Institutional Investors and the Time-Variation in Expected Stock Returns

Rüdiger Weber*

First draft: March 30, 2017
This version: March 8, 2019

Latest version: *click here*

Abstract

I document a new stylized fact: the higher the share of institutional ownership in a stock, the more its price-dividend ratio is driven by discount rate variation rather than by changes in dividend growth expectations. Hence, the dividend-price ratio of stocks with high institutional ownership predicts returns. Conversely, for stocks held mostly by individual investors, returns are not predictable. As a general equilibrium outcome, return predictability crucially depends on the properties of the marginal investor. More strongly time-varying volatility in the marginal utility of institutions acting as marginal investors in the respective stocks provides a natural explanation for the observed pattern. In an equilibrium model, time-varying redemption risks generate the observed predictability patterns among a priori identical stocks. My findings help explain the weak return predictability of small and value stocks, the postwar predictability reversal, and the fact that dividend smoothing cannot explain that reversal.

Keywords: predictability, institutional ownership, asset pricing, marginal investor, redemption risk, intermediaries, financialization

JEL: G12, G17, G23

---

* The Stephen M. Ross School of Business, University of Michigan, 701 Tappan Ave, Ann Arbor, MI 48109. E-mail: rcweber@umich.edu

I would like to thank Jules van Binsbergen, Robert Dittmar, Marco Grotteria, Michael Halling, Alexander Hillert, Byoung-Hyoun Hwang, Donald Keim, Yiming Ma, Paulo Maio, Christoph Meinerding, Philippe Mueller, Tyler Muir, Christian Opp, Nikolai Roussanov, Martin Schmalz, Christian Schlag, Tobias Sichert, Robert Stambaugh, Luke Taylor, Julian Thimme, Annika Weber, Ivo Welch, Amir Yaron and seminar participants at Frankfurt (Goethe), Philadelphia (Wharton), Karlsruhe (KIT) and Ann Arbor (Michigan Ross) for valuable comments and insightful discussions.

Parts of this paper were written while I was visiting Wharton. I gratefully acknowledge funding from DAAD German Academic Exchange Service during that time.

1 Introduction

The Campbell-Shiller identity states that the dividend-price ratio, or dividend yield, comprises information about future expected returns and dividend growth. Hence, changes in the dividend yield must predict either returns or dividend growth. Investigating the predictability of returns and dividend growth by the dividend yield therefore provides an answer to what is arguably one of the most fundamental questions in asset pricing: What drives the variation in stock prices normalized by the level of dividends?

Whether time variation in the dividend yield is mostly driven by time variation in expected cash flows or expected returns, i.e. discount rates, depends on the period and the portfolios under study. For example, it has been shown that the dividend yield predicts returns in the postwar sample but not in the prewar sample (Chen, 2009; Golez and Koudijs, 2017). Likewise, returns are predictable for the growth and big portfolios, but not for value and small portfolios (Rytchkov, 2010; Maio and Santa-Clara, 2015).

In this paper, I show that the predictive role of the dividend yield depends on the degree of institutional ownership (IO), with a greater share of dividend yield variation due to discount rate changes for high IO stocks. I show that this novel result is not driven by stock characteristics but by investor characteristics and that this can provide a deeper economic rationale for the aforementioned stylized facts.

Faced with the question of what drives the differences in return predictability between samples, it is not enough to only look at stock characteristics such as market capitalization or accounting measures. This is because expected return variation is a general equilibrium outcome on the stock market. As such, it is inevitably linked to the pricing kernel of the marginal investor, i.e., whoever prices the firm's random cash flow.

If the marginal investor cares about state variables that affect the volatility of her marginal utility, then this will be reflected in the expected returns of the assets she prices. As state variables are time-varying, the expected returns on the assets she prices will also be time-varying and the dividend yield of these assets will predict expected return variation.

Therefore, it is crucial to think about who the marginal investor in a stock may be. While this characteristic is not directly observable, comparing stocks held by different groups of investors is instructive. Sias and Starks (1997) show that institutional ownership is a good predictor of who is the marginal investor.[1] In line with their findings, I argue that the degree of institutional ownership in a stock (and by extension, of a portfolio) serves as a proxy for the likelihood of the marginal investor being institutional and that high IO stocks are more likely to be priced by institutions rather than households. Strong return predictability of high IO stocks is therefore driven by strongly time-varying covariation with the marginal utility of institutions.

In my analysis, I proceed as follows: I first establish the stylized fact that return predictability is stronger for stocks with higher institutional ownership. This paper is the first to document this phenomenon. To that end, I sort stocks into portfolios based on their degree of institutional ownership and run predictive regressions of portfolio returns and dividend growth on the dividend yield as in Cochrane (2008). The results indicate that expected return variation drives virtually all of the variation in the dividend yield of high IO stocks. Conversely, for low IO stocks, expected return variation is negligible and the null of no return predictability cannot be rejected.

Typical features of stocks with high IO or low IO that are also known to affect predictability cannot explain my result. For example, low IO stocks tend to have low market capitalization and high book-to-market ratios and Maio and Santa-Clara (2015) show that returns on the small and value portfolios are not predictable. However, within the small and value portfolios, stocks with high IO have predictable returns whereas the low IO parts of the big and growth portfolios have predictable dividend growth (and no return predictability). Furthermore, the data do not support explanations based on binding short-sale constraints for low IO stocks. There are

[1]Sias and Starks (1997) record a larger fraction of "price-setting orders" for stocks with high institutional ownership attributable to institutional investors. Several studies have since used institutional ownership as a proxy for the likelihood of an institutional marginal investor. Examples include Dhaliwal et al. (2003, 2007)

no differences in cash flow duration between high and low IO stocks that could mechanically explain the result. Explanations based on noisy predictor variables or differing dividend payout policies cannot account for my result, either.

In the theoretical part of my paper, I argue that the underlying rationale for my result is that institutions act as marginal investors in the stocks with high institutional ownership whereas private households are the marginal investors for low IO stocks. Moreover, there are agency frictions between institutions and households and institutions are subject to specific risks due to their provision of liquidity transformation (Chernenko and Sunderam, 2016). Hence, they are not mere pass through entities and it is reasonable to assume that pricing kernels of institutions and private households differ.

Specifically, I suggest that the marginal utility of institutional investors has more time-variation in its volatility than that of households. This is because they are more strongly exposed to time-varying risks. Examples of these risks include liquidity risks that are mostly borne by the institutional sector because it provides liquidity transformation. Examples are the redemption risk faced by mutual funds or the risk of a run on deposits faced by banks. Not only are these risks time-varying and highly correlated with corporate cash flows, they are also systematic to the institutional sector. Hence, in times of higher institution-specific risks institutional investors demand higher risk premia for holding their stock portfolio. For groups of assets for which institutions are marginal investors, these risk premia will translate into time-varying expected returns, i.e., return predictability. Conversely, households get most of their income from labor, hold considerable non-financial wealth, and have less systematic exposure to the aforementioned liquidity risks.

One example of time-varying risks that affect institutions but not households (at least, if at all, to a much lesser degree) is the redemption risk faced by mutual funds. The reason is that mutual funds provide a liquidity transformation to households (Chernenko and Sunderam, 2016). Moreover, households tend to redeem shares in mutual funds rather than selling the stocks they hold directly (Chang et al., 2016; Dorn and Weber, 2017). Hence, the institutional sector bears the risk of sudden liquidity needs and associated fire sales (Coval and Stafford, 2007). I formalize this intuition in an equilibrium model in which heterogeneous investors and time-varying redemption risk generate the pattern observed in the data. Empirical support for the model comes from the fact that mutual fund cash holdings – which proxy for unobserved redemption risk – positively predict returns for high IO stocks, but not for low IO stocks.

The contribution of my paper is threefold. First, my results help to give a more nuanced answer to one of the most fundamental questions in asset pricing: what drives today's asset prices relative to the current level of cash flows? Variation in the dividend yield must reflect variation in the first moments of dividend growth and returns. This follows from the log-linearized return identity:

$$dp_t \approx -k_0 + k_1 \cdot dp_{t+1} - \Delta d_{t+1} + r_{t+1}, \tag{1}$$

where r, dp and Δd denote log return, log dividend yield and log dividend growth, respectively.[2] Equation (1) implies that the dividend yield *must* predict expected returns and/or dividend growth when those are time-varying, or equivalently, that those two quantities must drive the variation in the dividend yield.[3] As general equilibrium outcomes, the evolution of expected returns and dividend growth is a central (and testable) feature of any asset pricing model.[4] Understanding the circumstances under which expected returns are high is therefore not only interesting from a practical investment perspective, it is also of utmost theoretical importance. For example, it may make sense to avoid samples that span both the prewar and postwar years when estimating asset pricing models if the representative investor changed from household to institutional investor. It has been shown for other asset classes that intermediary risk-bearing capacity is a crucial factor in the pricing of asset classes such as CDS, commodities and sovereign

[2]See Appendix A for the derivation.

[3]This is of course leaving aside the possibility of a "bubble": $k_1\rho > 1$ would imply a persistent rise in prices relative to dividends that is not substantiated by higher future cash flows or lower discount rates but only by higher expected valuations.

[4]For example, in Campbell and Cochrane (1999), dividends are assumed to be i.i.d. and all variation in the dividend yield is due to discount rate variation. Conversely, the long-run risk literature following Bansal and Yaron (2004) asserts that cash flows are also predictable.

bonds, see e.g. Haddad and Muir (2018). This paper is the first to show that at least for subsets of the stock market, institution-specific risks predicts returns.

Furthermore, my findings help explain a number of puzzling stylized facts. Against the backdrop of the rise in institutional ownership in the decades after the war (Blume and Keim, 2017) and the tilt of the aggregate institutional portfolio towards growth and big stocks, my results help explain the puzzling results of Chen's (2009) predictability reversal, the fact that dividend smoothing cannot explain this feature of the data (Chen et al., 2012) and the weak return predictability of value and small stocks (Maio and Santa-Clara, 2015).

