Is Institutional Buying More Informative Than Selling? Evidence From Book-to-Market Ratios

Dr. Marius M. Mihai
Assistant Professor of Economics
Widener University
School of Business Administration
One University Place
Chester, PA 19013
mmmihai@widener.edu

Abstract

I document that book-to-market ratios of stocks that experience stronger institutional buying pressure are more informative for future returns. The degree of information incorporated by these valuation proxies for stocks with different levels of institutional price pressure has direct implications for return predictability. Book-to-market ratios of stocks in the highest quintile of institutional buying pressure generate higher out-of-sample R^2s and economic gains compared to stocks in the lowest quintile. In interpreting my empirical findings, I focus the discussion on a few primary causes that may explain why institutional buying is more informative than selling for future market returns. First, the information embedded in valuation ratios of portfolios with the highest degree of buying pressure cannot be replicated from stocks that are outside institutions' trading sets. Second, there tends to be less disagreement among similar valuation ratios (e.g. sales-to-price, earnings-to-price) in portfolios of stocks belonging to the highest quintile of institutional buying pressure. Third, the source of predictability is most likely due to the ability of this particular subset of book-to-market ratios to anticipate the cash-flow news component of future returns.

Keywords: book-to-market ratio; institutional buying pressure; cash-flow news; economic gains

JEL Classification: G12, G20, G23

1. Introduction

Institutional investor trades represent a large proportion of total volume in US financial markets. Their demand affects financial markets outcomes in significant ways. Koijen and Yogo (2019) show that around 80% of variance in returns can be explained by institutional latent demand after controlling for other important cross-sectional characteristics. The informational role of institutions in financial markets, and in particular the connection between stock returns and institutional trades has been a topic of interest for asset pricing researchers.[1] There is ample evidence that institutional ownership and future returns are strongly linked. The study of Nofsinger and Sias (1999) argue that institutional herding and positive feedback trading play an important role in generating these type of lead-lag relationships between ownership and returns. An alternative view is offered by Gompers and Metrick (2001) who argue the relationship is triggered by demand shocks due to shifts in compositional ownership caused by institutions. Cai and Zheng (2004), in contrast to previous results, further explore this relationship in the context of Granger-Causality tests and find that institutional trading predicts future returns with a negative effect. Yan and Zhang (2009) bring forward this lack of directional relationship and attribute it to a weak assumption made by previous studies to include all institutional investors as one large homogenous group. By focusing on short-term institutional investors, they find strong evidence of a positive relationship between ownership and future returns. One common factor among these papers is they all investigate the informational role of institutional investors in financial markets by directly estimating relationships between some proxy of institutional demand and stock returns. An important question that still remains unanswered in this literature is the information role of buying and selling for the aggregate market. For example, is buying more informative than selling in the context of predicting future returns? If it is, then what drives these differences? The premise of my paper is to answer these questions and contribute to our understanding of the informational role of trading by institutions through exploring the heterogeneity of the effect of book-to-market ratios on aggregate markets triggered by different levels of institutional herding. Empirically, I differentiate my results from prior literature by focusing on understanding the effects of institutional herding on future market returns through the informational content of book-to-market ratios.

The choice of book-to-market is not arbitrary. These ratios have historically represented an accurate measure of value and an important factor for the cross-section as argued by Fama and French (1993). One interpretation of book-to-market values is that they reflect prospects for efficiency and investment. Firms

[1]Important works here include Gompers and Metrick (2001), Nofsinger and Sias (1999), Yan and Zhang (2009), Sias et al. (2006), Bennett et al. (2003), Choi and Sias (2012).

with better growth opportunities as reflected by changes in assets or with a superior managerial ability as reflected by return on equity, should display lower book-to-market ratios if markets are efficient. If institutions trade on private information and represent a finer filter for screening out more efficient firms with superior growth opportunities, that should be reflected in their book-to-market values. Hence, the value of information these ratios incorporate ought to reflect outcomes of returns and establish a connection between institutional trades and the aggregate market.

To assess the informational effect of institutional trading on future market returns, I create a proxy variable called institutional buying pressure (IBP). IBP is simply a measurement of upward and downward price movement created by institutions at the stock level. Using IBP as a sorting variable, I find strong evidence that buying is more informative than selling for future returns since book-to-market ratios of portfolios with high IBP forecasts the market more accurately in contrast to stocks with low IBP. In classic asset-pricing time-series tests, book-to-market ratios corresponding to stocks in the highest institutional buying pressure quintile generate out-of-sample R^2s that are two to three times greater compared to stocks in the lowest quintile. In addition to statistical gains, I also uncover larger economic gains for book-to-market ratios of stocks with high IBP. The ratios of firms that institutions sell are less informative since they correlate highly with the general cross-section. On the other hand, ratios in firms with high institutional IBP have a lower magnitude and comove only moderately with the rest of the cross-section. All this undoubtedly points to a high degree of information heterogeneity across book-to-market ratios generated by institutional demand. In particular, selling is more liquidity driven while buying is more information driven.

To rationalize the finding that book-to-market is more informative only among firms for which institutions exhibit a higher degree of positive herding, I argue there are three important differences between institutional buying and selling. First, it is much harder to replicate book-to-market ratios of firms institutions buy as opposed to those of firms that institutions sell. I illustrate it through a comparison exercise between original IBP and pseudo portfolios which include firms outside institutions' trading sets that may look attractive based upon various accounting characteristics or through correlations with relevant cross-sectional factors. Accounting characteristics include variables such as earnings-to-price, dividend-to-price, size, asset growth, operating profitability, revenue growth, dividend growth, analyst disagreement and momentum. Regardless of the matching procedure considered, book-to-market ratios of portfolios with high IBP stocks have lower correlations with their counterparts from pseudo portfolios compared to book-to-market ratios in low IBP portfolios. Ratios of low IBP portfolios are easier to recreate from outside institutions' trading set. Second, the informational gap between book-to-market ratios is reinforced by a lower degree of disagreement between and within ratios of portfolios with high institutional buying pressure. The between disagreement is measured through time-series correlations of book-to-market ratios, sales-to-price and earnings-to-price. Within the higest IBP quintile, firms exhibit a stronger positive correlation among the three valuation measures as opposed to firms in the lowest IPB quintile. Since buying is more information driven, the comovement in the time-series among similar valuation ratios is stronger. The within disagreement is measured by the standard deviation of book-to-market ratios in IBP sorted portfolios. Within the higest IBP quintile, firms display a smaller standard deviation of book-to-market ratios compared to those in the lowest IBP quintile. These results strengthen the narrative that buying is more informative since in the time-series there is less disagreement among portfolios of stocks with high institutional demand. Third, by decomposing the variance of future returns, I show the source of predictive power in book-to-market ratios of firms with high IBP stems from its ability to anticipate the cash-flow news component more accurately than book-to-market ratios of firms with high selling pressure. Building on Pontiff and Schall (1998), aggregate book-to-market is shown to be a less accurate predictor for future returns as it lost its ability to anticipate the cash-flow component. The outcome is consistent with Park et al. (2019) and Choi (2021) who argue the loss of predictive power by book-to-market is rooted in book values needing various accounting changes to restore its ability to anticipate cash flows again. Their solutions reflect these adjustments. Park et al. (2019) create an intangible-adjusted ratio while Choi (2021) modify the original book-to-market to account for assets related to knowledge, organization capital and goodwill. This paper suggests an alternative approach to restoring informative book-to-market ratios: allowing institutional demand act as a filter. My results show there is a significant gain in the predictive ability of book-to-market ratios in the highest IBP quintile to anticipate cash flow news.

The contributions of this article add to two important strands of literature. The first one is in the area of the role of institutional trades for future stock returns. I establish a link between institutional demand and the aggregate stock market returns. By allowing institutional trading patterns dictate the composition of book-to-market ratios, I show a novel result that buying is more informative than selling for future returns through the effect of book-to-market ratios. The second contribution is on their predictive power for market outcomes. Welch and Goyal (2008) argue that book-to-market in its current form is irrelevant in forecasting returns. By using trades of institutions, I build a book-to-market ratio which is statistically more informative compared to one that is obtained through an average in an index like the Dow Jones or SP500. Since institutional buying is more informative and selling is driven by liquidity needs, I find that aggregating book-to-market ratios across stocks that are in high institutional demand generate better return forecasts compared to a simple market aggregated book-to-market.

The rest of the paper is organized as follows. Section 2 describes the data and empirical tests employed in the paper. Section 3 illustrates the empirical results and Section 4 concludes.

2. Data and Empirical Methodology

A. Data

Data on institutional common stock holdings is from the Thomson Reuters Institutional Holdings Database (s34 file). It includes Securities and Exchange Commission (SEC) quarterly filings of all institutional managers with investment accounts exceeding $100M in total market value. The corresponding SEC form is known as 13F. Form 13F not only reports current institutional holdings but also changes in holdings on a quarterly basis. It does not collect any information on short positions, cash or bond holdings. Accounting data on assets, earnings, revenues, dividends, book values, and any other relevant fundamentals are collected from the CRSP-Compustat database. Data on analyst expectations is collected from I/B/E/S. The institutional holdings database is merged with the CRSP-Compustat database by CUSIP number and any securities that do not match are dropped from the analysis. Stocks with prices below $5 and above $1000 are excluded. I also omit stocks with negative book-values.

Data on cross-sectional asset pricing factors is from Andrew Chen of the Federal Reserve Board. In Chen and Zimmermann (2021), authors publish a series of 161 predictors based on anomalies and factors collected from the asset pricing literature. Since the returns are computed at a monthly frequency, I convert it to quarterly units in order to match with the analysis in the merged CRSP-Compustat-13F database. Any other relevant macroeconomic indicators are extracted from the Federal Reserve Economic Database of the St. Louis Fed. All observations from the merged CRSP-Compustat-13F database and other related predictors used in the paper span a time series starting in Q1 1980 and ending in Q4 2020.

B. BM ratios and Institutional Buying Pressure

Book-to-market(B/M) ratios are calculated as the fraction of book value of equity over the market value of equity. The conventional way to compute an aggregate book-to-market measure is at the stock level or more generally at an index level. For example, Welch and Goyal (2008) estimate aggregate BM as the ratio of book value to market value for the Dow Jones Industrial Average. Since aggregate BM ratios incorporate some function of price in the denominator, it can identify departures from fundamentals. According to the mispricing view argued by Lewellen (2004), when valuations are high (low BM), stocks are overvalued with respect to their fundamentals and returns are expected to decrease. Conversely, when they are low (high

BM), expected returns are anticipated to increase. Hence, these ratios should be positively correlated with future returns.

In this paper I propose a slightly different approach to estimate aggregate book-to-market ratios. Instead of averaging them across a general index such as the SP500 or the Dow Jones, I focus on selecting a subsample of firms based on the demand of institutional investors from the 13F reports. My main motivation to look at valuation ratios through the lens of institutional investors is inspired by Koijen and Yogo (2019). Their study shows that latent demand, modeled as a residual of institutional demand after controlling for a wide variety of stock characteristics (e.g. dividends, earnings, book equity, leverage), is one of the most important drivers of returns explaining about 80% of their quarterly variance. Since institutional demand plays an important role in market fluctuations, I ask if average book-to-market ratios, aggregated at the institutional demand level instead of a market index level, change predictability implications in the time series of returns.

The 13F data, described in a previous section, contains information on asset holdings for any institutions with more than \$100M in assets under management (AUM). It also reports any buying and selling of stocks by these institutions on a quarterly basis. To calculate financial ratios based on institutional demand, I define a measure of price pressure for each stock represented by a discrete dummy $p_{i,j,t}$:

$$p_{i,j,t} = \begin{cases} 1, & \text{if } Change_{i,j,t} > 0 \\ 0, & Change_{i,j,t} < 0 \end{cases} \tag{1}$$

where $Change_{i,j,t}$ is the amount of shares of stock i traded by institution j in quarter t. I exclude stocks that were not traded (e.g. $Change_{i,j,t} = 0$) and compute the mean of $p_{i,j,t}$ at the asset level summarized across all institutions which executed trades in that particular quarter:

$$IBP_{i,t} = \frac{\sum_j p_{i,j,t}}{n_{i,t}} \tag{2}$$

$n_{i,t}$ is the number of institutional investors who traded stock i in quarter t. I define $IBP_{i,t}$ as institutional buying pressure in stock i at time t which essentially captures price pressure in either direction at the stock level. Statistically, it is the mean of a Bernoulli variable. The intuition behind $IBP_{i,t}$ is straightforward. For example, a value of 0.3, means that only three institutions bought stock i in quarter t and seven institutions sold. Hence, there would be more selling pressure. Conversely, a value of 0.7 means that seven institutions bought and only three sold and that creates more buying pressure. More generally, values over 0.5 coincide with an upward price pressure whereas values under 0.5 correspond to a downward price pressure.

IBP is a parsimonious way to capture price effects generated by changes in institutional demand at the stock level. To connect financial ratios and institutional demand, I sort stocks into portfolios based on IBP. A stock is required to be traded by at least 50 managers in each quarter before being placed in a portfolio. I split firms into quintiles which range from the lowest IBP with more selling pressure to the highest IBP which represent stocks with the largest buying pressure from institutions. The sorting exercise is performed on a rolling basis in each quarter t ensuring there is no look-ahead bias in the measure. I start the procedure in Q2 1980 and continue until the sample is exhausted in Q4 2020. Once stocks are sorted into portfolios I calculate the average book-to-market as follows:

$$\overline{BM_{p,t}} = \frac{\sum_i BM_{i,p,t}}{m_{p,t}} \tag{3}$$

where $p = 1, 2, 3, 4, 5$ represent portfolios of stocks sorted by IBP and $m_{p,t}$ represents the number of stocks in portfolio p in quarter t. I winsorize book-to-market observations at the 99% level. $\overline{BM_{p,t}}$ represents the average book-to-market of firms in portfolio p at quarter t.

C. Predictability Tests

Equipped with financial ratios on stocks with different degrees of institutional buying pressure, I investigate implications for predictability in the time series of returns for various market indices. The predictability tests are in line with what the literature has proposed in this area of asset pricing. Rapach and Zhou (2021) present a recent survey of predictability tests for factors both in-sample and out-of-sample. Also Goyal et al. (2021) evaluate the newest set of predictors proposed in the asset pricing literature and analyze their performance out-of-sample. My tests are in line with theirs.

I begin with a simple predictive model that regresses the one-quarter ahead excess returns $r_{i,t+1}$ on some factor x_t:

$$r_{i,t+1} = \beta_{i,0} + \beta_{i,1} x_t + \epsilon_{i,t+1} \tag{4}$$

where test assets are indexed by subscript $i = 1, 2$ and ϵ_{t+1} is the zero-mean, unpredictable error term. Test assets throughout the entire paper are CRSP value-weighted returns (i=1), and SP500 returns (i=2). Predictors in x_t are the average book-to-market ratios $\overline{BM_{p,t}}$. In addition, I also investigate the classic measure of $\overline{BM_{mkt,t}}$ averaged across the entire market. The sample is quarterly and spans from Q2 1980 to Q4 2020. For in-sample tests, I focus on two statistical measures: the Newey-West (NW) t-stats and the R^2. Goyal et al. (2021) argue that before proceeding with any out-of-sample procedures, one needs to ensure that predictors achieve a NW t-stat of at least 1.65 and a positive R^2. In-sample tests are usually the first hurdle in asset pricing. It essentially gives the green light for additional analyses. If there is evidence of a better and information-rich factor, Welch and Goyal (2008) show out-of-sample tests are the last crucial step in finding predictability since investors do not have access to all of the sample data.

