sample (the first 15 observations are used to initialize the estimation). Panel B reports out-of-sample results for the post-oil-shock sample (the first 40 observations are used to initialize the estimation).