Finally, this is the first study to consider cross-sectional differences in predictability from the demand side, i.e., from an investor point of view. In my equilibrium model, I formally show how – among a priori identical dividend claims – predictability properties can differ dramatically solely depending on who the marginal investor is.

The paper is organized as follows: After a brief overview of the most closely related literature in Section 2 and the presentation of the data employed in my analyses in Section 3, I establish the predictability properties of stocks with different degrees of institutional ownership in Section 4. The empirical results are discussed in greater detail in Section 5. In Section 6, I discuss the theoretical implications of my result, establish an explicit link of my result to institutional risk bearing and present a general equilibrium model that can generate the observed differences in predictability. Section 7 concludes.

2 Literature

The question whether or not returns and dividend growth are predictable is a fundamental question in financial economics that has given rise to a vast and growing body of literature. It is far beyond the scope of this paper to review all literature on predictability, so I focus on the work related to the predictive power of the price-dividend ratio using the present-value relation of Campbell and Shiller (1988).

From the early 1980s onward, dividend yields have been used to predict stock returns, see e.g. Rozeff (1984); Shiller et al. (1984) and Fama and French (1988, 1989). In particular, Fama and French (1989) emphasize that return predictability by the dividend yield does not violate the efficient markets hypothesis. Rather, as a compensation for risk, expected returns vary with economic conditions. This is an important insight whose implications are crucial for this paper.

However, over the years, return predictability had come under attack. Statistical issues associated with persistent, not strictly exogenous predictor variables meant that some of the evidence in favor of return predictability was spurious. Most notably, Stambaugh (1999) shows that the bias of predictive regressions depends on the persistence of the predictor variable and its lack of strict exogeneity. In a comprehensive study, Goyal and Welch (2008) examine various predictor variables including the dividend yield. The authors find a poor performance in terms of coefficients of determination and out-of-sample performance.

Yet, as can be seen from Equation (1), the dividend yield must predict returns or dividend growth. Hence, it is debatable whether coefficients of determination such as R^2 are apt measures of the predictive power of the dividend yield for returns. Using the tight link between return and dividend growth predictability, Cochrane (2008) establishes that returns are indeed predictable. In Cochrane (2008), most of the support for the predictability of returns comes from the non-predictability of dividend growth rather than from the return predictability regression itself. This is because if dividend growth expectations are not reflected in the dividend yield, its variation must be due to expected returns. Overall, there seems to be a consensus that most of the variation in the dividend yield is due to expected return variation when considering the postwar market portfolio. This, however, does not mean that dividends are not predictable. In fact, there is ample evidence that dividend growth is predictable by other variables, such as the risk-free rate. See for example Bansal et al. (2012).

There are a small number of papers dealing with cross-sectional differences in the time-series predictability and changes in predictability over time which are related to my empirical findings and my theoretical argument. Maio and Santa-Clara (2015) show that returns on small and value stocks are not predictable. Chen (2009) finds that dividends are predictable in the prewar sample, whereas in the postwar sample, returns are predictable. It has been suggested that this could be due to dividend smoothing which led to unpredictable dividend growth and hence to the dominance of return predictability. Chen et al. (2012) show that dividend smoothing can indeed hamper dividend predictability. However, this does not mean that returns are not predictable for non-dividend smoothing firms: Rather, the authors arrive at the "intriguing finding [...] that returns are strongly predictable by [the] dividend yield in the postwar period for both 'dinosaurs' [i.e. firms that have been consistently paying dividends for at least 15 years] and 'non-dinosaurs'. In fact, returns are strongly predictable by dividend yield for all portfolios regardless of whether we separate firms by 'dinosaurs' by dividend smoothing, by earnings smoothing, or even by earnings volatility."(Chen et al., 2012, p. 1850). Thus, dividend smoothing is no exhaustive explanation for the predictability reversal. Finally, Chiang (2015) mentions that for Real Estate Investment Trusts (REITs), the predictability reversal coincides with the advent of institutional investors in the market.

More broadly, this paper is also related to the growing literature that emphasizes features of specific groups of investors and subsets of assets, such as Lettau et al. (2017) but also Drechsler and Drechsler (2016). In particular, this paper adds to the strand of the literature that deals with the impact of financial intermediaries on asset prices such as He and Krishnamurthy (2013); Adrian et al. (2014); He et al. (2017) and Haddad and Muir (2018). Obviously, this paper is related to the literature on the effects of institutional ownership on stock returns such as Gompers and Metrick (2001), Nagel (2005) or Phalippou (2008). In contrast to these studies, I do not examine the cross-section of expected returns but cross-sectional differences in the *time variation* of expected returns. In highlighting the theory of how time-varying dividend growth and time-varying marginal utility affect the dividend yield, this paper is related to Menzly et al. (2004); Santos and Veronesi (2005). However, my model serves the specific purpose of generating cross-sectional differences in predictability rather than providing explanations for the existence of return time-series predictability (Santos and Veronesi, 2005) or modeling how expected dividend growth can confound the predictability of returns (Menzly et al., 2004).

3 Data

I use data on individual stock returns from the Center for Research in Security Prices (CRSP). Dividend yields and dividend growth rates of value-weighted portfolios are computed at annual frequency as

$$\frac{D_{t+1}}{P_{t+1}} = \frac{P_{t+1}+D_{t+1}}{P_t}\frac{P_t}{P_{t+1}} - 1 = \frac{R_{t+1}}{R^X_{t+1}} - 1 \quad \text{and} \quad \frac{D_{t+1}}{D_t} = \frac{D_{t+1}}{P_{t+1}}\frac{P_t}{D_t}\frac{P_{t+1}}{P_t} = \frac{\frac{D_{t+1}}{P_{t+1}}}{\frac{D_t}{P_t}} R^X_{t+1}, \quad (2)$$

where R and R^X denote raw returns with and without dividends, respectively. Data on stock prices and the number of outstanding shares are also from CRSP. In the baseline setting, I only use ordinary common shares with share codes 10 and 11. Subtracting the risk-free rate on both sides of Equation (1) shows that the identity works just as well using excess returns and dividend growth in excess of the risk-free rate, subsequently referred to as excess dividend growth. This means that the dividend yield must also predict excess returns and excess dividend growth. To compute these quantities, I use the three month treasury bill secondary market rate from St. Louis Fed's FRED database as the risk-free rate. Accounting data are from the CRSP-Compustat merged database. Mutual fund cash holding data is from the CRSP Survivorship-free Mutual Fund Database.

I use institutional ownership data from Thomson Reuters that is based on 13(f) filings, adjusted for stock splits. The variable IO share (denoted IO) is the ratio of the number of stocks held by institutional investors over the total number of stocks outstanding as reported by CRSP. All in all, I consider 24,812 stocks, on average 6,408 in a given quarter. Because IO data is not available prior to 1980, I consider 134 quarterly IO observations, from the first quarter of 1980 to the second quarter of 2013.[5]

[5]Using the 140 observations from 1980 to 2014 yields qualitatively similar results but is potentially problematic due to the data quality issues with Thomson Reuters IO data reported after mid 2013 by Wharton

4 Empirical analysis

4.1 Portfolios

I sort the entire sample of stocks into portfolios according to the share of institutional ownership at the end of each year. I form the portfolios based on fixed cutoff points to ensure having both, sufficiently large portfolios, and comparable IO shares over time.

To ensure that portfolios stay large enough, I pick cutoff points of 5, 30 and 50 percent IO in my baseline setting. Figure 1 shows that while portfolio size inevitably varies across time, none of the portfolios becomes too small to only depend on a couple of stocks.[6] As illustrated in Figure 3, the average degree of IO remains fairly stable with the chosen cut-off points.

In Section 4.3, I show that the results prevail qualitatively with quantile portfolios that obviously yield sufficiently large portfolios. However, as shown in Figure 2, the share of overall institutional ownership rises considerably over time. Consequently, with quantile cutoffs, the average degree of IO in a portfolio would not be comparable over time. Hence, the interpretation of being in the high/low IO portfolio would differ dramatically depending on whether an observation is early or late in the sample. In particular, in line with Sias and Starks (1997), I use the degree of IO as a proxy for the likelihood of the marginal investor being institutional. For this interpretation, it is much more meaningful to consider portfolios with stable absolute levels of IO.

[FIGURES 1, 2 and 3 ABOUT HERE]

4.2 Methodology

I determine the predictive power of the dividend yield for returns and dividend growth of the four IO sorted portfolios in the standard way by running predictive regressions as in Cochrane

Research Data Services.

[6]The smallest number of stocks at any time in any portfolio is 148 which roughly coincides with the average number of stocks in a portfolio considered in Bansal et al. (2005) of 150 stocks.

(2008); Chen (2009); Maio and Santa-Clara (2015).

$$r_{t+1} = \mu_r + \beta_r \ dp_t + u_{t+1} \tag{3}$$

$$\Delta d_{t+1} = \mu_d + \beta_d \ dp_t + w_{t+1} \tag{4}$$

$$dp_{t+1} = \alpha + \rho \ dp_t + v_{t+1}. \tag{5}$$

Predictive regressions are econometrically challenging. Consider for example the predictive regression for returns under the assumption that the dp-ratio follows an AR(1) process:

$$r_{t+1} = \mu_r + \beta_r dp_t + u_{t+1} \tag{6}$$

$$dp_{t+1} = \alpha + \rho \ dp_t + v_{t+1} \tag{7}$$

It is reasonable to assume that $\text{Cov}(u, v) < 0$: When returns are unusually high in t, it is likely that the price (relative to dividends) at the end of that period t is high and hence all else equal, dp_t will be low. So while dp_t is independent of u_{t+1}, it is not strictly exogenous, i.e. $\text{Cov}(dp_t, [\ldots, u_{t-1}, u_t]) \neq 0$. Therefore, while the OLS estimator for β_r is consistent, it is not unbiased. As is shown in Stambaugh (1999), the bias is particularly strong when ρ is large. In particular, it holds that $E[\hat{\beta}_r - \beta_r] = -\frac{Cov(u,v)}{Var(v)} \frac{1+3\rho}{T}$. For details, see Appendix B.