The out-of-sample forecast of $r_{i,t+1}$ using only information until quarter t from (4) is:

$$\hat{r}_{i,t+1|t} = \hat{\beta}_{i,0,t} + \hat{\beta}_{i,1,t}x_t \tag{5}$$

where $\hat{\beta}_{i,0,t}$ and $\hat{\beta}_{i,1,t}$ are estimaed by OLS. To show robustness, I use both recursive and rolling windows in estimations. The recursive scheme consists of the following steps:

1. I train returns from beginning of sample to time t and forecast $\hat{r}_{t+1}$
2. I train returns from beginning of sample to time $t+1$ and forecast $\hat{r}_{t+2}$
3. I repeat step 2 until the sample is exhausted.

In this case, quarter t can be arbitrary depending on the start of out-of-sample predictions. In the empirical results, I exploit a variety of start dates.

The rolling scheme fixes windows of 60 quarters and is implemented as follows:

1. I train returns from time t to $t+60$ and forecast $\hat{r}_{t+61}$
2. I train returns from time $t+1$ to $t+61$ and forecast $\hat{r}_{t+62}$
3. I repeat step 2 until the sample is exhausted

where t=Q2 1980, Q3 1980, Q4 1980,... .

The idea behind out-of-sample tests in recursive and rolling windows is to compare two models: benchmark and competing. The competing forecast denoted by m is given in (5). The benchmark forecast denoted by subscript b is traditionally the arithmetic average of previous returns up until quarter t:

$$\overline{r}_{i,t+1} = \frac{1}{t}\sum_{s=1}^{t} r_s \tag{6}$$

To asses the strength of x_t versus a historical average benchmark I follow Campbell and Thompson (2008) and calculate the out-of-sample R^2 as follows:

$$R^2_{OOS} = 1 - \frac{MSFE_m}{MSFE_b} \tag{7}$$

where $MSFE_m$ represents the mean square forecast error of the competing model in (5) and $MSFE_b$ represents the mean square forecast error of the benchmark in (6). As suggested by Rapach and Zhou (2021), the statistical significance of R^2_{OOS} is evaluated using a Clark and West (2007) test. In addition to measuring statistical gains, I also compute the economic value of forecasts in relation to the benchmark. In doing so, I follow Rapach and Zhou (2021) and consider a mean-variance investor who allocates wealth between a risky and a risk-free asset:

$$R_{p,t+1} = (1 - \omega_{t+1|t})r_{f,t+1} + \omega_{t+1|t}r_{i,t+1} \tag{8}$$

The optimal allocation at time $t+1$ is given by:

$$\omega_{t+1|t} = \frac{1}{\gamma}\frac{\hat{r}_{i,t+1|t}}{\hat{\sigma}^2_{t+1|t}} \tag{9}$$

where γ is investors' relative risk aversion, $\hat{r}_{i,t+1|t}$ is the forecast. The investors' certainty equivalent is thus:

$$CER_p = \hat{\mu}_p - 0.5\gamma\hat{\sigma}_p \tag{10}$$

and the annual gain of the mean-variance investor who moves from the historical mean benchmark to the competing forecast is:

$$\text{Annual Gain} = (CER_m - CER_b) * 400 \tag{11}$$

where CER_m and CER_b are the certainty equivalents from the competing and benchmark predictions, respectively. Inference on annual gains from equation (11) is carried out through a bootstrapping procedure I develop based on McCracken and Valente (2018). Estimation details are described in detail in Appendix B.

Using both recursive and rolling windows, I evaluate statistical gains through the out-of-sample R^2 in (7) and economic gains through the difference in certainty equivalents from (11). In the case of recursive out-of-sample tests, I evaluate statistical end economic gains for three different time intervals. Since there is no priori regarding the start of the recursive scheme, I use three dates which are 10 years apart from each other: Q4 1990, Q4 2000 and Q4 2010. One-step ahead predictions are computed recursively starting these dates. The three out-of-sample evaluation periods are Q1 1991 to Q4 2020, Q1 2001 to Q4 2020 and Q1 2011 to Q4 2020.

D. Source of predictability

Pontiff and Schall (1998) argue that book-to-market ratio's ability to predict returns stems from book values being good proxies for future cash flows. I measure this ability through variance decomposition tests. Campbell (1991) and Campbell and Shiller (1988) show that unexpected future returns can be decomposed as follows:

$$r_{t+1} - E_t[r_{t+1}] = (E_{t+1} - E_t)\sum_{j=0}^{\infty} \rho^j \Delta d_{t+1+j} - (E_{t+1} - E_t)\sum_{j=1}^{\infty} \rho^j r_{t+1+j} \tag{12}$$

The first term of the equation represents news about future cash flows and the second term is news about discount rates; $\rho = \frac{1}{1+exp(\overline{d-p})}$ is the discount rate , $\overline{d-p}$ is the long term mean of $d_t - p_t$. I denote the discount rate news component as n^{DR} and the cash flows news component as n^{CF}; n^R represent the unexpected component of future returns.

$$\begin{aligned} \eta_{t+1}^{R} &= r_{t+1} - E_t[r_{t+1}] \\ \eta_{t+1}^{CF} &= (E_{t+1} - E_t)\sum_{j=0}^{\infty} \rho^j \Delta d_{t+1+j} \\ \eta_{t+1}^{DR} &= (E_{t+1} - E_t)\sum_{j=1}^{\infty} \rho^j r_{t+1+j} \end{aligned} \tag{13}$$

Combining (12) and (13) yields:

$$\eta_{t+1}^{R} = \eta_{t+1}^{CF} - \eta_{t+1}^{DR} \tag{14}$$

I use a VAR framework to extract the cash flow and discount rate news components:

$$Y_{t+1} = \Phi_1 Y_t + \epsilon_{t+1} \tag{15}$$

where Y_{t+1} represents a state vector, Φ_1 is a matrix of coefficients and ϵ_{t+1} is a vector of residuals with covariance matrix $E[\epsilon_{t+1}\epsilon_{t+1}'] = \Sigma$. In the spirit of Rapach et al. (2016), I present a classic approach and identify discount rate news (DR) first and back out the cash flow news (CF) component as a residual. Appendix A discusses issues associated with the classical method and presents an alternative approach based on Maio and Philip (2015) which identifies the cash flow news first and backs out the discount rate component as a residual. Specifically, I do so as a robustness check in view of the debate surrounding the reliability of discount and cash flow news estimates from the VAR. The results are qualitatively similar amongst the two approaches. Herein, I only discuss the classic method.

I denote a benchmark VAR state vector $Y_t = [r_{mkt,t} - r_{f,t}, \Delta D_t, \frac{D_t}{P_t}, cfnai_t]'$ where $r_{mkt,t} - r_{f,t}$ represents the excess market return, $\frac{D_t}{P_t}$ is the dividend-yield corresponding to the aggregate market index, ΔD_t is the quarterly change in dividends, and $cfnai_t$ is the Chicago Fed National Activity Index proposed by the Federal Reserve Bank of Chicago and described in more detail in Evans et al. (2002). The term-spread or default-spread are potential options to control for the stage of the business cycle, but a more effective way is to incorporate a larger set of macroeconomic controls. $Cfnai_t$ is a summary of 85 economic indicators and should represent a good proxy for the business cycle. The dividend-yield denoted by $\frac{D_t}{P_t}$ is computed as $\left(\frac{1+vwretd}{1+vwretx} - 1\right)$ where $vwretd$ is the value-weighted return including dividends from CRSP and $vwretx$ is the value-weighted return excluding distributions. The quarterly dividend growth is estimated as in Cochrane (2008) and denoted by $\Delta D_t = \frac{D_t/P_t}{D_{t-4}/P_{t-4}} * \frac{P_t}{P_{t-4}}$ where $\frac{P_t}{P_{t-4}}$ represents the quarterly return without distribution in CRSP ($vwretx$).

I identify the discount rate news component as follows:

$$\eta_{t+1}^{DR} = e_1' \sum_{j=1}^{\infty} \rho^j \Phi_1^j \epsilon_{t+1} = e_1' \rho \Phi_1 (I - \rho \Phi_1)^{-1} \epsilon_{t+1} \tag{16}$$

Innovations from the VAR forecast can be expressed as:

$$\eta_{t+1}^{R} = e_1' \epsilon_{t+1} \tag{17}$$

I back out the news about future cash flows from equation (14):

$$\eta_{t+1}^{CF} = \eta_{t+1}^{R} + \eta_{t+1}^{DR} \tag{18}$$

I use OLS to estimate equation (15) and extract estimates for $\hat{\Phi}_1$, $\hat{\epsilon}_{t+1}$ and $\hat{\rho} = \frac{1}{1+exp(\overline{\frac{D_t}{P_t}})}$ to calculate the sample counterparts of the discount rate news $\hat{\eta}_{t+1}^{DR}$, and the cash flow news $\hat{\eta}_{t+1}^{CF}$ in both the classic and alternative approaches of identification. Then, I analyze the predictive information of $\overline{BM_{p,t}}$ through predictive regressions of the individual components:

$$\begin{aligned} \hat{r}_{m,t+1}^{e} &= \alpha^E + \beta_p^E \overline{BM_{p,t}} + u_{t+1}^E \\ \hat{\eta}_{t+1}^{CF} &= \beta_p{}^{CF} \overline{BM_{p,t}} + u_{t+1}^{CF} \\ \hat{\eta}_{t+1}^{DR} &= \beta_p{}^{DR} \overline{BM_{p,t}} + u_{t+1}^{DR} \end{aligned} \tag{19}$$

where $\hat{r}_{m,t+1}^{e}$ are fitted values of one-quarter ahead predictions of returns on $\overline{BM_{p,t}}$. Estimates of $\beta_p{}^{CF}$ and $\beta_p{}^{DR}$ indicate which effect is significant: news about discount rates, or news about future cash flows.

3. Empirical Results

This section contains the main empirical results of the paper. I start by showing descriptive statistics of book-to-market ratios corresponding to portfolios sorted by IBP. Next, I investigate the predictability of these valuation ratios on future returns in a battery of in-sample and out-of-sample tests. In doing so, I provide estimates of statistical and economic gains through a classic framework using mean-variance investors. Lastly, my analysis aims to provide an empirical justification why book-to-market ratios of stocks with high IBP are more informative for the future performance of aggregate returns compared to alternative valuation ratios of stocks with low IBP or stocks in an index.

A. BM ratios in quintile portfolios

Demand of stocks by institutional investors produces an important amount of variation in returns as argued by Koijen and Yogo (2019). If IBP is a proxy measure of price pressures triggered by institutional demand, sorting stocks by IBP should generate important cross-sectional differences in valuation ratios, and along with that different levels of information embedded in them. I begin to show that by first providing a comparison of the average book-to-market ratios across portfolios sorted by institutional buying pressure. Portfolios sorts are executed using a rolling scheme from Q2 1980 to Q4 2020, and each quarter we place stocks in quintiles. I exclude stocks with prices less than \$5 and higher than \$1000, filter out firms with negative book-to-market ratios, and winsorize at the 99% level in each quarter of the rolling scheme. Portfolio 1 represents stocks that institutions sell the most, whereas portfolio 5 is comprised of stocks that institutions buy the most. The measure of institutional buying pressure from equation (2) is directional and indicates possible price changes. Within each portfolio, we use a simple average of the book-to-market ratios as in equation (3): for example, $\overline{BM_1}$ is the simple average of book-to-market ratios in portfolio 1, and $\overline{BM_5}$ is the simple average of book-to-market ratios in portfolio 5. I also compute $\overline{BM}_{mkt}$ which is the arithmetic mean of all firms in the cross section in the CRSP-COMPUSTAT-13F database.

There are alternative ways to calculate average book-to-market ratios for the aggregate market. For example, one could average these ratios at the level of an index such as the SP500 or the Dow Jones. Pontiff and Schall (1998) , Lewellen (2004), and Welch and Goyal (2008) perform predictability tests using only firms in the index. To obtain a more representative book-to-market ratio, I deviate from previous literature and incorporate a larger number of firms to compute $\overline{BM}_{mkt}$.

[Table 1 about here.]

These comparisons are provided in Table 1. Panel A illustrates a correlation table of book-to-market ratios, Panel B shows the difference in means of these ratios between portfolios, and Panel C indicates the Newey-West tstats associated with the differences from Panel B.

Panel A of Table 1 uncovers some interesting patterns among correlations that exist in these ratios. The first column shows pairwise correlations between $\overline{BM_{mkt}}$ and average book-to-market ratios of firms sorted by IBP. The degree of co-movement in these ratios is inversely related to IBP. While there is a strong positive correlation between book-to-market ratios of portfolio 1 and the general market, correlation coefficients with the general market decrease for portfolios of stocks with high institutional buying pressure. The correlation between $\overline{BM_{mkt}}$ and $\overline{BM_1}$, calculated from Q2 1980 to Q4 2020, is around 0.92. Conversely, between $\overline{BM_{mkt}}$ and $\overline{BM_5}$, I estimate a correlation of around 0.74 over the same period of time. This suggests institutions tend to sell stocks whose valuations level resemble closely that of the entire market and tend to buy stocks whose valuations have less co-movement with the market. Buying patterns of institutional investors vary more in relation to the market compared to selling patterns. This can be seen in the second column of Panel A, where the correlation between book-to-market ratios of portfolio 1 and portfolio 5 is only 0.66. Pairwise lower correlations of $\overline{BM_{mkt}}$, $\overline{BM_1}$ with $\overline{BM_5}$ suggest that valuation ratios aggregated across stocks that institutions buy more may be informationally different compared to those averaged at the index level or over stocks with more selling pressure. Table 1, Panel B, shows the difference in means between average book-to-market ratios of stocks sorted by IBP. These estimates are computed as the difference of valuation ratios between columns and rows. Panel C of Table 1 displays the NW t-stats associated with the estimates in Panel B. For example, the difference between $\overline{BM_{mkt}}$ and $\overline{BM_1}$ is -0.01 with a NW tstat of -1.24. The difference between $\overline{BM_{mkt}}$ and $\overline{BM_5}$ is -0.12 with a NW tstat of -9.34. There is no significant statistical difference between $\overline{BM_1}$ and $\overline{BM_{mkt}}$, and $\overline{BM_5}$ is significantly lower than $\overline{BM_{mkt}}$. Overall, firms with higher IBP have significantly lower book-to-market ratios compared to firms with low IBP or firms in the aggregate market. Moreover, these differences grow larger as IBP increases confirming a similar pattern with correlation coefficients in Panel A. This discrepancy in valuations between the market and a portfolio of stocks that institutions herd in, suggests this group of investors is more informed relative to the rest of market participants. According to the "smart investors" view of Gompers and Metrick (2001), institutions in this group ought to pick stocks with high expected returns. Fama and French (1995) defines this subset of stocks with lower book-to-market ratios as "growth" since they are more profitable and exhibit higher average returns on capital. Institutions certainly exhibit a strong preference for these "growth" type stocks.[2]

[2] I do not separate institutions based on orientation: value vs. growth. Certainly, depending on the goal, some institutions may buy more value while others buy growth. However, looking at the overall demand patterns over the length of the time-series, I find they buy firms with lower book-to-market ratios which are also known as "growth".