Hence, classic inference should be augmented with a simulation analysis. As in Cochrane (2008), Maio and Santa-Clara (2015) and similar to Chen (2009), under the two null hypotheses of i) no return predictability and ii) no dividend predictability, I simulate the complete system of predictive equations (3) to (5) and compare the parameter estimates obtained from the simulated data to those from the actual data. Demeaning (1) and projecting both sides on dp_t, yields

$$1 \approx \beta_r + k_1 \rho - \beta_d, \tag{8}$$

where ρ denotes the autocorrelation of the dividend yield and β_r and β_d denote the slope coefficients of predictive regressions of r_{t+1} and Δd_{t+1} on dp_t, respectively. Because of the identity (8), it makes sense to specify each null two-dimensionally, where the restriction on the third dimension follows from the identity. For example, when $\beta_r = 0$, this implies $\beta_d = k_1\rho - 1$, with a suitably specified null for the value of $\rho = \rho_{H_0}$. The resulting null hypotheses are $H_0^r : \{\beta_r = 0,\ \rho = \rho_{H_0}\}$ (with $\beta_d = k_1\rho_{H_0} - 1$), and $H_0^d : \{\beta_d = 0,\ \rho = \rho_{H_0}\}$ which implies $\beta_r = 1 - k_1\rho_{H_0}$.

The above predictive regressions are informative about the drivers of today's asset prices. To that end, the most informative measures are the long-run coefficients of the two predictive regressions (Cochrane, 2008). They are defined as

$$\beta_r^{LR} = \frac{\beta_r}{1 - k_1\rho} \quad \text{and} \quad \beta_d^{LR} = \frac{\beta_d}{1 - k_1\rho}. \tag{9}$$

Typically, because higher valuations (lower dp) are associated with higher future dividend growth, $\frac{\beta_d}{1-k_1\rho}$ will be negative. The parameter identity (8) implies that

$$\beta_r^{LR} - \beta_d^{LR} \approx 1. \tag{10}$$

What makes these long-run coefficients so informative is that their absolute values can be interpreted as the share of dividend yield variation due to either expected (excess) return or (excess) dividend growth variation. For the derivation of this (and other) interpretations of β_r^{LR} and β_d^{LR}, see Appendix C.

I run the predictive regressions (3) to (5) for each of the four portfolios introduced above. I compute the time t dividend yield as well as time $t+1$ returns, dividend growth and dividend yield for a fixed portfolio of stocks. In other words, the stocks on either side of Equations (3) to (5) are the same.

4.3 Regression results

The results for the predictive regressions are presented in Table 1 and Figure 4. The differences in the predictive power of the dividend yield for returns across IO-sorted portfolios are striking. Essentially all of the variation in the dividend yield is due to variation in expected returns for high IO stocks, whereas for the low IO portfolio, the slope coefficient β_d even has the wrong (negative) sign. The share of dividend yield variation attributable to expected return variation as measured by the long-run coefficients rises from essentially none of the variation to basically all of it.[7] Panel A of Figure 4 shows a graphical depiction. The red parts of the bars represent the share of variation in the dividend yield that can be attributed to discount rate variation. The results show a cross-sectional predictability reversal from dividend growth to returns as the share of institutional ownership rises. The results do not differ much when considering excess returns and excess dividend growth, as can be seen from Panel B of Table 1.

[FIGURE 4 ABOUT HERE]

R^2 shows a similar pattern, although because dp must predict either returns or dividend growth, R^2 is not as informative as the long-run coefficients and is inapt for comparing performance between samples. For example, it may be the case that in one regression, R^2 is rather low because returns are conditionally very volatile, even though there is a strong, stable positive relation between the dividend yield and expected future returns. Consistent with the importance of institutional investors in the overall market, the CRSP value weighted market portfolio is similar to the high IO portfolio in terms of return predictability.[8]

To make use of higher frequency observations and to avoid the issues related to time-aggregation[9], the analysis is repeated using data at quarterly frequency with results presented

[7] Shares can be below zero and above 100% respectively, because the variance decomposition from Appendix C is not a decomposition into orthogonal parts. See also Cochrane (2008, 2011).

[8] Also note that the CRSP value-weighted return includes not only stocks with share codes 10 and 11 but all equity.

[9] Kroencke (2018) shows that before recessions, using end-of-period prices relative to time-aggregated cash-flows falsely suggest that prices fall prior to cash flows, when in fact, prices and cash flows fall contemporaneously.

in Table 2 and Panel B of Figure 4. Overall, the reversal result is less extreme, with more dividend growth information reflected in dividend yields. This may be caused by dividend growth seasonality that is reflected in quarterly but not in annual data.

As shown in Table F.1 in the appendix, using IO quartiles as breakpoints instead yields qualitatively similar results. However, because the average degree of institutional ownership in the market is rising over time, results are not as meaningful and the pattern is less pronounced.

I also run direct weighted long-horizon regressions with a horizon of up to five years as in Cochrane (2008) and Maio and Santa-Clara (2015). The results confirm the above results. Details can be found in Appendix D.

[TABLES 1 and 2 ABOUT HERE]

4.4 Simulation analysis

As mentioned before, predictive regressions suffer from the Stambaugh bias. Moreover, the identity $\beta_r \approx -k_1\rho + \beta_d + 1$ implies that a lack of predictability of one quantity is evidence for predictability of the other. Therefore, to formally study the shift from dividend growth predictability to return predictability as IO increases, I test the null hypotheses of no return predictability, $H_0^r : \{\beta_r = 0, \ \rho = \hat{\rho}\}$ and no dividend growth predictability $H_0^d : \{\beta_d = 0, \ \rho = \hat{\rho}\}$ for each of the portfolios. In other words, under each null, I assume that there is no predictability of either returns or dividend growth, and that therefore all variation in the dividend yield must be due to varying expectations about the other quantity. I assume the persistence of the dividend yield to be equal to the one estimated from the data. The return identity implies that the absence of predictability for one quantity is evidence for the predictability of the other. Thus one should consider the results of both predictive regressions in order to assess the rejection or non-rejection of any null hypothesis.

Doing so improves the power of the tests: No return predictability implies strong dividend predictability. Conversely, whenever I do not observe dividend predictability, this is evidence in favor of return predictability. Moreover, $\beta_r \approx -k_1\rho + \beta_d + 1$ implies that the persistence of the dividend yield affects the size of both β_r and β_d. Taking evidence regarding β_d and ρ into account leads to fewer non-rejections of $\beta_r = 0$ given that, in fact, the true β_r is different from zero (Cochrane, 2008).

Equation (8) implies that the error terms of regressions (3) to (5) must also be related. Deducting the conditional time t expectation from each side of Equation (1) yields

$$r_{t+1} - E_t[r_{t+1}] \approx k_0 - k_0 - k_1\left(dp_{t+1} - E_t\left[dp_{t+1}\right]\right) + \Delta d_{t+1} - E_t\left[\Delta d_{t+1}\right] + dp_t - dp_t \tag{11}$$

$$\Leftrightarrow u_{t+1} \approx w_{t+1} - k_1 v_{t+1}. \tag{12}$$

Under the null of no return predictability, the data are generated by:

$$r_{t+1} = \mu_r + w_{t+1} - k_1 v_{t+1} \tag{13}$$

$$\Delta d_{t+1} = \mu_d + (k_1\rho - 1)dp_t + w_{t+1} \tag{14}$$

$$pd_{t+1} = \alpha + \rho\ dp_t + v_{t+1}. \tag{15}$$

Analogously, under the null of no dividend growth predictability, the data are generated by

$$r_{t+1} = \mu_r - (k_1\rho - 1)dp_t + u_{t+1} \tag{16}$$

$$\Delta d_{t+1} = \mu_d + u_{t+1} + k_1 v_{t+1} \tag{17}$$

$$pd_{t+1} = \alpha + \rho\ dp_t + v_{t+1}. \tag{18}$$

The simulated data are generated using the parameter estimates from the actual data but imposing the null, e.g. μ_r in (13) is given by $\hat{\mu}_r = \frac{1}{T-1}\sum_{t=1}^{T-1} r_{t+1}$.[10] The simulated error terms are drawn from a multivariate normal distribution, the covariance matrix of which is estimated under the respective null hypotheses. Using bootstrapped residuals according to the procedure

[10]Using instead $\rho = 0.99$ or $\rho = 0.975$ for all portfolios gives qualitatively similar, albeit more extreme rejections of either null hypothesis (not tabulated).

in Goyal and Welch (2008) instead does not alter the results qualitatively. Results are also robust to using a vector autoregression (VAR) specification (both not tabulated).

I simulate 5,000 data sets, each consisting of a dividend yield, return and dividend growth time series. Table 3 shows the results of the simulation. For portfolios with little IO, the evidence for no return predictability becomes stronger as compared to portfolios with a larger share of institutional ownership. In particular, for the low IO portfolio, I cannot reject the null hypothesis H_0^r that future returns cannot be predicted by the dividend yield. This implies that discount rates do not significantly influence prices relative to dividends in that portfolio. The most powerful test statistics are those in the rows indicated with "$\widehat{\Pr}[\,\cdot\,|H_0]$", showing the estimated probability of the listed events under the two-dimensional null hypotheses, $H_0^r:\ \{\beta_r = 0,\ \rho = \hat{\rho}\}$ (implying $\beta_d = k_1\rho - 1$) and $H_0^d:\ \{\beta_d = 0,\ \rho = \hat{\rho}\}$ (implying $\beta_r = 1 - k_1\rho$).

This is a very powerful test statistic. It shows the probability under the null that one would observe a long-run slope coefficient as extreme as the one estimated from the data. Take the example of H_0^r in the low IO portfolio: If returns were not predictable, how likely is it that one would observe a β_r^{LR} as large as the one from the data while at the same time observing a β_d^{LR} as large (i.e. not very strongly negative) as in the data? The simulation results indicate that it is quite likely. The null generates coefficients as extreme as in the sample in about 73 percent of the cases.

Figure 5 nicely illustrates the findings explained above. In the top left panel of the figure, the black cross marking the observed sample is right in the middle of the blue cloud of data points generated under the null of no return predictability. As the IO share rises, the evidence for return predictability becomes stronger. This is illustrated in Figure 5 where the cross depicting the observed parameter vector moves further away from the data points generated under H_0^r.

Using excess returns or quarterly data again gives qualitatively similar results (see Table 4 and Tables F.2 and F.3 in the appendix), albeit with non-rejection of H_r^0 at lower significance levels. In the upcoming section, I show that this result is robust to a large number of specifications.

[TABLES 3 AND 4 ABOUT HERE]

[FIGURE 5 ABOUT HERE]

5 Discussion

5.1 Size and value

Table 5 shows some descriptive statistics of the four IO-sorted portfolios. The low IO portfolio has a tilt towards small and value stocks, for which Maio and Santa-Clara (2015) find that returns are not predictable.[11]

[TABLE 5 ABOUT HERE]

To test whether my results just mirror the fact that the low IO portfolio contains more stocks with low market capitalizations and high book-to-market ratios, I form my sample equivalents of the small and big, growth and value portfolios as in Maio and Santa-Clara (2015) and split each into a part with low IO and a part with high IO.[12] I then repeat my analysis of predictive regressions and simulated coefficient distributions. The results are presented in Tables 6 and 7.

[TABLES 6 AND 7 ABOUT HERE]

[11]This is also consistent with Barber and Odean (2000) who find that individual investors tilt their common stock investments towards small and value stocks. Institutions tend to focus on larger stocks (see e.g. Gompers and Metrick (2001)), although recent evidence suggests that this has been changing (Blume and Keim, 2017).

[12]To ensure that portfolios remain sufficiently large but with meaningful cutoffs, the IO cutoff point is 30% for the big and growth portfolio and 5% for the small and value portfolios. The sorting variables size and book-to-market ratio do not differ much between the high and low IO parts of the portfolios. For descriptives, see Table F.4 in Appendix F. Alternatively, one could consider residual IO (RI) as in Nagel (2005) rather than IO. However, this does not yield meaningful cutoff points for my research question. I am interested in who is the dominant investor, not whether institutional ownership is high given the size of a stock. Moreover, a sort on RI yields substantial differences in mean market capitalization across RI-sorted portfolios.

The evidence is striking: Taking the high-powered joint hypothesis p-values of $\beta_r^{LR} \geq \hat{\beta}_r^{LR} \wedge \beta_d^{LR} \geq \hat{\beta}_d^{LR}$, I find that no matter whether I consider a portfolio of small, big, growth or value stocks, a considerably higher share of dividend yield variation is due to discount rate variation for the high IO portfolios. I cannot reject the null of no return predictability for the low IO part of the small, big, growth and value portfolio. Significance levels vary, but the overall pattern is overwhelming. In particular, I find that for the institutionally held parts of the value and small portfolio, returns are predictable. For value stocks, the result is most striking: In the high IO value portfolio, β_d has the wrong (i.e. positive) sign and in the low IO value portfolio β_r has the wrong (i.e. negative) sign.

5.2 Duration and longer forecast horizons

Maio and Santa-Clara (2015) find that the lack of return predictability is mostly driven by value stocks. I.e. dividend growth predictability for small stocks seems to be due to small value stocks. The authors suggest that this may be the result of growth stocks having a longer cash flow duration. Much like the price of a long maturity bond reacts more strongly to interest rate changes, the dividend-price ratios of long-duration stocks react more strongly to discount rate variation (Weber, 2018). Table 5 shows Dechow et al.'s (2004) implied equity duration as a measure for the timing of cash flows. There are virtually no differences in cash flow duration across the IO-sorted portfolios. High investment or high shares of research and development (R&D) expenses may also proxy for longer duration of cash flows. Table 5 shows no marked patterns in size or R&D expenses that would suggest a longer cash flow duration for high IO stocks.

Predictability patterns do not change for longer horizons. The regressions with two- and three-year predictive horizons confirm the pattern from the one-year predictive regression. As before, the share of discount rate-driven variation in the dividend yield rises in IO. The results are presented in Figure 6. Again, for the low IO portfolio, the null hypothesis of no return predictability cannot be rejected at both the two- and three-year horizons (untabulated).

[FIGURE 6 ABOUT HERE]

5.3 Noisy signals and cross-predictions

If a stock exhibits strong variation in mean dividend growth, it may be difficult to detect any expected return information in the dividend yield simply because the expected return information is confounded by dividend growth information. Therefore, another stock's dividend yield that is less influenced by dividend growth variation but shares discount rate components with the first stock may be a better predictor of the first stock's return. I therefore perform cross-predictions, i.e. regressions of one portfolio's return or dividend growth on another portfolio's dividend yield. The results are presented in Table 8. The entry in row i and column j of the table shows the p-value of the one-dimensional null hypotheses $\beta_r^{j,i} = 0$ and $\beta_d^{j,i} = 0$ estimated from a regression of portfolio j's return (dividend growth) on portfolio i's dividend yield. In this setup, $r_{t+1}^j \neq k_0 - k_1 \cdot dp_{t+1}^i + \Delta d_{t+1}^i + dp_t^i$ and the long-run coefficients need not add up to one. Hence, the p-value refers to the probability of observing a $\beta_d^{j,i}$ as small under the null of $H_0^{d,i,j} : \{\beta_d^{j,i} = 0\}$ and a $\beta_r^{j,i}$ as large under the null of $H_0^{r,i,j} : \{\beta_r^{j,i} = 0\}$. Under the nulls, data is simulated from

$$r_{t+1}^j = \hat{\mu}_r + u_{t+1}^j \tag{19}$$

$$dp_{t+1}^i = \hat{\alpha} + \hat{\rho}\ dp_t^i + v_{t+1}^i \tag{20}$$

and

$$\Delta d_{t+1}^j = \hat{\mu}_d + w_{t+1}^j \tag{21}$$

$$dp_{t+1}^i = \hat{\alpha} + \hat{\rho}\ dp_t^i + v_{t+1}^i, \tag{22}$$

respectively. Here, $\hat{\mu}_r = \frac{1}{T-1}\sum_{t=1}^{T-1} r_{t+1}^i$, $\hat{\mu}_d = \frac{1}{T-1}\sum_{t=1}^{T-1} \Delta d_{t+1}^i$. $\hat{\alpha}$ and $\hat{\rho}$ are as estimated from the actual data. For each simulated time series, a set of slope parameters $\{\beta_r^{j,i}, \beta_d^{j,i}\}$ is estimated. In this test setup, no evidence about predictability of the other quantity is taken into account.

The results in Table 8 indicate that the predictability patterns across IO-sorted stocks are not driven by strong dividend growth predictability in the low IO portfolio that "drowns" out the expected return variation in the dividend yield. This is in line with Chen et al. (2012) who find that the strong return predictability in the postwar sample is not due to weaker dividend growth predictability. The first columns in Panel A and B in Table 8 show that also when using other portfolios' valuation ratios no significant time-variation in expected (excess) returns can be detected in the low IO portfolio, i.e. the null of no return predictability cannot be rejected. The high p-values underline the importance of considering the other quantity when conducting inference.

[TABLE 8 ABOUT HERE]

Baker and Wurgler (2004) suggest that firm managers cater to investors' needs. Therefore, institutionally held firms have smoother dividends than other firms (Larkin et al., 2016). This or other differences in payout policy would make it more difficult to detect time-varying cash flow growth information in the dividend yield. To check if this channel explains the documented effects, I also repeat my analysis with the price-earnings ratio as predictor and earnings growth as a less endogenous measure of firms' ability to distribute dividends (Sadka, 2007). The results are shown in Table F.5 in the appendix. Overall, the pattern is similar as with dividend growth: Cash flow predictability is weak for high IO stocks, and their price-earnings ratio contains mostly expected return information. Again, this corroborates the results of Chen et al. (2012) who find that payout policy is no exhaustive explanation of the predictability reversal.

5.4 Limits-to-arbitrage

An explanation for the observed pattern may be found in short-sale constraints. As is well-known, it is difficult to short stocks with little institutional ownership (Nagel, 2005; Stambaugh et al., 2015; Drechsler and Drechsler, 2016). Therefore, less sophisticated non-institutional investors may fail to recognize an overvaluation and thus do not sell 'overvalued' stocks while at the same time, institutional investors who detect the overvaluation cannot correct the mispricing by short-selling because there is no institution they could borrow the stock from. Hence, low dividend yields could persist and would not be followed by lower returns. This would then bias the estimator β_r downwards. To test this hypothesis, I run the predictive regression of excess returns on the dividend yield again but divide the sample into above-median (high dp) or below-median (low dp) dividend yield observations that indicate low and high valuations, respectively. The results are presented in Table 9. The hypothesis outlined above would imply that for high valuations, the effect of dp on expected excess returns is markedly weaker than for low valuations. This is not the case. With the exception of portfolio 3, the slope estimates for the high valuation observations are always larger than for the low valuation observations, although the difference is insignificant in all portfolios but portfolio 2. This is inconsistent with the short-sale constraints argument that I outlined above.

[TABLE 9 ABOUT HERE]

6 Theoretical implications

6.1 Time-varying expected returns and marginal utility

In the previous sections, I have shown that my results are not driven by stock characteristics that can be understood as characteristics of the cash flow supply side. This suggests that the explanation may be found on the demand side, i.e. with the investors in the stock. After all, expected (excess) returns are determined by the covariation of returns with the stochastic discount factor (SDF) of the marginal investor.