[Figure 1 about here.]

Figure 1 plot the time-series of $\overline{BM_{mkt}}$, $\overline{BM_1}$, and $\overline{BM_5}$ from Q2 1980 to Q4 2020. The overall positive co-movement among the three ratios is obvious, however $\overline{BM_1}$ tracks $\overline{BM_{mkt}}$ much closer compared to $\overline{BM_5}$. Examples of divergence between the three ratios can be spotted during several time periods. However, they are more prominent in early 1980s, 1990s, and 2000s which coincide with turning points in the business cycle. The graph just confirms the negative differences computed in Table 1. Average book-to-market ratios in portfolios with high IBP are lower. Buying patterns, compared to selling patterns, generate lower valuation ratios that are less correlated with the rest of the market. This implies that information trades may be more tilted towards firms with high IBP in portfolio 5 and reinforces institutions' preferences for lower book-to-market stocks.

B. BM Predictability Revisited

I proceed to investigate the predictive power of valuation ratios of stocks with high IBP in forecasts of future market returns. The correlation analysis of book-to-market ratios in portfolios with different degrees of IBP suggests significant deviations in valuations between the aggregate market and stocks institutional investors trade. The premise is that institutional demand merely acts like a filter in picking stocks with more desirable valuations. This, in turn, yields more informative book-to-market ratios that are likely better proxies for future market returns. I evaluate my hypothesis in the context of in-sample and out-of-sample tests from the empirical procedure of Section 2. In addition, I also investigate if this subset of valuation ratios generate economic gains measured in the context of a mean-variance portfolio allocating wealth between the market and a risk-free asset.

B.1. In-Sample

In the spirit of Goyal et al. (2021), I seek in-sample Newey-West t-stats above 1.65 and positive R^2s. These are minimum conditions to establish if predictors of the equity premium are statistically solid on a forward-looking basis.

[Table 2 about here.]

Table 2 shows the results of in-sample tests for book-to-market in (4). I use two proxies for market returns: value-weighted returns (VW) and the SP500. OLS regression coefficients, NW t-stats and R^2s span columns one through six where $\overline{BM_1}$ represents the average book-to-market ratio of firms with the lowest IBP and $\overline{BM_5}$ is the average book-to-market ratio of stocks with the highest IBP. $\overline{BM_{mkt}}$ is averaged across all available firms in the CRSP-COMPUSTAT-13F database. Column six in Table 2 shows the aggregated book-to-market ratio yields weak one-quarter ahead predictions. Across value-weighted and SP500 indices, estimated t-stats are below 1.65 and in-sample R^2s close to null. While the result is not novel, I confirm previous findings of Welch and Goyal (2008), Park et al. (2019) and Choi (2021). Book-to-market, as a proxy for value in the aggregate market, is not informative of future returns. There are a couple of reasons why this happens. First, there may be too many distortions in terms of BM ratios when considering such a large set of firms. There is a plethora of macroeconomic and firm-specific factors that affect demand for stocks on a quarterly basis among both institutional and retail investors. One such example is disagreement in terms of valuations which can severely distort book-to-market ratios by skewing the denominator and increasing variance. Another example is information trades. Generally, institutions are considered "smart" in relation the rest of the market, hence their buying patterns should produce a more informative book-to-market that is more representative of anticipated changes in future returns.

The ability of institutional demand to act as a filter and generate more informative valuation measures can be seen in the predictive power of $\overline{BM_5}$ in column five of Table 2. $\overline{BM_5}$ performs better in forecasting one quarter ahead returns comapred to $\overline{BM}_{mkt}$. In the case of value-weighted returns, the NW t-stat corresponding to $\overline{BM_5}$ is 3.29 and the R^2s is 0.06. Explanatory power for one-quarter ahead returns is strong for book-to-market ratios for stocks with high IBP. For the SP500 index, $\overline{BM_5}$ also forecasts better than $\overline{BM}_{mkt}$ with a NW t-stat of 2.93, and an R^2 of 0.05. For SP500, only regression coefficients associated with $\overline{BM_5}$ are above 1.65. To asses statistical differences between $\overline{BM_5}$ and $\overline{BM}_{mkt}$ I apply a Wald test. Column eight of Table 2 illustrates p-values associated with the Wald test. The difference of 0.15 is associated with a p-value of 0.04 indicating $\overline{BM_5}$ is a stronger predictor compared to $\overline{BM}_{mkt}$.

There are also differences in the predictive power of book-to-market ratios triggered by buying and selling patterns of institutions. These differences can be seen in OLS betas of $\overline{BM_1}$ and $\overline{BM_5}$ and are consistent across the value-weighted and SP500 indices. Generally, we observe higher betas and R^2s for book to market ratios in portfolios of stocks with higher IBP compared to portfolios comprised of firms with lower IBP. Statistical differences between $\overline{BM_1}$ and $\overline{BM_5}$ are confirmed by Wald tests in column seven of Table 2. Average book-to-market ratios of firms which display stronger buying pressure from institutions tend to explain more of the time series variation in market returns. $\hat{\beta}_{VW}$ and $\hat{\beta}_{SP500}$ corresponding to $\overline{BM_5}$ are significantly higher compared to corresponding coefficients of $\overline{BM_1}$.

[Figure 2 about here.]

As an alternative to using all observations, I break the sample in intervals and estimate rolling fixed window one-quarter ahead regressions of (4). To compare the in-sample statistical fits of $\overline{BM_{p,t}}$ in these intervals, I plot R^2s in Figure 2. The top panel corresponds to value-weighted returns and the bottom panel represents the SP500. Windows are fixed in intervals of 60 observations. Dates on the horizontal axis in Figure 2 correspond to the ending quarter of each rolling window sample. For example, the R^2 of 0.07 corresponding to $\overline{BM_5}$ in Q2 1995 is estimated on the previous 60 observations that originate in Q2 1980. In general, patterns of R^2s confirm findings from Table 2. Book-to-market ratios of firms that institutions herd in, explain a larger time-series variation of market returns compared to $\overline{BM_1}$ or $\overline{BM_{mkt}}$. Throughout the sample, the black solid line corresponding to R^2s of $\overline{BM_5}$ are predominantly above the blue ($\overline{BM_{mkt}}$) and red ($\overline{BM_1}$) solid lines. One big exception is the Financial Crisis in 2008 where valuation ratios across the board performed poorly. In the case of the SP500 (bottom panel of Figure 2), with the exception of the Financial Crisis, $\overline{BM_5}$ generates positive R^2s all the way from the 2000s till the end of the sample in 2020. Conversely, R^2s of $\overline{BM_{mkt}}$ and $\overline{BM_1}$ stay much closer to the zero line. For the most part of early 2000 till 2010, $\overline{BM_{mkt}}$ actually generates negative R^2s and only rises slightly after the end of the Financial Crisis.

These results provide preliminary evidence that institutional demand is an important filter for book-to-market ratios as predictors of future returns. Next, I continue to reinforce my findings with a set of out-of-sample tests where I evaluate statistical gains through out-of-sample R^2s and economic gains measured by the ability of a mean-variance investor to allocate wealth based on forecasts using book-to-market ratios.

B.2. BM Ratio - OOS Statistical Gains

The procedure behind out-of-sample tests is outlined in detail in section 2. In recursive forecasts, I select three forecast origins which are 10 years apart. Evaluation of out-of-sample competing models of $\overline{BM_{p,t}}$ in relation to the benchmark constant prediction is provided for three distinct periods: Q1 1991 to Q4 2020, Q1 2001 to Q4 2020 and Q1 2011 to Q4 2020.

[Table 3 about here.]

Table 3 illustrates the out-of-sample R^2s for the three evaluation periods using value-weighted and SP500 indices as test assets. Columns one, two and three correspond to forecast start dates of "1990-12-01", "1990-12-01" and "1990-12-01". For value-weighted returns, financial ratios of stocks associated with more institutional buying pressure such as $\overline{BM_4}$ and $\overline{BM_5}$ predict a significantly higher amount of variation in one-quarter ahead returns compared to $\overline{BM_{mkt}}$. The out-of-sample R^2s for the three forecast origins associated with $\overline{BM_5}$ are 0.02, 0.09 and 0.12, respectively. In contrast, out-of-sample R^2s associated with $\overline{BM_{mkt}}$ are 0.01, 0.03 and 0.04. Differences between these values are large. Book-to-market ratios of portfolios with high IBP anticipate one-quarter ahead returns more accurately by factors of two and three compared to a more traditional book-to-market. In the case of SP500 , the results are qualitatively similar. $\overline{BM_5}$ is a stronger factor than $\overline{BM_{mkt}}$ for all start dates. Out-of-sample R^2s of $\overline{BM_{mkt}}$ and $\overline{BM_1}$ are comparable for value-weighted and SP500 returns. Numerically, this is not surprising given their correlation in Table 1. Statistically, the lower pairwise correlations between $\overline{BM_5}$ and $\overline{BM_{mkt}}$ or $\overline{BM_5}$ and $\overline{BM_1}$, suggest that patterns of institutional buying are more informative than selling. In addition, since institutions sell stocks that are closer in valuation to the aggregate market, $\overline{BM_{mkt}}$ remains a weaker factor relative to $\overline{BM_5}$. The first set of out-sample tests reinforce the in-sample findings that the bulk of information in valuation ratios useful to forecast future returns is concentrated in stocks with high IBP.

[Figure 3 about here.]

Figure 3 plots R^2s for additional forecast start dates. Instead of limiting tests to Q4 1990, Q4 2000 and Q4 2010, I also consider all years in between and show it graphically. The top panel illustrates results for value-weighted returns, while the bottom panel is for SP500. Lines represent out-of-sample R^2s for $\overline{BM_{p,t}}$ evaluated on forecast samples starting at different dates between Q4 1990 and Q4 2010. $\overline{BM_5}$ dominates the rest of book-to-market ratios for all other quarters. In general, buying patterns are more informative than selling patterns for the predictive content of financial ratios. R^2s of $\overline{BM_1}$ and $\overline{BM_{mkt}}$ are below those of $\overline{BM_5}$.

[Table 4 about here.]

As a final robustness check for statistical gains, I evaluate out-of-sample results in the context of rolling fixed window regressions. Windows are fixed at 60 quarters. The details of the rolling methodology are also described in section 2. Table 4 depict R^2s for value-weighted and SP500 returns in this alternative scheme. R^2s corresponding to $\overline{BM_4}$, $\overline{BM_5}$ are higher and statistically different than zero as shown by a Clark and West (2007) test. Irrespective of the preferred forecasting scheme, strong institutional buying generates more informative valuation ratios for future market returns.

B.3. Economic Gains

Rapach and Zhou (2021) argue that statistical value of forecasts do not necessarily imply economic value. A statistical accurate forecast is essential because it gives investors necessary information to decide how to allocate wealth between risky and risk-free assets (stocks vs. bonds). Because of this, I also compare the economic value of $\overline{BM_{p,t}}$ in out-of-sample forecast tests. Annualized economic gains estimations from equation (11) cover the same three evaluation periods as the out-of-sample R^2s. Test assets are the CRSP value-weighted index and the SP500.

[Table 5 about here.]

Table 5 displays annualized economic gains for three different out-of-sample forecast start dates. Statistically significant values from the bootstrap algorithm(Appendix B) are displayed in bold. In the case of value-weighted returns, economic gains resulting from forecasts using $\overline{BM_4}$ and $\overline{BM_5}$ are higher compared to $\overline{BM_1}$ or $\overline{BM_{mkt}}$. In the case of $\overline{BM_5}$, these are 42, 180 and 186 basis points at forecast origination dates December 1990, 2000 and 2010, respectively. In contrast, for $\overline{BM_1}$, annualized utility gains are 36, 94, and 74 basis points over the same forecast periods. Estimates of utility gains from predictions using $\overline{BM_{mkt}}$ are 32, 82 and 75 basis points. Hence, as in the case of out-of-sample R^2s, economic values of forecasts from $\overline{BM_{mkt}}$ are more comparable to $\overline{BM_1}$ as opposed to $\overline{BM_5}$ or $\overline{BM_4}$. The second panel presents results for the SP500. Statistical patterns match qualitatively. For three different forecast start dates, utility gains from $\overline{BM_5}(\overline{BM_1})$ are 92(17), 224(51), and 235(51) basis points.

[Figure 4 about here.]

Figure 4 plots all annualized economic gains for $\overline{BM_{p,t}}$ at all forecast start dates between December 1990 and December 2010. In both cases of value-weighted and SP500 returns, lines corresponding to $\overline{BM_4}$ and $\overline{BM_5}$ are higher compared to the rest of book-to-market ratios for a large proportion of forecast origination quarters. In contrast, $\overline{BM_1}$ and $\overline{BM_{mkt}}$ yield similar economic gains in the time series. As in the case of out-of-sample R^2s, utility gains from return forecasts suggest that institutional trading distorts the information content of valuation ratios. Book-to-market ratios of stocks with high IBP are of more economic value compared to those of stocks with low IBP. Institutional buying is more informative than selling because it generates higher statistical and economic content in valuation ratios to anticipate future returns more accurately. Additionally, book-to-market ratios of stocks with high IBP produce more economically informative forecasts compared to book-to-market ratios averaged in the entire market or an index.

C. Institutional buying is more informative than selling

Why is $\overline{BM_5}$ more informative than $\overline{BM_1}$ or $\overline{BM_{mkt}}$ for future returns? Both in-sample and out-of-sample regressions indicate significant statistical and economic gains for valuation ratios of firms associated with more institutional buying pressure. In contrast, firms that are being sold more heavily by institutions tend to have book-to-market ratios closer to the aggregate market and are less informative for future returns. In terms of valuation ratios, institutional buying is more informative than selling. I attempt to explain my findings with some insightful empirical observations.

C.1. Pseudo Portfolios

The first one is connected to the ease of replicating buying and selling patterns of institutional investors based on accounting information or cross-sectional asset pricing factors. I show it is more challenging to build portfolios of companies whose valuations replicate book-to-market ratios associated with buying patterns of institutions as opposed to selling. I refer to these as pseudo portfolios and construct them by searching for subsets of companies that may look like a buy or a sell to institutions, but are actually outside their trading sets.[3] In practice, it is very difficult to know the exact reasons why these classes of traders buy or sell a particular stock. While the reasons remain largely unknown, it is nevertheless possible to paint a general picture of the average accounting characteristics institutions seek when making trades. To form some predictions with respect to the average profile of firms that may look appealing to them, I sort portfolios by institutional buying pressure and evaluate them on the following factors: EP (earnings-to-price), DP (dividend-to-price), SZ (market capitalization), INV (yearly growth in total assets), OP (yearly growth in operating income), REV (yearly growth in revenues), DIV (yearly growth in dividends), ANN DIS (analyst disagreement in earnings per share), and MOM (prior yearly returns).[4]

[Figure 5 about here.]