This means that, ultimately, when expected returns are to be time-varying, it must be that the covariance of the pricing kernel of the marginal investor with returns is time-varying. Conversely, when returns are not predictable, this implies that the covariation is approximately constant over time. In this context, it is important to note that expected returns are not an "exogenous" stock characteristic but an equilibrium outcome. Institutions compete for expected returns and only when there is a risk that institutions need to be compensated for, high expected returns will not disappear. In other words, only in this case institutions will not "trade away" high expected returns entirely.

This is the rationale of the following explanation: Institutional investors are marginal in the high IO portfolios, while households are marginal investors in the low IO portfolios (Sias and Starks, 1997). Because institutions are not mere pass-through entities for household investment (i.e. there are frictions), the SDF of institutional investors (or, more accurately, that of institutional investment managers) differs from that of the typical household. In particular, it has more strongly time-varying volatility and is more strongly correlated with corporate cash flows. In other words, the sensitivity with respect to taking on the risk of holding stocks is more time-varying for institutions. Hence, there is more expected return variation for high IO stocks. That institutions refrain from investing in certain stocks may be due to explicit investment rules or concerns about individual stock liquidity (Gompers and Metrick, 2001). Market clearing requires that those shunned stocks are then held by households.

It is reasonable to argue that households' marginal utility is not as strongly correlated with corporate cash flows as that of institutional investment managers and that it also has less strongly time-varying volatility. Households derive most of their income from labor.[13] Labor income is less risky than corporate cash flows and it helps to think of labor income as akin to a debt claim on corporate revenues. Indeed, there is ample evidence that firms provide income insurance to their employees (see Guiso et al. (2005); Fagereng et al. (2017); Rettl et al. (2018) and others). Only the more persistent and pervasive corporate cash flow shocks get through to the household sector in a significant manner to affect *aggregate* household utility. Other shocks to households' individual wealth can be diversified away between households. They will thus leave the household sector's corporate cash flow risk bearing capacity largely untouched. This

[13] According to the 2016/2017 Consumer Expenditure Survey, among the households in the sixth to tenth decile of the income distribution (which are more likely to participate in the stock market) the average shares of labor income in total are at 78, 82, 86, 87 and 81 percent (CES, 2018).

holds even more so when these shocks are transient. Consequently, household sector risk bearing capacity does not exhibit high frequency variation.

Conversely, the wealth of institutions is mostly financial. Hence, shocks to financial wealth, i.e. mostly claims to future corporate cash flows, will affect a large part of aggregate institutional wealth. Moreover, institutional investors are not mere shells for household investment or pass-through entities. The largest group of institutional investors in the stock market are investment companies (mostly mutual funds) and investment advisers (Gompers and Metrick, 2001; Blume and Keim, 2017). The compensation of the decision makers in these firms is not fully aligned with the utility of their clients who have substantial sources of income besides financial assets. Mutual fund managers, for example, are typically compensated based on their fund's performance and assets under management (AUM) (Ma et al., 2017). Investment advisers are usually compensated based on their AUM (SEC, 2011). Hence, investment managers' labor income does not insure against financial market and corporate cash flow risk but is rather strongly exposed to it.[14]

6.2 Institution-specific risks

If institutional wealth and manager compensation paid from this wealth is subject to time-varying and systematic risks that are not shared perfectly with the household sector, then these risks will have a time-varying effect on the marginal utility of institutional investors that does not affect the marginal utility of households in the same way. One of the risks in the economy that is primarily borne by institutional investors are risks associated with the provision of liquidity. This is because the institutional sector as a whole provides liquidity transformation to the rest of the economy. Consider for example mutual funds. Mutual fund clients can redeem their shares and have the right to receive cash. This very liquid liability contrasts with less

[14]Passive index funds that are closer to actually being pass-through entities have become popular in recent years. However, they still only account for a rather small fraction of total mutual fund assets. According to the Investment Company Institute Factbooks (2010, 2016), their share of equity mutual fund total net assets rose from 4% in 1995 to about 11% in 2005 and 20% in 2014 at the end of my sample. The vast majority of mutual fund assets has been and still is actively managed.

liquid asset holdings which in the case of redemptions have to be transformed into cash. It is well known that mutual funds hold cash in order to accommodate redemptions, see Chernenko and Sunderam (2016) among others. When the risk of redemptions for the mutual fund industry is high, mutual funds hold more cash than when that risk is low. Therefore, the percentage of equity mutual funds' aggregate assets held in cash (mutual fund cash holdings, MFCH) can serve as a proxy for redemption risk. Hence, MFCH should positively predict stock returns for those stocks that are priced by institutions.

Note that this argument extends to other groups of institutional investors but not to households. MFCH can reasonably be assumed to be correlated with other institutional sector risks, such as investment advisors' risk of losing AUM, funding liquidity risks, etc. Moreover, attempts of one large group of institutions to engage in risk sharing with other institutions will affect those other institutions' risk, leaving the aggregate sectoral risk bearing capacity largely untouched, no matter where the risk originated. In other words, it is systematic to the institutional sector. Conversely, redemption risk is not shared with households. This is because the institutional sector provides liquidity transformation to the economy. For example, a household that redeems its shares in a mutual fund receives cash. The fund needs to come up with that cash, potentially by selling assets at a discount. This leads to costly inefficient risk-sharing within the sector that the household sector does not profit from.[15] Existing empirical evidence suggests that households rather sell their shares in funds than their holdings of individual stocks (Chang et al., 2016; Dorn and Weber, 2017). Moreover, households that invest in funds tend to be different from those that hold more stocks directly. In particular, households that have managed accounts have less stable income.[16] This further supports the argument that liquidity shocks affect institutionally and privately held stocks asymmetrically.

[15] While some firms may profit from buying assets at a discount, the new allocation of assets constitutes a loss in efficiency at the sectoral level. This is because buyers in fire sales tend to be not as specialized as sellers. They therefore incur additional costs (Coval and Stafford, 2007). This leads to a loss to the institutional sector as a whole that the household sector as a whole does not profit from.

[16] According to the Survey of Consumer Finance (SCF), the median share of stable income defined as wage income and social security income including pensions divided by total income is at 73% for households without managed account and at 34% for those with a managed account.

Conveniently for my analysis, fund cash holdings were hardly regulated until 2018, so they were largely at the companies' discretion in my sample period.

Hence, MFCH provide a good opportunity to test the assertion that returns on the high IO portfolio are related to the risk-bearing capacity of institutional investors as opposed to the returns of the low IO portfolio. This is indeed what I find in the data. Table 11 shows the results of predictive regressions of the returns on each of the four IO-sorted portfolios in year $t+1$ on MFCH levels at the end of year t, similar in spirit to Haddad and Muir (2018). For the low IO portfolio, MFCH has no explanatory power for (excess) returns. For the intermediate portfolios, explanatory power for excess returns is at 5% and 6%, respectively. For the high IO portfolio, MFCH explains 14% of excess return variation.[17] Consistent with the findings from Section 5.3, I find that also the returns of the market portfolio are predictable by MFCH, albeit less so than those on the high IO portfolio.

[TABLE 11 ABOUT HERE]

In the next subsection, I show how the observed predictability patterns may arise in the presence of time-varying redemption risk in a segmented market equilibrium. Note that redemption risk is just one way in which the observed patterns may come about, others are also conceivable. This includes all mechanisms that cause institutions to have more strongly time-varying volatility of marginal utility than households. Examples are habit formation utility of investment managers but also anything working in the "opposite direction", i.e. households having less time-varying pricing kernels than institutions, e.g. because they have stable background income.

[17] Yan (2006) interprets a similar result as evidence for *negative market timing ability*. I.e. he argues that if mutual funds had market timing skills, they would hold less cash when expected market returns are high. However, this is a partial equilibrium view. If mutual funds are not mere price takers, then subsequently higher mean market returns may rather be a compensation for risks that MFCH proxies for.

6.3 Model

The model economy is similar to the economy in Wachter (2013) but with time-varying expected dividend growth and heterogeneous agents. There are two heterogeneous investors, representing the household sector h and the institutional sector i. The existence of institutions (or institutional investment managers) is a model assumption. It is based on the empirical finding that households holding stocks directly differ from those that have managed accounts.

The main mechanism driving return predictability for stocks held by the institutional sector is that there is imperfect sharing of time-varying liquidity risk between the institutional and the household sector. This is because the former provides liquidity transformation for the latter. This risk – even if not realized – affects marginal utility in a recursive utility framework leading to higher expected returns and lower price-dividend ratios in times of higher risk. In the following, I present the model setup with the formal solution and the derivation of all results left to Appendix E.

Both agents have stochastic differential utility where the intertemporal elasticity of substitution (IES) is set equal to one, i.e. continuation utility at date t is given by

$$V_t = E_t\left[\int_t^{\infty} f(c_s, V_s)ds\right], \tag{23}$$

with aggregator function

$$f(c_s, V_s) = \beta(1-\gamma)V_s \ln(c_s) - \beta V_s \ln\left((1-\gamma)V_s\right), \tag{24}$$

where β denotes the time preference rate and γ the degree of relative risk aversion. There is a continuum of dividend claims, 'the stock market', each of which pays state-by-state identical cash flows δ^j with dynamics

$$d\delta_t^j = \delta_t^j(\mu_t dt + \sigma_\delta dB_t^\delta). \tag{25}$$

In other words, there is an infinite number of identical dividend claims with the exact same payoffs. This means that a priori, there is no difference between stocks that will be held by the institutional sector or the household sector. In the following, the index j is dropped for convenience. These claims are indivisible, exhibiting what could be called an 'atomic structure'. The indivisibility of one dividend claim between several owners guarantees that there is one investor that is undoubtedly the decisive, marginal investor in a specific stock. This is of course a strong assumption. When transferring this idea to the data, I relax this assumption and rather classify stocks where one group of investors owns a very large share of a stock's market capitalization as being influenced by this group of investors.

Dividend growth dynamics are predictable by assumption. Expected dividend growth at time t, μ_t, is governed by

$$d\mu_t = \kappa_\mu(\bar{\mu} - \mu_t)dt + \sigma_\mu dB_t^\mu. \tag{26}$$

As in many other models, both agents have other sources of income that make consumption growth less volatile than dividend growth. The discrepancy between dividends and consumption is modeled with a 'smoothing parameter' $\xi < 1$ that works in the exact opposite way of the usual leverage parameter (Abel, 1999):

$$c_t = \delta_t^\xi.^{18} \tag{27}$$

Thus, consumption dynamics are

$$dc_t = \delta_t^\xi((\xi\mu_t + 0.5\xi(\xi - 1)\sigma_\delta^2)dt + \xi\sigma_\delta dB_t^\delta) = c_t(\mu_{c,t}dt + \sigma_c dB). \tag{28}$$

The only difference between the two agents is their exposure to liquidity shocks that cause sudden withdrawals of funds and associated losses of $(1 - e^{-\xi L})$ to the share of wealth ϕ that the agent has invested in the stock market. I assume that $L_i \gg L_h \geq 0$. The reason for this parameterization is that institutions provide liquidity transformation to their clients. Hence

[18] Different exposure to corporate cash flows can be modeled by setting $\xi_h < \xi_i$. I do not do so in the baseline setting to focus on exposure to jump risk L that I introduce in the next paragraph.

they have to bear losses due to firesales and illiquid holdings. This is modeled in a very reduced form way. I do not attempt to explicitly model the liquidity needs and how they lead to frictions when the institution attempts to obtain that liquidity but rather refer to the literature on mutual fund cash holdings, redemptions and fire sales and the discussion of how liquidity shocks affect institutions and households asymmetrically in Section 6.1. Liquidity outflow events arrive with time-varying intensity λ:

$$d\lambda_t = \kappa(\bar{\lambda} - \lambda_t)dt + \sigma_\lambda \sqrt{\lambda} dB_t^\lambda. \tag{29}$$

To emphasize the mechanism in the model, I set $L_h = 0$ in the baseline. As long as $L_i \gg L_h$, I obtain qualitatively similar results. For ease of notation, the subscript i is dropped in the baseline case, i.e. $L = L_i$. The compound process of dividends and redemption outflows that the institution obtains from holding a certain measure of the continuum of dividend claims is given by

$$d\tilde{\delta}_t = \tilde{\delta}_t \left(\mu_t dt + \sigma_\delta dB_t^\delta + (e^{-L} - 1)dN_t\right). \tag{30}$$

The resulting consumption growth is given by

$$d\tilde{c}_t = \tilde{c}_t \left(\mu_{c,t} dt + \sigma_c dB_t^\delta + (e^{-\xi L} - 1)dN_t\right). \tag{31}$$

The quantity $\tilde{\delta}$ captures all the cash flows that the institution receives from holding a certain measure of the continuum of dividend claims, i.e. the 'institutional portfolio'. Without outflows, these cash flows are identical to the dividends in (25). If there is an outflow event, the absolute level of dividends flowing from holding the institutional portfolio drops by $\delta_t(1 - e^{-L})$. Subsequent continuous dividend growth of the institutional portfolio accrues to a lower absolute level of dividends.

The household is immune to these risks. As a consequence, the jump terms in (30) and (31) do not show up for the household and neither does the time-varying withdrawal risk λ_t in the optimization problem of the household. Consequently, it does not show up in the household's

SDF whereas the state-price density of the institution is then given by

$$\frac{d\pi_t}{\pi_{t^-}} = \mu_\pi dt - \gamma\sigma_\delta dB_t^\delta + A_2\sigma_\mu dB_t^\mu + \sqrt{\lambda_t}A_1\sigma_\lambda dB_t^\lambda + \left(e^{\xi\gamma L} - 1\right) dN_t. \tag{32}$$

In the case of the household, $L = 0$ and $A_1 = 0$. Moreover, as shown in Appendix E, the dividend-prices ratio of the institutional portfolio and its expected excess returns increase in λ:

$$r_t^j - r_t = \text{const.} - \underbrace{A_1\frac{F_\lambda'}{F}\sigma_\lambda^2}_{<0}\lambda_t + \lambda_t \underbrace{\left[(e^{\gamma\xi L} - 1)(1 - e^{-L})\right]}_{>0}. \tag{33}$$

Consequently, the dividend-price ratio of the institutional portfolio positively predicts returns whereas the dividend-price ratio of the household portfolio does not. A constraint prevents the household from taking levered positions in the stock market. In equilibrium, this implies that all wealth is held in form of the dividend claims and the institution and the household hold fractions of the stock market according to their relative wealth. Intuitively, this also means that the institution cannot raise sufficient capital to trade away time-varying higher expected returns of the institutionally-held part of the stock market that do not compensate for risks that the household cares about. For further discussion, see Appendix E.

7 Conclusion

I show that whether a stock's dividend yield predicts dividend growth or excess returns depends on the degree of institutional ownership (IO). Sorting stocks into portfolios based on the share of stocks held by institutional investors and running predictive regressions as in Cochrane (2008) shows that the relative importance of expected returns for the variation in the price-dividend ratio increases in the pervasiveness of IO. In other words, stocks with higher institutional ownership are characterized by strongly time varying expected returns, whereas expected returns for stocks with little IO do not vary much over time.

My results cannot be explained by cross-sectional differences in established drivers of predictability. This includes stock characteristics such as cash flow duration, market capitalization or the book-to-market ratio. Explanations based on short-sale constraints, noisy predictor variables or differences in payout policy are not supported by the data, either. This corroborates the result of Chen et al. (2012), that independent of payout policy, returns are predictable in the postwar sample.

The fact that cash flow supply-side driven explanations can be ruled out suggests a demand-side, i.e., an investor-based explanation. This is because expected returns are equilibrium outcomes that are influenced by the properties of the marginal investor. In rational models, return predictability is due to time-varying covariation of the pricing kernel with returns. Indeed, I argue that the differences in predictability are driven by strongly time-varying volatility in the stochastic discount factor of institutional investors acting as marginal investors for high IO stocks. In my model, I show how limited sharing of time-varying redemption risk can generate the observed predictability patterns among a priori identical stocks. In the model, stocks held by institutional investors are cheaper in times of higher redemption risk. Empirical support for the model comes from my finding that equity mutual fund cash holdings, which proxy for redemption risk, positively predict the returns of high IO portfolios but not of those with low IO.

My findings speak to several established and puzzling stylized facts about predictability. Maio and Santa-Clara's (2015) result that returns are not predictable for the small and value portfolios can be explained by the fact that individual investors hold large fractions of these stocks. Moreover, my findings are consistent with the predictability reversal in the US stock market which coincided with the rise of IO: "The proportion of equities managed by institutional investors hovered around five percent from 1900 to 1945. But after World War II, institutional ownership started to increase, reaching 67 percent by the end of 2010"(Blume and Keim, 2017, p. 4). My results suggest that one should take a more nuanced look at who the marginal investor in the stock market is. In many cases, one should think of the marginal investor as an institution.

References

Abel, A. B. (1999): "Risk premia and term premia in general equilibrium," *Journal of Monetary Economics*, 43, 3–33.

Adrian, T., E. Etula, and T. Muir (2014): "Financial intermediaries and the cross-section of asset returns," *Journal of Finance*, 69, 2557–2596.

Amihud, Y. (2002): "Illiquidity and stock returns: cross-section and time-series effects," *Journal of Financial Markets*, 5, 31–56.

Baker, M. and J. Wurgler (2004): "A catering theory of dividends," *Journal of Finance*, 59, 1125–1165.

Bansal, R., R. F. Dittmar, and C. T. Lundblad (2005): "Consumption, Dividends, and the Cross Section of Equity Returns," *Journal of Finance*, 60, 1639–1672.

Bansal, R., D. Kiku, and A. Yaron (2012): "An Empirical Evaluation of the Long-Run Risks Model for Asset Prices," *Critical Finance Review*, 1, 183–221.

Bansal, R. and A. Yaron (2004): "Risks for the Long-Run: A Potential Resolution of Asset Pricing Puzzles," *Journal of Finance*, 59, 1481–1509.

Barber, B. M. and T. Odean (2000): "Trading Is Hazardous to Your Wealth: The Common Stock Investment Performance of Individual Investors," *Journal of Finance*, 55, 773–806.

Blume, M. E. and D. B. Keim (2017): "The changing nature of institutional stock investing," *Critical Finance Review*, 6.

Campbell, J. and R. Shiller (1988): "The Dividend-Price Ratio and Expectations of Future Dividends and Discount Factors," *Review of Financial Studies*, 1, 195–228.

Campbell, J. Y. and J. H. Cochrane (1999): "By force of habit: A consumption-based explanation of aggregate stock market behavior," *Journal of Political Economy*, 107, 205–251.

CES (2018): "Consumer Expenditure Survey 2016/2017," Tech. rep., Bureau of Labor Statistics.

Chang, T. Y., D. H. Solomon, and M. M. Westerfield (2016): "Looking for someone to blame: Delegation, cognitive dissonance, and the disposition effect," *Journal of Finance*, 71, 267–302.

Chen, L. (2009): "On the reversal of return and dividend growth predictability: A tale of two periods," *Journal of Financial Economics*, 92, 128 – 151.

Chen, L., Z. Da, and R. Priestley (2012): "Dividend Smoothing and Predictability," *Management Science*, 58, 1834–1853.

Chernenko, S. and A. Sunderam (2016): "Liquidity transformation in asset management: Evidence from the cash holdings of mutual funds," Tech. rep., National Bureau of Economic Research.

Chiang, K. C. (2015): "What Drives REIT Prices? The Time-Varying Informational Content of Dividend Yields," *Journal of Real Estate Research*, 37, 173–190.

Cochrane, J. H. (2008): "The Dog That Did Not Bark: A Defense of Return Predictability," *Review of Financial Studies*, 21, 1533–1575.