Figure 5 reveals general characteristics of trading patterns by institutional investors. These characteristics are obtained by time-series averages in portfolios sorted by institutional buying pressure. Typically, they buy companies with higher earnings-to-price ratios, lower dividend-to-price ratios, smaller size, high investment, high profitability, high revenue growth, high dividend growth, low analyst disagreement and high prior one-year returns. To show that institutional buying is more information driven compared to selling, I build pseudo portfolios from the CRSP-COMPUSTAT-13F database that match trading patterns of 13F companies on characteristics illustrated in Figure 5. The procedure is straightforward and is applied on a rolling basis:

1. In each quarter, I identify firms traded by institutional investors through 13F and exclude them from CRSP-COMPUSTAT database.
2. Using the subset of firms not traded by institutions from the previous step, in each quarter and for each characteristic from Figure 5, I assign scores ranging from one to five : one represents a "hypothetical" strong sell by an institutional investor whereas five is a "hypothetical" strong buy (this matches the IBP portfolios). To build pseudo portfolios, I take the mode of all the scores. For example, a firm may receive the following scores based on EP(3), DP(2), SZ(1), INV(5), OP(4), REV(4), DIV(5), ANN

[3] I refer to them as pseudo portfolios because they include firms outside of institutions' information sets but are similar to what institutions trade based on accounting characteristics or cross-sectional factors. The matching is done in both ways.

[4] These factors are known to drive the highest amount of variation in the cross-section of US stocks.

DIS(5), MOM(5). Given the mode of these scores is five and that this particular stock was not traded by any institution but appears to be a "hypothetical" strong buy, I place it in pseudo portfolio five. Each stock in the subset is assigned to a pseudo portfolio in this manner.

3. In each quarter, I calculate average book-to-market ratios across pseudo portfolios ($\overline{BM_{p,t}}^{pseudo}$) to compare to the original $\overline{BM_{p,t}}$

Alternatively, these pseudo portfolios can also be built from a matching procedure that accounts for a larger number of cross-sectional factors. Chen and Zimmermann (2021) compile a large list of 205 predictors that are built on certain characteristics which are important for the cross-section. I use the list as a basis for constructing pseudo portfolios sourcing a larger information set. Since some of the long-short portfolios have missing data, I exclude them from my analysis.[5] The rest are used to identify possible subsets that could explain the difference between returns of stocks with high IBP versus stocks with low IBP. For each of these portfolios sorted by institutional buying pressure, I calculate value- and equally-weighted returns and then regress them individually on each factor in order to identify which ones matter.

[Figure 6 about here.]

Panel A of Figure 6 shows all Chen and Zimmermann (2021) factors with Newey-West t-stats above 2.5, while Panel B plots all factors with Newey-West t-stats above 1.96. In Panel A, only a handful of factors emerge as highly significant. These are *DebtIssuance* of Spiess and Affleck-Graves (1999), *FirmAge* of Barry and Brown (1984), *IndRetBig* of Hou (2007) and *Herf*, *HerfAsset*, *HerfBE* of Hou and Robinson (2006). Using the characteristics of Chen and Zimmermann (2021) to rank replicated factors, all variables that are found significant in Panel A range from "Likely Predictors" to "Clear Predictors". In other words, these few factors are more likely to explain the differential of value- and equally-weighted returns between high and low IBP portfolios. The emergence of these predictors as possible candidates for explaining the IBP return differential has further implications for the general picture of institutional demand of stocks and adds valuable information to the average trade profile of institutions built from characteristics in Figure 5. For example, *Herf*, *HerfAsset* and *HerfBE* measure the returns of firms by level of industry concentration using the Herfindahl index based on sales, asset and book-value of equity. An increase in the long-short return based on industry concentration is associated with a larger IBP return differential, suggesting that institutional

[5] I excluded the following factors: Activism1, Activism2, AgeIIPO, betaVIX, ChangeInRecommendation, ConsRecomm, CredRatDG, DelBreadth, DelDRC, DownRecomm, EarningsForecastDisparity, EarningsStreak, ExclExp, fgr5yrLag, FR, Governance, iomomcust, iomomsupp, OptionVolume1, OptionVolume2, PctTotAcc , PredictedFE, ProbInformedTrading, RDcap, skew1 , SmileSlope, UpRecomm, AnalystRevision, BetaLiquidityPS, BidAskSpread, ChForecastAccrual, ChNAnalyst, CitationsRD, FEPS, ForecastDispersion, Mom6mJunk, PatentsRD, PriceDelayRsq, PriceDelaySlope, PriceDelayTstat, retConglomerate, REV6, RIODisp, CoskewACX, DivOmit. For a complete definition of these factors, I refer the reader to the open source cross-sectional asset pricing of Andrew Chen and Tom Zimmermann at https://www.federalreserve.gov/econres/feds/open-source-cross-sectional-asset-pricing.htm

investors are more likely to exhibit strong buying pressure on portfolios of firms in lower concentration industries. Likewise, *IndRetBig* is a significant factor with a positive beta on the IBP return differential; this implies that more institutional buying pressure may occur in the largest companies by market value in the Fama-French 48 industries. Furthermore, the positive effects of *FirmAge* and *DebtIssuance* on the IBP return differential indicate strong potential institutional buys for firms that issue less straight or convertible debt, or have been listed for shorter period of times as measured by their CRSP coverage.

Using observations from Panel A of Figure 6, I apply a similar rolling procedure to build pseudo portfolios by assigning stocks in possible "buy" or "sell" categories in the subsequent steps:[6]

1. I fix a 10 year (40 observations in quarterly) rolling window and start the procedure in Q1 1980.
2. Over the length of the rolling window, I identify firms traded by institutional investors through 13F and exclude them from CRSP-COMPUSTAT database.
3. Using the subset of firms not traded by institutions from the previous step, I run separate regressions of quarterly returns on each cross-sectional factor from Figure 6. Based on time-series regression betas for each factor, stocks are placed into quintiles. For each stock, depending on the portfolio assignment, there is a score associated with that individual factor. One is a "hypothetical" strong sell by an institutional investor whereas five is a "hypothetical" strong buy. To assign a final ranking to the stock, I take the mode of all the scores. For example, a firm may receive the following scores based on *Herf*(5), *HerfAsset*(5), *HerfBE*(5), *IndRetBig*(4), *FirmAge*(2), *DebtIssuance*(1). Given the mode of these scores is five, I would place this particular stock in pseudo portfolio five. Each stock in the subset is assigned to a pseudo portfolio in this manner.
4. At the end of each rolling period, I calculate average book-to-market ratios across pseudo portfolios ($\overline{BM_{p,t}}^{pseudo}$) to compare to the original $\overline{BM_{p,t}}$

[Table 6 about here.]

The comparison between book-to-market ratios are presented in the form of a correlation analysis in Table 6. Panel A illustrates results for pseudo portfolios built using accounting characteristics in Figure 5. Average book-to-market correlations between IBP and pseudo portfolios are higher for $\overline{BM_1}^{pseudo}$ or $\overline{BM_2}^{pseudo}$ compared to $\overline{BM_4}^{pseudo}$ or $\overline{BM_5}^{pseudo}$. For example, the correlation between $\overline{BM_1}$ and $\overline{BM_1}^{pseudo}$ is 0.94. By comparison, the same correlation between $\overline{BM_5}$ and $\overline{BM_5}^{pseudo}$ is much lower at 0.63. There is a general

[6]Panel B of Figure 6 illustrates a larger set of predictors that can explain the IBP return differential with Newey-West t-stats above 1.96. Appendix C presents the correlation results of book-to-market ratios between original IBP and pseudo portfolios built from a larger number of cross-sectional asset pricing factors.

decreasing pattern in the interdependence of $\overline{BM_{p,t}}$ with $\overline{BM_{p,t}}^{pseudo}$ suggesting the matching procedure is more successful in replicating average book-to-market ratios of stocks with high institutional selling pressure as opposed to stocks with high institutional buying pressure. In the case of stocks with high IBP, the matching procedure does not do well. This suggests that buying of institutions is driven much more by private information compared to selling which appears to be liquidity driven. There must be additional latent factors institutions consider when buying stocks since a parsimonious matching on financial ratios does not suffice. This is consistent with Koijen and Yogo (2019) who show that a large proportion of return variance is driven by latent demand. On the other hand, selling patterns are more straightforward to replicate with respect to valuation ratios. The implication of a low correlation between $\overline{BM_5}^{pseudo}$ and $\overline{BM_5}$ in the context of return predictability is that $\overline{BM_5}^{pseudo}$ does not forecast future returns while $\overline{BM_5}$ remains an important factor. $\overline{BM_1}^{pseudo}$ is also weak predictor due to its high co-movement with $\overline{BM_1}$ and $\overline{BM_{mkt}}$.

Panel B of Table 6 shows correlation outcomes between book-to-market ratios for pseudo portfolios built from asset pricing factors in Figure 6. I confirm a similar pattern found in Panel A. The correlation between $\overline{BM_1}$ and $\overline{BM_1}^{pseudo}$ is 0.80. In contrast, the correlation between $\overline{BM_5}$ and $\overline{BM_5}^{pseudo}$ is 0.58. This gap reinforces the narrative that it is more difficult to replicate book-to-market portfolios with high institutional buying pressure, which may explain why ratios associated with this subset of firms are more informative for future market returns.

The results stay robust even when I match portfolios on a smaller subset of characteristics or on statistically weaker cross-sectional factors. Appendix C shows correlation patterns in an alternative scheme with no alterations of main findings.

C.2. Less disagreement between and within valuation ratios

A second argument that reinforces institutional buying being more informative than selling is associated with the level of disagreement between and within valuation ratios in portfolios with different IBP. I measure the between disagreement in terms of stock fundamentals by estimating time-series of correlations at the portfolio level between average book-to-market and two alternative ratios: sales-to-price and earnings-to-price. These two can be viewed as substitutes to book-to-market since they signal a stock becoming under- or over-valued. For the within disagreement, I build time-series of standard deviation estimates from average book-to-market ratios in each portfolio sorted by institutional buying pressure. A lower level of disagreement within book-to-market and between book-to-market and alternative fundamental valuation metrics should indicate more consensus among institutional demand for a particular stock and more informative book-to-market ratios.

[Figure 7 about here.]

The top two graphs of Figure 7 illustrate the between disagreement measured by the correlations of book-to-market with sales-to-price and earnings-to-price in portfolios with high degree of institutional buying and selling. Portfolios sorts are consistent with the procedure explained in more detail in section 2. For simplicity, I show the correlation estimates for portfolio one and five. Portfolio one (red line) corresponds to stocks that are sold the most (lowest IBP), while portfolio five (blue line) are stocks that have the highest IBP. I also include a time-series of correlation estimates between the same fundamental ratios for the entire market (black line) as proxied by all firms available in the cross section in each quarter. The first graph of Figure 7 shows a higher positive correlation between book-to-market and sales-to-price among stocks with high IBP compared to stocks with low IBP or the aggregate market. Similarly, the second graph reveals a higher positive correlation between book-to-market and earnings-to-price in portfolio five contrary to portfolio one or the market. This pinpoints to a higher level of agreement between fundamental ratios corresponding to portfolios of stocks where institutions generate more buying pressure. The third graph of Figure 7 displays the within disagreement by plotting time-series estimates of cross-sectional standard deviations of book-to-market ratios in portfolio one, portfolio five and the aggregate market. Standard deviation estimates of book-to-market ratios in portfolio five are lower compared to portfolio one or the market. The level of within agreement among book-to-market ratios of stocks with higher institutional buying pressure is larger.

[Table 7 about here.]

Table 7 computes the time-series means of correlation and standard deviation estimates from Figure 7. Between book-to-market and sales-to-price, the average correlation is around 0.13 for portfolio one, 0.21 for portfolio 5 and 0.04 for the market. In general, the correlation increases for stocks with higher IBP. Differences of average correlations between portfolio five and portfolio one, portfolio five and the market are 0.07 and 0.17, respectively. In both cases, Newey-West test statistics confirm a higher correlation in the time series between book-to-market and sales to price for stocks with more institutional buying pressure. Between book-to-market and earnings-to-price, the average correlation is around 0.3 for portfolio one, 0.42 for portfolio 5 and 0.32 for the market. Differences of 0.12 and 0.1 along with their significant Newey-West test statistics confirm larger correlation in the time series between book-to-market and earnings-to-price for portfolio five compared to portfolio one or the aggregate market. As far as the within disagreement, average standard deviation estimates for portfolio five are significantly lower compared to portfolio one or the market.

Overall, the patterns are clear: there is a higher degree of between agreement among book-to-market ratios and its counterparts (e.g. sales-to-price and earnings-to-price), and there is more within agreement in the ratio itself. The graphical analysis of Figure 7 along with numerical estimates from Table 7 serve as additional important evidence that institutional buying is more informative than selling and further explains why $\overline{BM_5}$ is more informative than $\overline{BM_1}$ or $\overline{BM_{mkt}}$.

C.3. CF News

As a third supporting argument, I show institutional buying is more informative than selling for future market returns because book-to-market ratios of firms in the highest quintile of IBP are more informative about future cash-flows compared to firms in the lowest quintile. Pontiff and Schall (1998) argue book-to-markets' predictive ability for returns is the outcome of book values proxying for future cash flows. Overall, the variance decomposition results are consistent with the findings of Pontiff and Schall (1998) since the bulk of the regression coefficient estimate corresponding to each $\overline{BM_{p,t}}$ can be attributed to news about cash flows. However, there are significant differences in the magnitude of cash flow news between book-to-market ratios induced by institutional demand. Using IBP as a sorting variable, I decompose the effect of each $\overline{BM_{p,t}}$ and find that $\overline{BM_5}$ is the most accurate book-to-market ratio predictor for future market returns since it displays a statistically significant cash-flow component while the competing ratios do not.

[Table 8 about here.]

Table 8 illustrates in detail the decomposition exercise whose methodology is described in detail in section 2. I decompose each OLS beta corresponding to $\overline{BM_{p,t}}$ into a cash-flow component $\hat{\beta}_{CF}$, a discount rate component $\hat{\beta}_{DR}$ and an expected return component $\hat{\beta}_{E}$. Test assets are the value-weighted index (top panel) and the SP500 (bottom panel). For $\overline{BM_5}$, $\hat{\beta}^{OLS}_{VW}$ is 0.21 for value-weighted returns. This figure matches quantitatively the estimate from Table 2. When breaking it down between the two components, I find that cash-flow $\hat{\beta}_{CF}$ is 0.15 with a NW t-stat of 2.30 and discount rate $\hat{\beta}_{DR}$ is 0.01 with a NW t-stat of 1.46. The power of $\overline{BM_5}$ to anticipate future returns is due to its statistically ability to predict cash-flow news in contrast to discount rate news. When contrasting institutional buying to selling, I find once more that selling is less informative as $\overline{BM_1}$ is a weak factor for both variance components of future unexpected returns. In the case of value-weighted returns, $\hat{\beta}^{OLS}_{VW}$ for $\overline{BM_1}$ is 0.07 with $\hat{\beta}_{CF}$ around 0.03 and $\hat{\beta}_{DR}$ close to null. None of the estimates are statistically significant. For the aggregate book-to-market ratio $\overline{BM_{mkt}}$, I find similar news effects as in $\overline{BM_1}$. This is not surprising due to their high correlation of 0.92. In the case of $\overline{BM_{mkt}}$, neither $\hat{\beta}_{CF}$ nor $\hat{\beta}_{DR}$ are significant. For SP500, the news effects estimates are qualitatively similar as $\overline{BM_5}$ predicts the cash flow component better than $\overline{BM_1}$ or $\overline{BM_{mkt}}$. Overall, the variance decomposition procedure of book-to-market predictors offers important supplemental evidence that institutional buying is more informative than selling for future market returns. Furthermore, I confirm the source of predictive power for $\overline{BM_5}$ to remain the cash-flow news component. The ability of $\overline{BM_{mkt}}$ to predict future returns has reduced significantly because it is a weaker proxy for future cash flows.[7] In contrast, $\overline{BM_5}$ is a stronger proxy for future cash flows in view of the fact that high institutional buying pressure acts like an information filter for the cross-section. As shown in Table 6, it is very difficult to recreate $\overline{BM_5}$ from the rest of the cross-section, and that may explain why $\overline{BM_5}$ has more desirable predictive properties for future returns compared to $\overline{BM_{mkt}}$ or $\overline{BM_1}$.