——— (2011): "Presidential Address: Discount Rates," *Journal of Finance*, 66.

Coval, J. and E. Stafford (2007): "Asset fire sales (and purchases) in equity markets," *Journal of Financial Economics*, 86, 479–512.

Dechow, P. M., R. G. Sloan, and M. T. Soliman (2004): "Implied Equity Duration: A New Measure of Equity Risk," *Review of Accounting Studies*, 18, 197–228.

Dhaliwal, D., L. Krull, and O. Z. Li (2007): "Did the 2003 Tax Act reduce the cost of equity capital?" *Journal of Accounting and Economics*, 43, 121–150.

Dhaliwal, D., O. Z. Li, and R. Trezevant (2003): "Is a dividend tax penalty incorporated into the return on a firm's common stock?" *Journal of Accounting and Economics*, 35, 155 – 178.

Dorn, D. and M. Weber (2017): "Losing Trust in Money Doctors," Working Paper.

Drechsler, I. and Q. Drechsler (2016): "The Shorting Premium and Asset Pricing Anomalies," Working Paper.

Fagereng, A., L. Guiso, and L. Pistaferri (2017): "Portfolio Choices, Firm Shocks, and Uninsurable Wage Risk," *Review of Economic Studies*, 85, 437–474.

Fama, E. F. and K. R. French (1988): "Dividend Yields and Expected Stock Returns," *Journal of Financial Economics*, 22, 3–25.

——— (1989): "Business Conditions and Expected Returns on Stocks and Bonds," *Journal of Financial Economics*, 25, 23–49.

Golez, B. and P. Koudijs (2017): "Four centuries of return predictability," *Journal of Financial Economics*.

Gompers, P. A. and A. Metrick (2001): "Institutional Investors and Equity Prices," *The Quarterly Journal of Economics*, 116, 229–259.

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

Guiso, L., L. Pistaferri, and F. Schivardi (2005): "Insurance within the firm," *Journal of Political Economy*, 113, 1054–1087.

Haddad, V. and T. Muir (2018): "Do Intermediaries Matter for Aggregate Asset Prices?" Tech. rep., Working Paper, UCLA.

He, Z., B. Kelly, and A. Manela (2017): "Intermediary asset pricing: New evidence from many asset classes," *Journal of Financial Economics*, 126, 1–35.

He, Z. and A. Krishnamurthy (2013): "Intermediary asset pricing," *American Economic Review*, 103, 732–70.

Investment Company Institute Factbooks (2010): "Investment Company Institute factbook," Tech. rep., ICI Statistical Release.

——— (2016): "Investment Company Institute factbook," Tech. rep., ICI Statistical Release.

Kroencke, T. A. (2018): "Recessions and the Stock Market," Working Paper.

Larkin, Y., M. T. Leary, and R. Michaely (2016): "Do Investors Value Dividend-Smoothing Stocks Differently?" *Management Science*.

Lettau, M., S. C. Ludvigson, and S. Ma (2017): "Capital share risk in U.S. stock pricing," Tech. rep., National Bureau of Economic Research.

Ma, L., Y. Tang, and J.-P. Gómez (2017): "Portfolio manager compensation in the US mutual fund industry," Working Paper.

Maio, P. and P. Santa-Clara (2015): "Dividend Yields, Dividend Growth, and Return Predictability in the Cross Section of Stocks," *Journal of Financial and Quantitative Analysis*, 50, 33–60.

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

Nagel, S. (2005): "Short sales, institutional investors and the cross-section of stock returns," *Journal of Financial Economics*, 78, 277–309.

Newey, W. and K. West (1987): "A Simple, Positive Semi-definite, Heteroskedasticity and Autocorrelation Consistent Covariance Matrix," *Econometrica*, 55, 703–708.

Phalippou, L. (2008): "Where Is the Value Premium?" *Financial Analysts Journal*, 64, 41–48.

Rettl, D. A., A. Stomper, and J. Zechner (2018): "The Stability of Dividends and Wages: Effects of Competitor Inflexibility," CFS Working Paper.

Rozeff, M. S. (1984): "Dividend yields are equity risk premiums," *Journal of Portfolio Management*, 11, 68–75.

Rytchkov, O. (2010): "Expected returns on value, growth, and HML," *Journal of Empirical Finance*, 17, 552–565.

Sadka, G. (2007): "Understanding stock price volatility: The role of earnings," *Journal of Accounting research*, 45, 199–228.

Santos, T. and P. Veronesi (2005): "Labor income and predictable stock returns," *Review of Financial Studies*, 19, 1–44.

SEC (2011): "SEC Study on Investment Advisers and Broker-Dealers," Technical Report.

Shiller, R. J., S. Fischer, and B. M. Friedman (1984): "Stock prices and social dynamics," *Brookings papers on economic activity*, 1984, 457–510.

Sias, R. W. and L. T. Starks (1997): "Return autocorrelation and institutional investors," *Journal of Financial Economics*, 46, 103–131.

Stambaugh, R. F. (1999): "Predictive regressions," *Journal of Financial Economics*, 54, 375–421.

Stambaugh, R. F., J. Yu, and Y. Yuan (2015): "Arbitrage Asymmetry and the Idiosyncratic Volatility Puzzle," *Journal of Finance*, 70, 1903–1948.

Wachter, J. A. (2013): "Can Time-Varying Risk of Rare Disasters Explain Aggregate Stock Market Volatility?" *Journal of Finance*, 68, 987–1035.

Weber, R. (2018): "Value, Duration and Predictability," Working Paper.

Yan, X. (2006): "The determinants and implications of mutual fund cash holdings: Theory and evidence," *Financial Management*, 35, 67–91.

Table 1: **Predictive regressions, yearly frequency**

Panel A: Raw returns and dividend growth

		low IO	2	3	high IO	Mkt.
r_{t+1}	β_r	-0.01	0.07***	0.13***	0.13**	0.12**
		(-0.25)	(3.58)	(2.77)	(2.52)	(2.48)
	R^2	0.11%	6.41%	14.08%	8.74%	8.30%
	β_r^{LR}	-0.07	0.31***	1.18***	1.25***	1.06***
		(-0.38)	(3.72)	(4.96)	(5.30)	(4.93)
Δd_{t+1}	β_d	-0.16*	-0.15**	0.02	0.03	0.01
		(-1.76)	(-1.84)	(0.74)	(1.00)	(0.16)
	R^2	12.17%	6.39%	0.40%	0.54%	0.02%
	β_d^{LR}	-1.09***	-0.70***	0.18	0.24	0.05
		(-3.75)	(-2.98)	(0.81)	(1.65)	(0.28)

Panel B: Excess returns and excess dividend growth

		low IO	2	3	high IO	Mkt.
$r_{t+1} - r_{f,t+1}$	β_r	-0.05	0.03	0.09	0.08	0.07
		(-1.03)	(1.38)	(1.49)	(1.24)	(1.16)
	R^2	1.93%	1.26%	6.51%	3.07%	2.79%
	β_r^{LR}	-0.32	0.14*	0.81***	0.75***	0.62***
		(-1.62)	(1.56)	(3.19)	(3.03)	(2.74)
$\Delta d_{t+1} - r_{f,t+1}$	β_d	-0.20**	-0.19**	-0.02	-0.03	-0.04
		(-2.30)	(-2.20)	(-0.49)	(-0.66)	(-0.89)
	R^2	17.64%	9.49%	0.38%	0.55%	1.24%
	β_d^{LR}	-1.33***	-0.87***	-0.19	-0.26*	-0.39**
		(-4.64)	(-3.57)	(-0.75)	(-1.45)	(-1.98)

Panel C: Dividend yield autoregression

		low IO	2	3	high IO	Mkt.
dp_{t+1}	ρ	0.86***	0.80***	0.92***	0.91***	0.91***
		(8.13)	(8.00)	(24.63)	(22.11)	(21.97)

Predictive regression and dp-ratio autoregression results. Numbers in brackets are NW-t-statistics with 10 lags. ***, ** and * for one-period slope estimates indicate significance at the ten, five and one percent level, respectively. For long-run coefficients, stars refer to the respective significance levels of one-sided tests computed according to the delta method.

Table 2: **Predictive regressions, quarterly frequency**

Panel *A*: Raw returns and dividend growth

		low IO	2	3	high IO	Mkt.
r_{t+1}	β_r	0.01	0.03***	0.036**	0.05***	0.04***
		(1.06)	(3.29)	(2.26)	(2.70)	(2.58)
	R^2	0.06%	6.10%	4.44%	4.08%	4.19%
	β_r^{LR}	0.12**	0.25***	0.48***	0.62***	0.66***
		(1.65)	(4.55)	(5.10)	(6.36)	(6.02)
Δd_{t+1}	β_d	−0.11**	−0.10***	−0.04**	-0.03	-0.03
		(-1.85)	(-2.34)	(-1.87)	(-1.04)	(-0.93)
	R^2	4.29%	4.08%	1.59%	0.86%	0.5%
	β_d^{LR}	−0.89***	−0.75***	−0.52***	−0.38***	−0.39***
		(-4.38)	(-4.15)	(-3.93)	(-4.08)	(-3.68)

Panel *B*: Excess returns and excess dividend growth

		low IO	2	3	high IO	Mkt.
$r_{t+1} - r_{f,t+1}$	β_r	0.01	0.18***	0.03**	0.04**	0.04**
		(0.91)	(3.55)	(2.08)	(2.49)	(2.38)
	R^2	0.5%	5.33%	3.88%	3.53%	3.62%
	β_r^{LR}	0.10*	0.03***	0.45***	0.58***	0.61***
	t-stat	(1.39)	(3.02)	(4.74)	(5.92)	(5.58)
$\Delta d_{t+1} - r_{f,t+1}$	β_d	−0.11*	−0.10**	−0.04**	-0.03	-0.03
		(-1.90)	(-2.40)	(-1.99)	(-1.15)	(-1.04)
	R^2	4.46%	4.25%	1.78%	1.07%	0.63%
	β_d^{LR}	−0.91***	−0.77***	−0.55***	−0.42***	−0.43***
		(-4.49)	(-4.25)	(-4.16)	(-4.54)	(-4.09)

Panel *C*: Dividend yield autoregression

		low IO	2	3	high IO	Mkt.
dp_{t+1}	ρ	0.88***	0.87***	0.93***	0.93***	0.94***
		(14.26)	(20.24)	(33.41)	(30.41)	(32.85)

Predictive regression and dp-ratio autoregression results. Numbers in brackets are NW-t-statistics with 10 lags. ***, ** and * for one-period slope estimates indicate significance at the ten, five and one percent level, respectively. For long-run coefficients, stars refer to the respective significance levels of one-sided tests computed according to the delta method.