In Appendix A, I discuss the variance decomposition results in an alternative scheme where cash-flows news is identified first and discount-rate news is backed as a residual. Estimates of $\hat{\beta}^{CF}$, $\hat{\beta}^{DR}$ and $\hat{\beta}^{E}$ from this alternative approach can be found in Table A1. Furthermore, Table A2 presents the results of an alternative specification for state vector Y_t from (15) where $cfnai_t$ is replaced with both the term- and default-spread. There are no significant differences arising from robustness tests in the variance-decomposition. Thus, in the interest of space, the discussion of these tables is available in Appendix A.

[7]See Park et al. (2019) and Choi (2021).

4. Conclusion

I establish a new link between institutional trades and stock returns through book-to-market effects. Buying and selling patterns of institutions create different levels of information embedded in these ratios and produce new implications for return predictability. Using institutional buying pressure as an information filter in the cross-section, I show that book-to-market ratios associated with stocks that see the largest increases in institutional involvement (highest IBP) are more informative than those of stocks that institutions divest off (lowest IBP). The difference in information across book-to-market ratios generated by IBP in the context of future market returns are assessed in asset pricing tests through out-of-sample R^2s and economic gains. I argue that buying of institutions produces a more informative book-to-market for a handful of reasons. Buying is driven much more by private information in contrast to selling since it is indisputably more challenging to replicate these ratios in stocks with high IBP as opposed to stocks with low IBP. Following a strategy to generate pseudo portfolios of institutions based on cross-sectional factors that may explain IBP, I am able to reproduce much higher correlations among book-to-market ratios for stocks in the lowest IBP quintile. Additional evidence that reinforces the narrative is also presented in the form of a between disagreement across comparable valuation ratios spanning the cross-section of IBP portfolios. There is a higher positive correlation between book-to-market, sales-to-price and earnings-to-price for high IBP stocks. This suggests there is a lower level of disagreement in terms of fundamental valuations for this group of stocks. Finally, I argue that IBP generates a better book-to-market because it restores its ability to forecast the cash-flow news component of future market returns.

The informational role of institutions in asset prices remains an important topic for this literature. One possible extension of this work is to obtain a better understanding on the type information that makes book-to-market of high IBP stocks more informative of the cash-flow news component of future returns. On one hand it could be a private signal which cannot be observed directly from their trading patterns. On the other hand, as suggested by Park et al. (2019) and Choi (2021) , it could be these book values of stocks with high IBP are better proxies of future cash-flows since they reflect more accurately the growth of intangible assets, knowledge , goodwill or organizational capital on their balance sheet. Disentangling the source of information in book-to-market ratios remains a topic for future research.

Figure 1: Average BM ratios sorted by IBP

Note: This figure illustrates a quarterly time-series of average book-to-market ratios in firms split by IBP. Details of the sorting procedure can be found in section 2. The blue line represents average BM ratios in the highest IBP quintile ($\overline{BM_5}$). The red line represents average BM ratios in the lowest IBP quintile ($\overline{BM_1}$). The black line shows an average BM ratio of the full cross-section ($\overline{BM_{mkt}}$). The time series spans from Q2 1980 to Q4 2020.

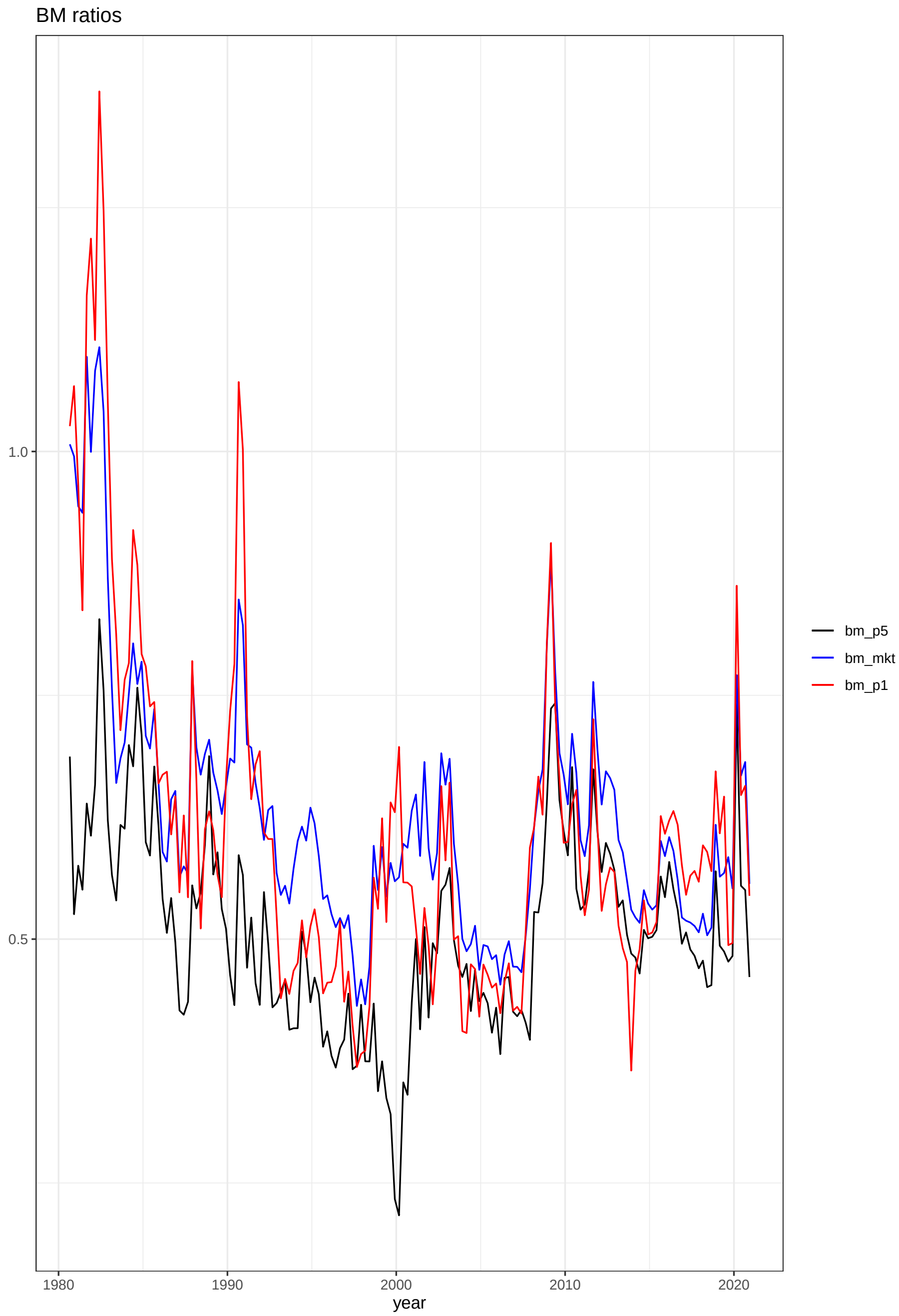

Figure 2: Rolling forecasts R^2 - in sample

Note: The figure illustrates R^2s obtained from rolling fixed window (60 observations) in-sample forecasts of value-weighted and SP500 returns using various specifications of $\overline{BM_p}$. Dates on the horizontal axis correspond to the ending quarter of each rolling window sample. The top panel shows results for value-weighted returns and the bottom panel shows results for SP500. The black line corresponds to R^2s obtained from regressing one quarter ahead returns on average book-to-market of stocks in the highest IBP quintile ($\overline{BM_5}$). The blue line plots R^2s obtained from regressing one quarter ahead returns on average book-to-market of stocks in the lowest IBP quintile ($\overline{BM_1}$). The red line shows R^2s obtained from regressing one quarter ahead returns on average book-to-market of stocks in the full cross-section ($\overline{BM_{mkt}}$).

Figure 3: Recursive forecasts R^2 - out of sample

Note: This figure illustrates out-of-sample R^2s from recursive forecasts of one quarter ahead returns on $\overline{BM_p}$ and $\overline{BM_{mkt}}$. The top panel shows statistical gain in the case of value-weighted returns while the bottom panel is for SP500. Details on the computation of out-of-sample R^2s are outlined in section 2. Dates on the horizontal axis correspond to forecast origins of the out-of-sample procedure. Estimations are performed using quarterly data that spans from Q2 1980 to Q4 2020.

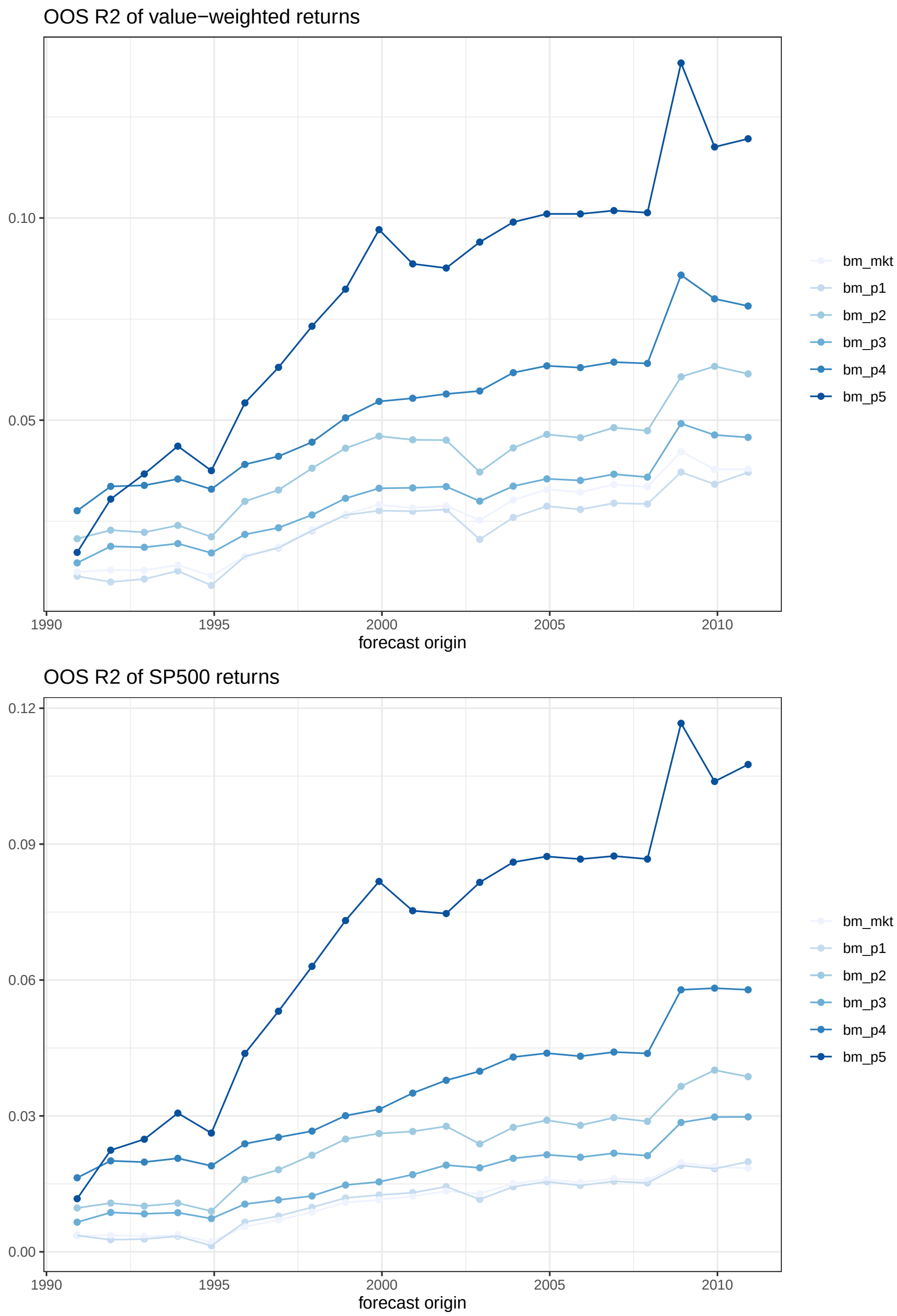

Figure 4: Recursive forecasts economic gains - out of sample

Note: This figure illustrates out-of-sample economic gains from recursive forecasts of one quarter ahead returns on $\overline{BM_p}$ and $\overline{BM_{mkt}}$. The top panel shows economic gains in the case of value-weighted returns while the bottom panel is for SP500. Economic gains measure the added economic utility in terms of a mean-variance investor who moves from a benchmark to a competing forecast that may include various specifications of book-to-market ratios sorted by IBP. Details on the computation are outlined in section 2. Dates on the horizontal axis correspond to forecast origins of the out-of-sample procedure. Estimations are performed using quarterly data that spans from Q2 1980 to Q4 2020.

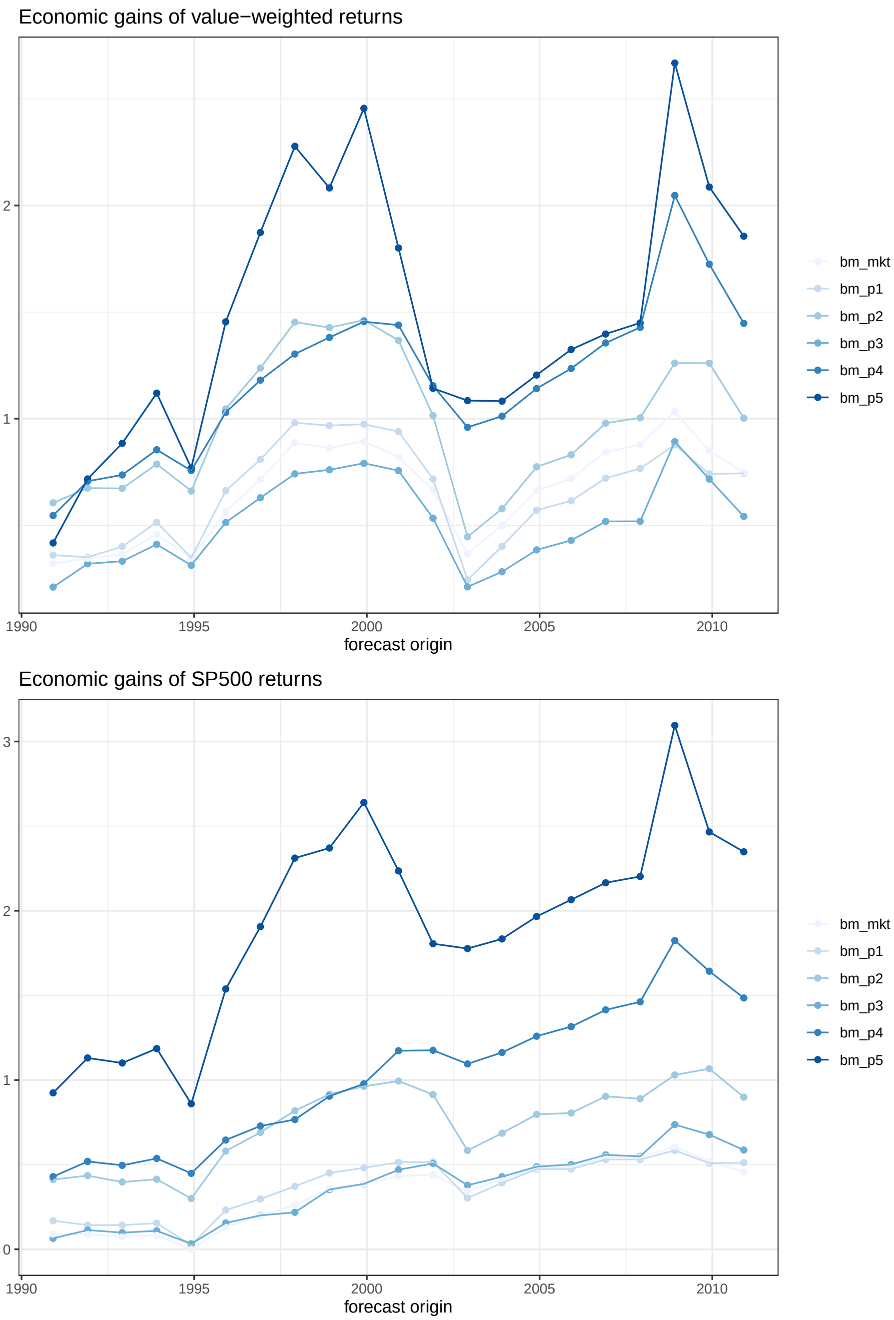

Figure 5: Accounting characteristics by IBP portfolio

Note: This figure illustrates time-series averages of accounting characteristics split by IBP portfolio. EP is earnings-to-price, DP is dividend-to-price, SZ is measured by the log of market capitalization, INV is yearly growth in total assets, OP is yearly growth in operating profitability, REV is the yearly growth in total revenue, DIV is the yearly growth in dividend distributions, ANN DIS is the analyst disagreement measured by the standard deviation of earnings-per-share forecasts and MOM is momentum proxied by yearly returns. The time series means are built from quarterly data starting in Q2 1980 and ending in Q4 2020.

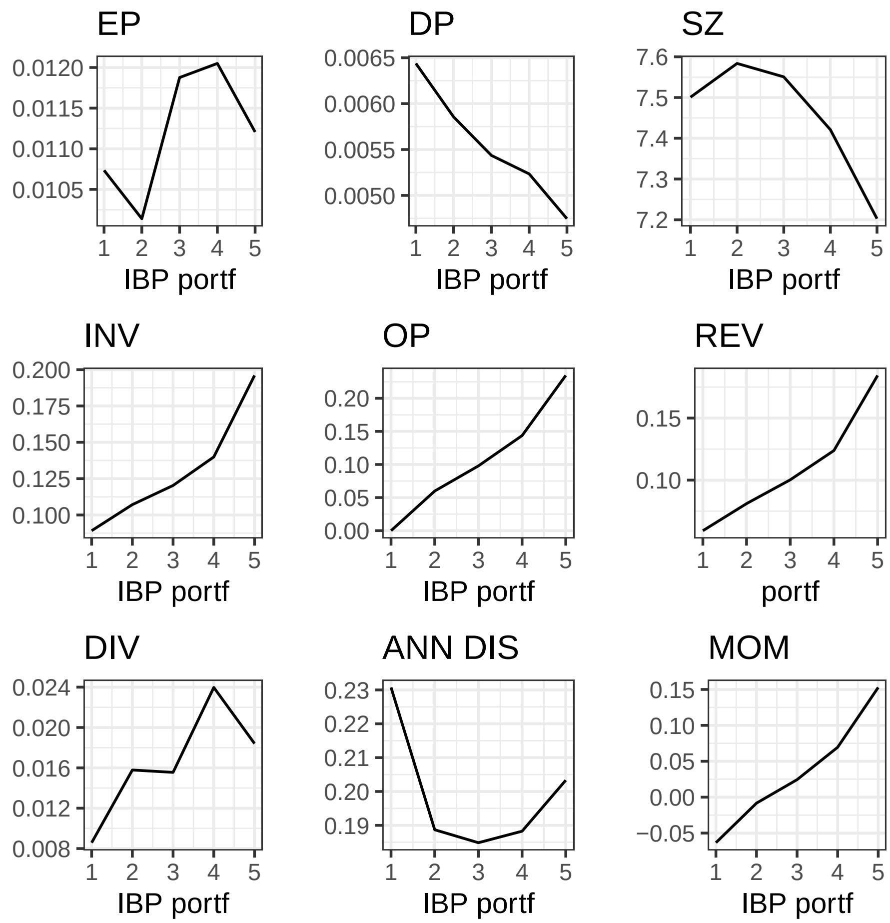

Figure 6: Cross-sectional asset pricing factors and IBP sorted returns

Note: This figure illustrates cross-sectional factors of Chen and Zimmermann (2021) that are relevant in explaining return differences between stocks in the highest and lowest IBP quintiles. To capture which ones may be relevant for explaining institutional buying pressure, I regress them on return differences of IBP portfolios (p5-p1). Black represent t-stats associated with differences in value-weighted returns while the light grey represents t-stats associated with differences in equally-weighted returns of IBP portfolios. Panel A shows highly significant factors determined by Newey-West t-stats greater than 2.5. Factors that emerge significant are *DebtIssuance* of Spiess and Affleck-Graves (1999), *FirmAge* of Barry and Brown (1984), *IndRetBig* of Hou (2007) and *Herf*, *HerfAsset*, *HerfBE* of Hou and Robinson (2006). Panel B shows factors determined by a lower level of significance or Newey-West t-stats greater than 1.96. The complete definition list can be found in Chen and Zimmermann (2021).

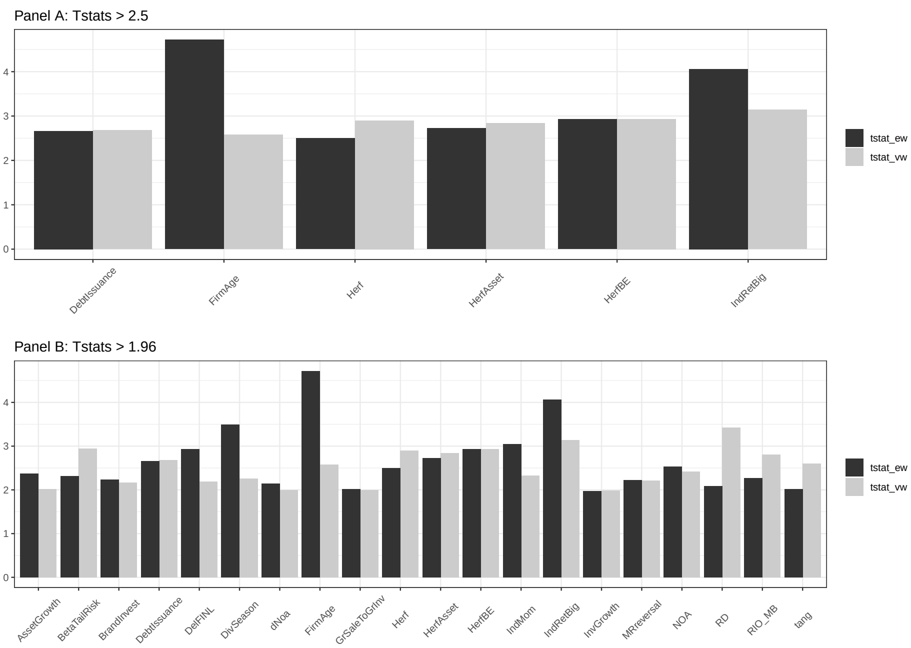

Figure 7: Between and within disagreement of BM ratios in IBP portfolios

Note: This figure illustrates various time-series correlations based on cross-sectional differences among comparable valuation ratios. The top panel plots correlation series between average book-to-market and sales-to-price ratios. The middle panel plots correlations series between book-to-market and earnings-to-price ratios. The bottom panel plots time-series of standard deviations measured in each $\overline{BM_p}$. Blue lines represent stocks in the highest IBP quintile, red lines correspond to stocks in the lowest IBP quintile and black lines correspond to the entire of cross-section.

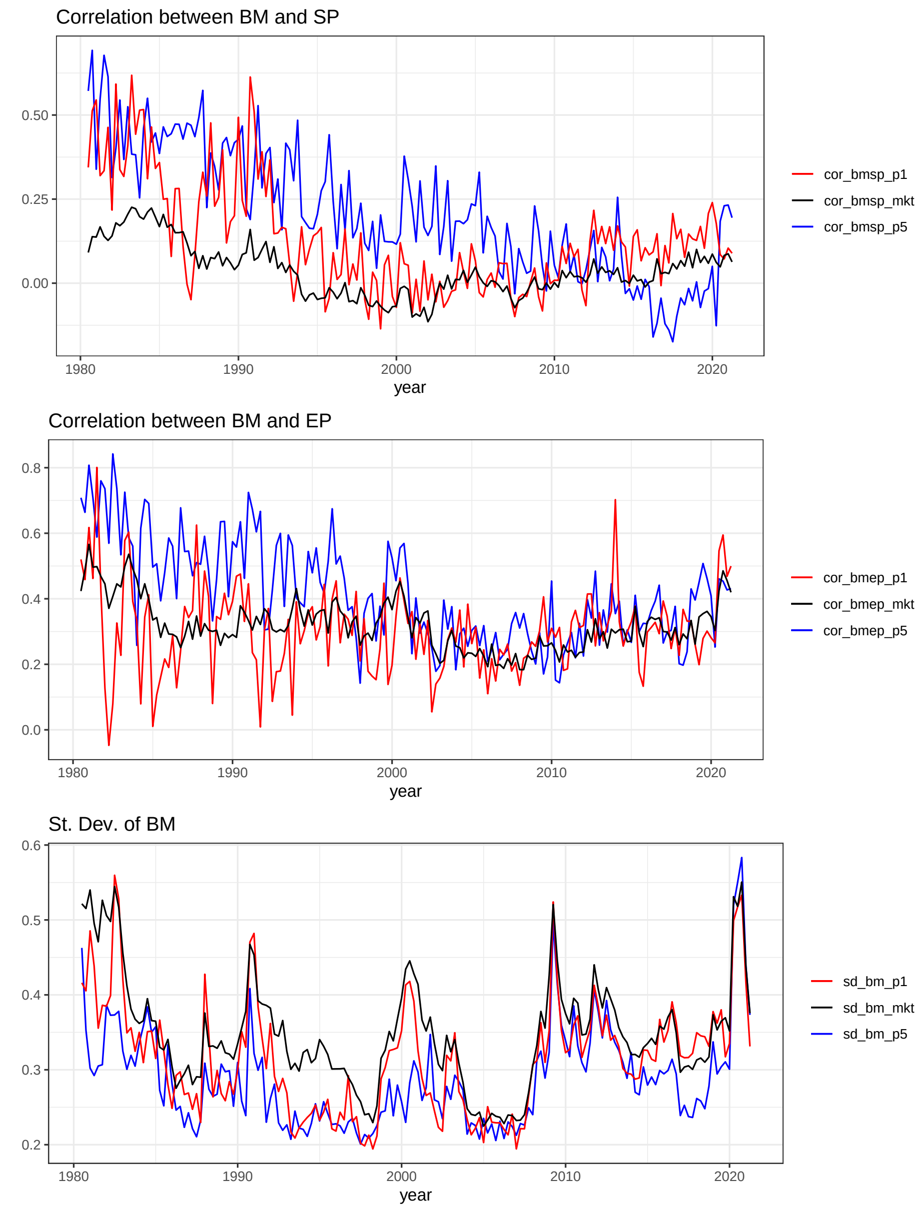

Table 1: Correlations between $\overline{BM_p}$

Note: The table below reports pairwise correlations and their differences between specifications of book-to-market ratios in portfolios sorted by IBP. Panel A displays magnitudes of these correlations. Panel B shows their pairwise mean differences, and Panel C illustrates the Newey-West t-stats associated with the differences from Panel B. Data is quarterly and spans from Q2 1980 to Q4 2020.

Panel A: Correlation Table

	$\overline{BM_{mkt}}$	$\overline{BM_1}$	$\overline{BM_2}$	$\overline{BM_3}$	$\overline{BM_4}$	$\overline{BM_5}$
$\overline{BM_{mkt}}$						
$\overline{BM_1}$	0.92					
$\overline{BM_2}$	0.95	0.93				
$\overline{BM_3}$	0.92	0.86	0.93			
$\overline{BM_4}$	0.87	0.79	0.88	0.93		
$\overline{BM_5}$	0.74	0.66	0.75	0.80	0.86	

Panel B: Mean Differences

	$\overline{BM_{mkt}}$	$\overline{BM_1}$	$\overline{BM_2}$	$\overline{BM_3}$	$\overline{BM_4}$	$\overline{BM_5}$
$\overline{BM_{mkt}}$						
$\overline{BM_1}$	-0.01					
$\overline{BM_2}$	-0.06	-0.05				
$\overline{BM_3}$	-0.08	-0.06	-0.02			
$\overline{BM_4}$	-0.09	-0.08	-0.03	-0.01		
$\overline{BM_5}$	-0.12	-0.11	-0.06	-0.04	-0.03	

Panel C: Mean Differences NW tstats

	$\overline{BM_{mkt}}$	$\overline{BM_1}$	$\overline{BM_2}$	$\overline{BM_3}$	$\overline{BM_4}$	$\overline{BM_5}$
$\overline{BM_{mkt}}$						
$\overline{BM_1}$	-1.24					
$\overline{BM_2}$	-11.73	-4.05				
$\overline{BM_3}$	-8.94	-5.92	-3.13			
$\overline{BM_4}$	-9.66	-4.44	-3.83	-2.17		
$\overline{BM_5}$	-9.34	-6.25	-5.85	-4.23	-4.28	

Table 2: $\overline{BM_p}$ predictability

Note: The table below reports in sample results of equation (4) for each $\overline{BM_p}$. Panel A illustrates betas, Newey-West t-stats and R^2s for CRSP value-weighted returns. Panel B illustrates betas, Newey-West t-stats and R^2s for SP500 returns. The last two columns show differences between betas corresponding to $\overline{BM_1}$ and $\overline{BM_5}$ and Wald tests associated with these differences. Regressions are quarterly and span a time series from Q2 1980 to Q4 2020.

	$\overline{BM_1}$	$\overline{BM_2}$	$\overline{BM_3}$	$\overline{BM_4}$	$\overline{BM_5}$	$\overline{BM_{mkt}}$	$\overline{BM_5}$-$\overline{BM_1}$	$\overline{BM_5}$-$\overline{BM_{mkt}}$
i = CRSP Value Weighted Returns (VW)								
$\hat{\beta}_{VW}$	0.06	0.11	0.1	0.14	0.21	0.09	$\hat{\beta}_{p5,VW}$ - $\hat{\beta}_{p1,VW} = 0.15$	$\hat{\beta}_{p5,VW}$ - $\hat{\beta}_{mkt,VW} = 0.12$
NW tstat	1.60	1.83	1.59	1.99	3.29	1.45	Wald Test p-val = 0.04	Wald Test p-val = 0.04
R^2	0.01	0.03	0.02	0.03	0.06	0.01		
i = SP500								
$\hat{\beta}_{SP500}$	0.04	0.08	0.07	0.11	0.18	0.06	$\hat{\beta}_{p5,SP}$ - $\hat{\beta}_{p1,SP} = 0.14$	$\hat{\beta}_{p5,SP}$ - $\hat{\beta}_{mkt,SP} = 0.12$
NW tstat	1.06	1.34	1.23	1.49	2.93	0.95	Wald Test p-val = 0.02	Wald Test p-val = 0.01
R^2	0.00	0.01	0.01	0.02	0.05	0.00		

Table 3: $\overline{BM_p}$ out of sample R^2 - recursive forecasts

Note: The table below reports out-of-sample R^2s of equation (5) in a recursive scheme for each $\overline{BM_p}$. Details on the procedure are covered in section 2. The top panel presents statistical gains for CRSP value-weighted returns while the bottom panel presents out-of-sample R^2s in the case of SP500 returns. I test $H_a : R^2_{OOS} > 0$ using the methodology suggested by Clark and West (2007). Statistically significant estimates at 5% are denoted in bold. The three forecast start dates considered are ten years apart: "1990-12-01", "2000-12-01" and "2010-12-01". Regressions are quarterly and span a time series from Q2 1980 to Q4 2020.

	OOS Forecast Start: "1990-12-01"	OOS Forecast Start: "2000-12-01"	OOS Forecast Start: "2010-12-01"
	Statistical Gains: OOS R^2 for VW		
$\overline{BM_1}$	**0.01**	**0.03**	**0.04**
$\overline{BM_2}$	**0.02**	**0.04**	**0.06**
$\overline{BM_3}$	**0.01**	**0.03**	**0.05**
$\overline{BM_4}$	**0.03**	**0.06**	**0.08**
$\overline{BM_5}$	**0.02**	**0.09**	**0.12**
$\overline{BM_{mkt}}$	**0.01**	**0.03**	**0.04**
	Statistical Gains: OOS R^2 for SP500		
$\overline{BM_1}$	0.00	**0.01**	**0.02**
$\overline{BM_2}$	**0.01**	**0.03**	**0.04**
$\overline{BM_3}$	0.01	**0.02**	**0.03**
$\overline{BM_4}$	**0.02**	**0.04**	**0.06**
$\overline{BM_5}$	**0.01**	**0.07**	**0.11**
$\overline{BM_{mkt}}$	0.00	**0.01**	**0.02**

Table 4: $\overline{BM_p}$ out of sample R^2 - rolling fixed window forecasts

Note: The table below reports out-of-sample R^2s from equation (5) in a rolling fixed window forecast for each $\overline{BM_p}$. Details on the procedure are covered in section 2. The first column illustrates statistical gains for CRSP value-weighted returns while the second column presents out-of-sample R^2s in the case of SP500 returns. I test $H_a : R^2_{OOS} > 0$ using the methodology suggested by Clark and West (2007). Statistically significant estimates at 5% are denoted in bold. Regressions are quarterly and span a time series from Q2 1980 to Q4 2020.

	OOS R^2 Value-weighted	OOS R^2 SP500
$\overline{BM_1}$	0.00	0.00
$\overline{BM_2}$	0.02	0.00
$\overline{BM_3}$	-0.01	-0.02
$\overline{BM_4}$	**0.03**	**0.02**
$\overline{BM_5}$	**0.04**	**0.03**
$\overline{BM}_{mkt}$	0.01	0.00

Table 5: $\overline{BM_p}$ out of sample economic gains - recursive forecasts

Note: The table below reports out-of-sample economic gains in a recursive scheme for each $\overline{BM_p}$ as measured by changes in certainty equivalents across forecasting models from equation (11). The benchmark forecast is the arithmetic average of past returns. Details on the procedure are covered in section 2. The top panel presents the annualized economic gains for CRSP value-weighted returns while the bottom panel presents the annualized economic gains in the case of SP500 returns. I test H_a : Annual Gain > 0 using a bootstrap procedure described in Appendix B. Statistically significant estimates at 5% or lower are denoted in bold. The three forecast start dates considered are ten years apart: "1990-12-01", "2000-12-01" and "2010-12-01". Regressions are quarterly and span a time series from Q2 1980 to Q4 2020.

	OOS Forecast Start: "1990-12-01"	OOS Forecast Start: "2000-12-01"	OOS Forecast Start: "2010-12-01"
	Econonomic Gains for VW: $(CER_m - CER_b) * 400$		
$\overline{BM_1}$	0.36	**0.94**	**0.74**
$\overline{BM_2}$	**0.61**	**1.37**	**1.00**
$\overline{BM_3}$	0.21	**0.76**	**0.54**
$\overline{BM_4}$	0.55	**1.44**	**1.45**
$\overline{BM_5}$	0.42	**1.80**	**1.86**
$\overline{BM_{mkt}}$	0.32	**0.82**	**0.75**
	Econonomic Gains for SP500: $(CER_m - CER_b) * 400$		
$\overline{BM_1}$	0.17	**0.51**	**0.51**
$\overline{BM_2}$	0.41	**0.99**	**0.90**
$\overline{BM_3}$	0.06	**0.47**	**0.59**
$\overline{BM_4}$	0.43	**1.17**	**1.48**
$\overline{BM_5}$	0.92	**2.24**	**2.35**
$\overline{BM_{mkt}}$	0.09	**0.43**	**0.46**

Table 6: Correlation of $\overline{BM_p}$ between IBP and pseudo portfolios

Note: The table below illustrates correlations between average book-to-market ratios in IBP portfolios ($\overline{BM_p}$) and average book-to-market ratios in pseudo portfolios ($\overline{BM_p}^{pseudo}$). Pseudo portfolios are built in two ways. The first is by identifying attractive firm characteristics that could explain buying pressure as seen in Figure 5. Using this matching, Panel A illustrates pairwise correlations of corresponding book-to-market ratios. The second method employed in building pseudo portfolios is by identifying highly significant cross-sectional factors with a t-stat greater than 2.5 that could explain buying pressure as seen in Figure 6. Using this alternative matching scheme, Panel B shows corresponding pairwise correlations. Correlations results for pseudo portfolios built from a smaller number of accounting characteristics or weaker cross-sectional factors can be found in Appendix C, Table C1.

	Panel A: Accounting Characteristics (Full Set)				
	$\overline{BM_1}$	$\overline{BM_2}$	$\overline{BM_3}$	$\overline{BM_4}$	$\overline{BM_5}$
$\overline{BM_1}^{pseudo}$	0.94	0.9	0.85	0.77	0.61
$\overline{BM_2}^{pseudo}$	0.89	0.86	0.8	0.72	0.56
$\overline{BM_3}^{pseudo}$	0.86	0.83	0.77	0.69	0.51
$\overline{BM_4}^{pseudo}$	0.82	0.82	0.78	0.71	0.54
$\overline{BM_5}^{pseudo}$	0.84	0.86	0.84	0.78	0.63
	Panel B: Cross-Sectional Factors Newey-West t-stats >2.5				
	$\overline{BM_1}$	$\overline{BM_2}$	$\overline{BM_3}$	$\overline{BM_4}$	$\overline{BM_5}$
$\overline{BM_1}^{pseudo}$	0.80	0.77	0.77	0.79	0.74
$\overline{BM_2}^{pseudo}$	0.78	0.80	0.82	0.86	0.81
$\overline{BM_3}^{pseudo}$	0.73	0.75	0.79	0.84	0.76
$\overline{BM_4}^{pseudo}$	0.68	0.70	0.75	0.79	0.73
$\overline{BM_5}^{pseudo}$	0.50	0.51	0.56	0.63	0.58

Table 7: Between and within disagreement in $\overline{BM_p}$

Note: The table below illustrates the between and within disagreement among $\overline{BM_p}$. The between disagreement is measured as a time series mean of correlations between book-to-market, sales-to-price and earnings-to-price. The within disagreement is measured as a time series mean of standard deviation in each $\overline{BM_p}$. The last two columns measure the differences between correlations or standard deviations of stocks in the highest and lowest IBP quintiles. Data spans from Q2 1980 to Q4 2020.

	p1	p2	p3	p4	p5	mkt	p1-p5	mkt-p5
$\overline{Cor(BM,SP)}$	0.13	0.17	0.18	0.21	0.21	0.04	0.07	0.17
NW tstat	5.45	6.18	7.72	9.76	7.44	1.53	2.81	6.3
$\overline{Cor(BM,EP)}$	0.3	0.35	0.38	0.41	0.42	0.32	0.12	0.1
NW tstat	21.98	20.89	25.23	19.77	22.81	27.61	4.94	6.76
$\overline{Sd(BM)}$	0.32	0.3	0.29	0.29	0.29	0.35	0.03	0.06
NW tstat	30.8	31.25	38.11	38.09	31.02	39.09	4.44	7.86

Table 8: Information content book-to-market ratios

Note: The table below reports the variance decomposition results of the effect of $\overline{BM_p}$ on market returns from a classic approach that identifies DR news first and backs out CF news as a residual. The state vector is $Y_t = [r_{mkt,t} - r_{f,t}, \Delta D_t, \frac{D_t}{P_t}, cfnai_t]'$. Panel A illustrates results for CRSP value-weighted returns while Panel B shows the outcome for SP500 returns. In each panel, across $\overline{BM_p}$, I report the OLS beta from equation (4), the cash-flow and discount-rate betas from equation (19), and Newey-West t-stats associated with their corresponding estimates. Regressions are quarterly and span a time series from Q2 1980 to Q4 2020.

	$\overline{BM_1}$	$\overline{BM_2}$	$\overline{BM_3}$	$\overline{BM_4}$	$\overline{BM_5}$	$\overline{BM_{mkt}}$
	CRSP Value Weighted Returns (VW)					
$\hat{\beta}^{OLS}_{VW}$	0.07	0.11	0.1	0.15	0.21	0.09
NW tstat	1.6	1.83	1.59	1.99	3.29	1.45
$\hat{\beta}^{CF}$	0.03	0.06	0.05	0.08	0.15	0.04
NW tstat	0.75	1.01	0.86	1.13	2.3	0.69
$\hat{\beta}^{DR}$	-0.00	-0.00	0.00	0.00	0.01	-0.00
NW tstat	-0.83	-0.12	0.78	0.77	1.46	-0.39
$\hat{\beta}^{E}$	0.03	0.05	0.05	0.06	0.06	0.05
NW tstat	5.27	5.86	5.77	7.58	9.16	4.78
	SP500					
$\hat{\beta}^{OLS}_{SP500}$	0.04	0.08	0.07	0.11	0.18	0.05
NW tstat	1.06	1.34	1.23	1.49	2.93	0.95
$\hat{\beta}^{CF}$	0.03	0.05	0.04	0.08	0.15	0.03
NW tstat	0.64	0.91	0.8	1.05	2.32	0.55
$\hat{\beta}^{DR}$	-0.00	-0.00	0.00	0.00	0.01	-0.00
NW tstat	-1.11	-0.52	0.44	0.44	1.61	-0.88
$\hat{\beta}^{E}$	0.02	0.02	0.02	0.04	0.04	0.02
NW tstat	2.83	3.23	3.09	3.78	4.24	2.51

Appendix

Appendix A. Variance Decomposition - CF news first

An alternative approach is to identify cash flow news directly and back out the discount rate news. Chen and Zhao (2009) argue the cash flow component, which is calculated as a residual in the main identification methodology of section 2, may end up misspecified if the return forecast is also misspecified. Consequently, identifying cash-flow news first may lead to very different estimates of individual components in the variance decomposition. On the other hand, if the VAR is properly specified, Campbell et al. (2010) and Engsted et al. (2012) argue that the two identification approaches should yield similar results regardless of cash flow or discount rate news being computed directly or one of them being backed out as a residual. To ensure the VAR is properly specified in the main results of the paper, I estimate the components of the variance decomposition by identifying the cash flow news directly and then backing out the discount rate news. Following Maio and Philip (2015), I estimate the cash flow component from the present-value relation as:

$$\eta_{t+1}^{CF} = e_2'(I - \rho\Phi_1)\epsilon_{t+1} \tag{A.1}$$

where e_2 is a column vector that takes the value of 1 for the position of the cash flow component ΔD_t in the VAR and zero otherwise. Then, I infer the discount-rate news component as:

$$\eta_{t+1}^{DR} = \eta_{t+1}^{CF} - \eta_{t+1}^{R} \tag{A.2}$$

I analyze the information content of $\overline{BM_{p,t}}$ in this alternative scheme through equation (19). Results are shown in Table A1. For both value-weighted and SP500 returns, estimates of $\hat{\beta}^{CF}$ and $\hat{\beta}^{DR}$ are almost identical with a few immaterial differences in the case of cash flow betas for $\overline{BM_1}$ and $\overline{BM_5}$. Even in the alternative approach, cash flow news betas are significant which reinforces findings from Table 8 that $\overline{BM_5}$ anticipates the cash flow news component of future returns more accurately than $\overline{BM_1}$ or $\overline{BM_{mkt}}$. Given the similarity of estimates between Table A1 and Table 8, there is sufficient evidence to believe the VAR model is properly specified and to ensure the stability of my results. In this process I also confirm the findings of Engsted et al. (2012), that in absence of misspecification in the VAR, estimates should be similar regardless of the component identified first.

As an additional robustness check, I also replace $cfnai_t$ in the benchmark VAR with the term-spread (ts_t) computed as the difference between long- and short-term yields and the default-spread (ds_t) computed as the difference between yields of BAA and AAA rated bonds. Both are well known controls for the business cycle. Thus, I estimate another VAR with state vector $Y_t = [r_{mkt,t} - r_{f,t}, \Delta D_t, \frac{D_t}{P_t}, ts_t, ds_t]'$. The variance decomposition exercise is executed in a standard way with discount-news identified first. Table A2 illustrates the results with these different choices as business cycle controls. No significant differences emerge as the outcome of the exercise matches qualitatively with Table 8 and Table A1. This further strengthens the finding that the source of predictability of book-to-market ratios of firms with high IBP to forecast future returns is rooted in its predictive power for cash flow news.

Appendix B. Bootstrap Algorithm Economic Gains

I use McCracken and Valente (2018) to create a block bootstrap algorithm that computes the level of statistical significance of estimated economic gains corresponding to $\overline{BM_{p,t}}$ in Table 5. According to McCracken and Valente (2018), generating bootstrapped critical values is based on an extension of a standard block-bootstrap methodology from Calhoun (2015) and is applicable to conduct inference in out-of-sample tests. In this case, the bootstrap methodology is preferred over asymptotic critical values since the former allows for estimation error in the standard errors of economic gains. Bootstrapped critical values are constructed in each of the out-of-sample scenarios consisting of different forecast start dates (e.g. "1990-12-01", "2000-12-01" and "2010-12-01"). The number of bootstrap replications is set at 2000 repetitions and the block length is set at 5 observations. For each out-of-sample forecast start date, I simulate the critical values in the following steps:

1. Denote by T=("1990-12-01", "2000-12-01","2010-12-01") the set of start dates for out-of-sample forecasts.
2. Generate 2000 bootstrap samples by reshuffling blocks of the predetermined length in the evaluation period. For example, if T="1990-12-01", the evaluation period is "1991-03-01" till "2020-03-12" (end of sample). I generate 2000 samples by reshuffling observations from the evaluation period. I denote by B the corresponding bootstrap replication with $B = 1, 2, ..., 2000$.
3. For each bootstrap replication B, I estimate rolling out-of-sample forecasts using only information up until quarter t by estimating $\hat{r}^B_{i,t+1|t} = \hat{\beta}^B_{i,0,t} + \hat{\beta}^B_{i,1,t} x^B_t$ where x^B_t are the bootstrapped sets of $\overline{BM_{p,t}}$ from step 2 and $t = T+1, T+2, .., T+P$. Thus, I generate P^B out-of-samples return forecasts of $\hat{r}^B_i$ associated with bootstrap sample B where i =(VW,SP500).
4. For each B, I compute the optimal allocation $\omega^B_{t+1|t} = \frac{1}{\gamma}\frac{\hat{r}^B_{i,t+1|t}}{\hat{\sigma}^{2B}_{t+1|t}}$, and the investors' certainty equivalent $CER^B_p = \hat{\mu}^B_p - 0.5\gamma\hat{\sigma}^B_p$. The coefficient of relative risk aversion γ is set at 5 which is in line with the asset pricing literature.

5. Following the suggestion of McCracken and Valente (2018) I center each CER_p^B by subtracting the mean across all bootstrapped replications $B = 1, 2, ..., 2000$. I denote it as $CER_p^{B,N}$ compute it as $CER_p^{B,N} = CER_p^B - \overline{CER_p^B}$.

6. Gains for bootstrap sample B that measure the additional utility a mean-variance investor generates by moving from a benchmark return forecast with just a constant to a competing return forecast with $\overline{BM_p}$ are computed as $CER_m^{B,N} - CER_b^{B,N}$ where $m = (\overline{BM_1}, \overline{BM_2}, \overline{BM_3}, \overline{BM_4}, \overline{BM_5}, \overline{BM_{mkt}})$. Distribution of 2000 bootstrapped centered economic gains $CER_m^{B,N} - CER_b^{B,N}$ are used to generate empirical densities around the point estimate $CER_m - CER_b$ for all $\overline{BM_p}$. Bootstrapped p-values associated with the alternative hypothesis $H_a : CER_m - CER_b > 0$ are used to determine if the economic gain of the corresponding $\overline{BM_p}$ is statistically meaningful.

Appendix C. Alternative Pseudo Portfolios

I employ two additional specifications in building pseudo portfolios. The outcomes in terms of pairwise correlations between $\overline{BM_p}$ and $\overline{BM_p}^{pseudo}$ can be found in Table C1. Panel A considers pseudo portfolios matched on a reduced set of accounting characteristics which only include SZ, MOM , INV and OP. Results from Panel A of Table C1 qualitatively match the correlation patterns from Panel A of Table 6. The correlation between $\overline{BM_1}$ and $\overline{BM_1}^{pseudo}$ (0.94) is significantly higher than the correlation between between $\overline{BM_5}$ and $\overline{BM_5}^{pseudo}$ (0.59). In Panel B of Table C1, I also consider a reduced set of cross-sectional factors from Chen and Zimmermann (2021) that are of lower statistical significance in explaining the buying pressure of institutions. Those factors can be found graphically in Panel B of Figure 6. In reproducing the correlation patterns, I confirm the narrative that it's more challenging to replicate book-to-market ratios of high IBP portfolios compared to low IBP portfolios. In this alternative construction mechanism of pseudo portfolios, the correlation between $\overline{BM_1}$ and $\overline{BM_1}^{pseudo}$ (0.72) is still higher when compared to the correlation between between $\overline{BM_5}$ and $\overline{BM_5}^{pseudo}$ (0.58).

Table A1: Information content book-to-market ratios -CF First

Note: The table below reports the variance decomposition results of the effect of $\overline{BM_p}$ on market returns from an alternative specification that identifies CF news first and backs out DR news as a residual. The state vector is $Y_t = [r_{mkt,t} - r_{f,t}, \Delta D_t, \frac{D_t}{P_t}, cfnai_t]'$. Details on the procedure are in Appendix A. Panel A illustrates results for CRSP value-weighted returns while Panel B shows the outcome for SP500 returns. In each panel, across $\overline{BM_p}$, I report the OLS beta from equation (4), the cash-flow and discount-rate betas from equation (19), and Newey-West t-stats associated with their corresponding estimates. Regressions are quarterly and span a time series from Q2 1980 to Q4 2020.

	$\overline{BM_1}$	$\overline{BM_2}$	$\overline{BM_3}$	$\overline{BM_4}$	$\overline{BM_5}$	$\overline{BM_{mkt}}$
	CRSP Value Weighted Returns (VW)					
$\hat{\beta}^{OLS}_{VW}$	0.07	0.11	0.10	0.15	0.21	0.09
NW tstat	1.6	1.83	1.59	1.99	3.29	1.45
$\hat{\beta}^{CF}$	0.04	0.06	0.05	0.08	0.13	0.05
NW tstat	1.07	1.23	0.88	1.23	2.10	0.94
$\hat{\beta}^{DR}$	0.01	0.00	0.00	0.00	-0.01	0.01
NW tstat	0.88	0.48	0.02	-0.12	-1.45	0.62
$\hat{\beta}^{E}$	0.03	0.05	0.05	0.06	0.06	0.05
NW tstat	5.27	5.86	5.77	7.58	9.16	4.78
	SP500					
$\hat{\beta}^{OLS}_{SP500}$	0.04	0.08	0.07	0.11	0.18	0.05
NW tstat	1.06	1.34	1.23	1.49	2.93	0.95
$\hat{\beta}^{CF}$	0.04	0.06	0.05	0.08	0.13	0.05
NW tstat	1.05	1.28	0.9	1.22	2.10	1.00
$\hat{\beta}^{DR}$	0.01	0.01	0.01	0.01	-0.01	0.02
NW tstat	1.79	1.56	0.53	0.51	-0.52	1.55
$\hat{\beta}^{E}$	0.02	0.02	0.02	0.04	0.04	0.02
NW tstat	2.83	3.23	3.09	3.78	4.24	2.51

Table A2: Information content book-to-market ratios -DR First and alternative business cycle controls

Note: The table below reports the variance decomposition results of the effect of $\overline{BM_p}$ on market returns based on a classic approach with DR news identified first and CF news being backed as a residual. I consider an alternative specification of the state vector $Y_t = [r_{mkt,t} - r_{f,t}, \Delta D_t, \frac{D_t}{P_t}, ts_t, ds_t]'$ where I replace $cfnai$ with ts (term spread) and ds (default spread) to control for the business cycle. Panel A illustrates results for CRSP value-weighted returns while Panel B shows the outcome for SP500 returns. In each panel, across $\overline{BM_p}$, I report the OLS beta from equation (4), the cash-flow and discount-rate betas from equation (19), and Newey-West t-stats associated with their corresponding estimates. Regressions are quarterly and span a time series from Q2 1980 to Q4 2020.

	$\overline{BM_1}$	$\overline{BM_2}$	$\overline{BM_3}$	$\overline{BM_4}$	$\overline{BM_5}$	$\overline{BM_{mkt}}$
	CRSP Value Weighted Returns (VW)					
$\hat{\beta}^{OLS}_{VW}$	0.07	0.11	0.1	0.15	0.21	0.09
NW tstat	1.6	1.83	1.59	1.99	3.29	1.45
$\hat{\beta}^{CF}$	0.04	0.07	0.06	0.1	0.16	0.05
NW tstat	1.09	1.23	1.02	1.37	2.35	0.92
$\hat{\beta}^{DR}$	0.00	0.00	0.00	0.00	0.00	0.00
NW tstat	-1.07	-0.79	0.22	-0.09	0.27	-0.68
$\hat{\beta}^{E}$	0.02	0.04	0.04	0.05	0.06	0.04
NW tstat	2.84	3.28	3.15	3.42	5.23	2.57
	SP500					
$\hat{\beta}^{OLS}_{SP500}$	0.04	0.08	0.07	0.11	0.18	0.05
NW tstat	1.06	1.34	1.23	1.49	2.93	0.95
$\hat{\beta}^{CF}$	0.04	0.06	0.06	0.09	0.15	0.04
NW tstat	0.97	1.18	1.06	1.3	2.38	0.81
$\hat{\beta}^{DR}$	0.00	0.00	0.00	0.00	0.00	0.00
NW tstat	-1.05	-0.79	0.32	0.14	0.71	-0.81
$\hat{\beta}^{E}$	0.00	0.01	0.01	0.02	0.03	0.01
NW tstat	0.49	0.96	0.92	1.13	1.92	0.64

Table C1: Actual vs. Pseudo Portfolios

Note: The table below illustrates correlations between average book-to-market ratios in IBP portfolios ($\overline{BM_p}$) and average book-to-market ratios in pseudo portfolios ($\overline{BM_p}^{pseudo}$). In this alternative specification, pseudo portfolios are built using a smaller subset of accounting characteristics including only SZ, MOM, INV, and OP. Using this matching, Panel A illustrates pairwise correlations of corresponding book-to-market ratios. I also use a smaller number of cross-sectional factors that can be found in Panel B of Figure 6. Under the scheme with fewer factors, Panel B shows corresponding pairwise correlations of book-to-market ratios.

	Panel A: Accounting Characteristics (Reduced Set)				
	$\overline{BM_1}$	$\overline{BM_2}$	$\overline{BM_3}$	$\overline{BM_4}$	$\overline{BM_5}$
$\overline{BM_1}^{pseudo}$	0.94	0.91	0.86	0.78	0.62
$\overline{BM_2}^{pseudo}$	0.87	0.86	0.8	0.71	0.54
$\overline{BM_3}^{pseudo}$	0.88	0.85	0.79	0.71	0.54
$\overline{BM_4}^{pseudo}$	0.83	0.82	0.78	0.7	0.54
$\overline{BM_5}^{pseudo}$	0.83	0.83	0.81	0.75	0.59
	Panel B: Cross-Sectional Factors T-stats >1.96				
	$\overline{BM_1}$	$\overline{BM_2}$	$\overline{BM_3}$	$\overline{BM_4}$	$\overline{BM_5}$
$\overline{BM_1}^{pseudo}$	0.72	0.71	0.76	0.81	0.83
$\overline{BM_2}^{pseudo}$	0.75	0.76	0.81	0.85	0.84
$\overline{BM_3}^{pseudo}$	0.73	0.72	0.79	0.8	0.76
$\overline{BM_4}^{pseudo}$	0.69	0.67	0.74	0.72	0.73
$\overline{BM_5}^{pseudo}$	0.50	0.45	0.56	0.58	0.58

References

Barry CB, Brown SJ. Differential information and the small firm effect. Journal of Financial Economics 1984;13(2):283–94.

Bennett JA, Sias RW, Starks LT. Greener pastures and the impact of dynamic institutional preferences. The Review of Financial Studies 2003;16(4):1203–38.

Cai F, Zheng L. Institutional trading and stock returns. Finance Research Letters 2004;1(3):178–89.

Calhoun G. A simple block bootstrap for asymptotically normal out-of-sample test statistics. Working Paper, Iowa State University[426, 428, 431, 434] 2015;.

Campbell J. A variance decomposition for stock returns. Economic Journal 1991;101(405):157–79.

Campbell JY, Polk C, Vuolteenaho T. Growth or glamour? fundamentals and systematic risk in stock returns. The Review of Financial Studies 2010;23(1):305–44.

Campbell JY, Shiller RJ. The dividend-price ratio and expectations of future dividends and discount factors. The Review of Financial Studies 1988;1(3):195–228.

Campbell JY, Thompson SB. Predicting excess stock returns out of sample: Can anything beat the historical average? The Review of Financial Studies 2008;21(4):1509–31.

Chen AY, Zimmermann T. Open source cross-sectional asset pricing. Critical Finance Review, Forthcoming 2021;.

Chen L, Zhao X. Return decomposition. The Review of Financial Studies 2009;22(12):5213–49.

Choi KS. Going by the Book: Valuation Ratios and Stock Returns. Ph.D. thesis; Massachusetts Institute of Technology; 2021.

Choi NY, Sias RW. Why does financial strength forecast stock returns? evidence from subsequent demand by institutional investors. The Review of Financial Studies 2012;25(5):1550–87.

Clark TE, West KD. Approximately normal tests for equal predictive accuracy in nested models. Journal of Econometrics 2007;138(1):291–311.

Cochrane JH. The dog that did not bark: A defense of return predictability. The Review of Financial Studies 2008;21(4):1533–75.

Engsted T, Pedersen TQ, Tanggaard C. Pitfalls in var based return decompositions: A clarification. Journal of Banking & Finance 2012;36(5):1255–65.

Evans CL, Liu CT, Pham-Kanter G. The 2001 recession and the chicago fed national activity index: Identifying business cycle turning points. Economic Perspectives-Federal Reserve Bank of Chicago 2002;26(3):26–43.

Fama EF, French KR. Common risk factors in the returns on stocks and bonds. Journal of Financial Economics 1993;33(1):3–56.

Fama EF, French KR. Size and book-to-market factors in earnings and returns. The Journal of Finance 1995;50(1):131–55.

Gompers PA, Metrick A. Institutional investors and equity prices. The Quarterly Journal of Economics 2001;116(1):229–59.

Goyal A, Welch I, Zafirov A, et al. A Comprehensive Look at the Empirical Performance of Equity Premium Prediction II. Technical Report; Swiss Finance Institute; 2021.

Hou K. Industry information diffusion and the lead-lag effect in stock returns. The Review of Financial Studies 2007;20(4):1113–38.

Hou K, Robinson DT. Industry concentration and average stock returns. The Journal of Finance 2006;61(4):1927–56.

Koijen RS, Yogo M. A demand system approach to asset pricing. Journal of Political Economy 2019;127(4):1475–515.

Lewellen J. Predicting returns with financial ratios. Journal of Financial Economics 2004;74(2):209–35.

Maio P, Philip D. Macro variables and the components of stock returns. Journal of Empirical Finance 2015;33:287–308.

McCracken MW, Valente G. Asymptotic inference for performance fees and the predictability of asset returns. Journal of Business & Economic Statistics 2018;36(3):426–37.

Nofsinger JR, Sias RW. Herding and feedback trading by institutional and individual investors. The Journal of Finance 1999;54(6):2263–95.

Park H, et al. An intangible-adjusted book-to-market ratio still predicts stock returns. Critical Finance Review 2019;25(1):207–36.

Pontiff J, Schall LD. Book-to-market ratios as predictors of market returns. Journal of Financial Economics 1998;49(2):141–60.

Rapach D, Zhou G. Asset pricing: Time-series predictability. Available at SSRN 3941499 2021;.

Rapach DE, Ringgenberg MC, Zhou G. Short interest and aggregate stock returns. Journal of Financial Economics 2016;121(1):46–65.

Sias RW, Starks LT, Titman S. Changes in institutional ownership and stock returns: Assessment and methodology. The Journal of Business 2006;79(6):2869–910.

Spiess DK, Affleck-Graves J. The long-run performance of stock returns following debt offerings. Journal of Financial Economics 1999;54(1):45–73.

Welch I, Goyal A. A comprehensive look at the empirical performance of equity premium prediction. The Review of Financial Studies 2008;21(4):1455–508.

Yan X, Zhang Z. Institutional investors and equity returns: Are short-term institutions better informed? The Review of Financial Studies 2009;22(2):893–924.