Table 3: **Simulated $p-$values, yearly frequency**

Panel A: Portfolio 1, low IO

$H_0 : \beta_r = 0, \rho = \hat{\rho}$	β_r	p-value	$Q_{0.05}$	$Q_{0.95}$	β_d	p-value	$Q_{0.05}$	$Q_{0.95}$
short-run	-0.01	0.7132	-0.08	0.17	-0.16	0.3639	-0.41	-0.06
long-run	-0.07	0.7291	-0.54	0.60	-1.09	0.7390	-1.54	-0.40
p joint hyp.	$\beta_r^{LR} \geq \hat{\beta}_r^{LR} \wedge \beta_d^{LR} \geq \hat{\beta}_d^{LR}$: 0.7291							

$H_0 : \beta_d = 0, \rho = \hat{\rho}$	β_r	p-value	$Q_{0.05}$	$Q_{0.95}$	β_d	p-value	$Q_{0.05}$	$Q_{0.95}$
	β_r	p-value	$Q_{0.05}$	$Q_{0.95}$	β_d	p-value	$Q_{0.05}$	$Q_{0.95}$
short-run	-0.01^{**}	0.0070	0.06	0.32	-0.16	0.1880	-0.29	0.08
long-run	-0.07^{**}	0.0079	0.24	1.63	-1.09^{**}	0.0076	-0.76	0.63
p joint hyp.	$\beta_r^{LR} \leq \hat{\beta}_r^{LR} \wedge \beta_d^{LR} \leq \hat{\beta}_d^{LR}$: 0.0076							

Panel B: Portfolio 2

$H_0 : \beta_r = 0, \rho = \hat{\rho}$	β_r	p-value	$Q_{0.05}$	$Q_{0.95}$	β_d	p-value	$Q_{0.05}$	$Q_{0.95}$
short-run	0.07	0.1233	-0.10	0.10	-0.15	0.1062	-0.58	-0.12
long-run	0.31	0.0680	-0.34	0.36	-0.70	0.0713	-1.34	-0.64
p joint hyp.	$\beta_r^{LR} \geq \hat{\beta}_r^{LR} \wedge \beta_d^{LR} \geq \hat{\beta}_d^{LR}$: 0.068							

$H_0 : \beta_d = 0, \rho = \hat{\rho}$	β_r	p-value	$Q_{0.05}$	$Q_{0.95}$	β_d	p-value	$Q_{0.05}$	$Q_{0.95}$
short-run	0.07^{**}	0.0134	0.12	0.32	-0.15	0.3176	-0.38	0.09
long-run	0.31	0.0652	0.28	1.60	-0.70	0.0616	-0.72	0.60
p joint hyp.	$\beta_r^{LR} \leq \hat{\beta}_r^{LR} \wedge \beta_d^{LR} \leq \hat{\beta}_d^{LR}$: 0.0130							

Table continues on next page

Table continues from previous page

Panel C: Portfolio 3

$H_0 : \beta_r = 0, \rho = \hat{\rho}$	β_r	p-value	$Q_{0.05}$	$Q_{0.95}$ β_d	p-value	$Q_{0.05}$	$Q_{0.95}$	
short-run	0.13	0.1606	-0.06	0.21	0.02**	0.0051	-0.31	-0.04
long-run	1.18**	0.0070	-0.58	0.74	0.18**	0.0070	-1.58	-0.26
p joint hyp.		$\beta_r^{LR} \geq \hat{\beta}_r^{LR} \wedge \beta_d^{LR} \geq \hat{\beta}_d^{LR}$: 0.0070						

$H_0 : \beta_d = 0, \rho = \hat{\rho}$	β_r	p-value	$Q_{0.05}$	$Q_{0.95}$	β_d	p-value	$Q_{0.05}$	$Q_{0.95}$
short-run	0.13	0.1780	0.10	0.24	0.02	0.8424	-0.06	0.04
long-run	1.18	0.8994	0.66	1.27	0.18	0.8998	-0.34	0.27
p joint hyp.		$\beta_r^{LR} \leq \hat{\beta}_r^{LR} \wedge \beta_d^{LR} \leq \hat{\beta}_d^{LR}$: 0.8994						

Panel D: Portfolio 4, high IO

$H_0 : \beta_r = 0, \rho = \hat{\rho}$	β_r	p-value	$Q_{0.05}$	$Q_{0.95}$	β_d	p-value	$Q_{0.05}$	$Q_{0.95}$
short-run	0.13	0.2606	-0.10	0.32	0.03	0.0617	-0.29	0.04
long-run	1.25*	0.0451	-1.14	1.22	0.25*	0.0456	-2.14	0.22
p joint hyp.		$\beta_r^{LR} \geq \hat{\beta}_r^{LR} \wedge \beta_d^{LR} \geq \hat{\beta}_d^{LR}$: 0.0451						

$H_0 : \beta_d = 0, \rho = \hat{\rho}$	β_r	p-value	$Q_{0.05}$	$Q_{0.95}$	β_d	p-value	$Q_{0.05}$	$Q_{0.95}$
	β_r	p-value	$Q_{0.05}$	$Q_{0.95}$	β_d	p-value	$Q_{0.05}$	$Q_{0.95}$
short-run	0.13	0.3156	0.09	0.20	0.03	0.8777	-0.04	0.04
long-run	1.25	0.9353	0.70	1.28	0.25	0.9307	-0.30	0.28
p joint hyp.		$\beta_r^{LR} \leq \hat{\beta}_r^{LR} \wedge \beta_d^{LR} \leq \hat{\beta}_d^{LR}$: 0.7394						

Parameter estimates and simulated p-values for the estimates, yearly regression. Q_p denotes the p-quantile of the simulated distribution. *,** and *** denote significance on the one, five and ten percent level, respectively (based on 5,000 simulations and with respect to the respective null hypothesis, i.e. whether the estimate is consistent with the null. **Stars indicate that it is not**). The rows labeled "$\widehat{\Pr}[\, \cdot \, | H_0]$" show the estimated probability of the noted events occuring given that the respective null hypothesis is true.

Table 4: **Simulated $p-$values, yearly frequency, excess returns and dividend growth**

Panel A: Portfolio 1, low IO

$H_0 : \beta_r = 0, \rho = \hat{\rho}$	β_r	p-value	$Q_{0.05}$	$Q_{0.95}$	β_d	p-value	$Q_{0.05}$	$Q_{0.95}$
short-run	-0.05	0.8751	-0.08	0.18	-0.20	0.5325	-0.41	-0.06
long-run	-0.32	0.8916	-0.56	0.61	-1.33	0.8946	-1.56	-0.39
p joint hyp.		$\beta_r^{LR} \geq \hat{\beta}_r^{LR} \wedge \beta_d^{LR} \geq \hat{\beta}_d^{LR}$: 0.8916						
$H_0 : \beta_d = 0, \rho = \hat{\rho}$	β_r	p-value	$Q_{0.05}$	$Q_{0.95}$	β_d	p-value	$Q_{0.05}$	$Q_{0.95}$
	β_r	p-value	$Q_{0.05}$	$Q_{0.95}$	β_d	p-value	$Q_{0.05}$	$Q_{0.95}$
short-run	-0.05^{***}	0.0031	0.06	0.33	-0.20	0.1256	-0.28	0.08
long-run	-0.32^{***}	0.0030	0.24	1.63	-1.33^{***}	0.0030	-0.76	0.63
p joint hyp.		$\beta_r^{LR} \leq \hat{\beta}_r^{LR} \wedge \beta_d^{LR} \leq \hat{\beta}_d^{LR}$: 0.0030						

Panel B: Portfolio 2

$H_0 : \beta_r = 0, \rho = \hat{\rho}$	β_r	p-value	$Q_{0.05}$	$Q_{0.95}$	β_d	p-value	$Q_{0.05}$	$Q_{0.95}$
short-run	0.03	0.3013	-0.10	0.11	-0.19	0.1915	-0.58	-0.12
long-run	0.14	0.2492	-0.36	0.36	-0.87	0.2605	-1.36	-0.64
p joint hyp.		$\beta_r^{LR} \geq \hat{\beta}_r^{LR} \wedge \beta_d^{LR} \geq \hat{\beta}_d^{LR}$: 0.2492						
$H_0 : \beta_d = 0, \rho = \hat{\rho}$	β_r	p-value	$Q_{0.05}$	$Q_{0.95}$	β_d	p-value	$Q_{0.05}$	$Q_{0.95}$
short-run	0.03^{**}	0.0050	0.11	0.33	-0.19	0.2435	-0.38	0.10
long-run	0.14^{**}	0.0124	0.27	1.62	-0.87^{**}	0.0111	-0.73	0.62
p joint hyp.		$\beta_r^{LR} \leq \hat{\beta}_r^{LR} \wedge \beta_d^{LR} \leq \hat{\beta}_d^{LR}$: 0.0111						

Table continues on next page

Table continues from previous page

Panel C: Portfolio 3

$H_0 : \beta_r = 0, \rho = \hat{\rho}$	β_r	p-value	$Q_{0.05}$	$Q_{0.95}$	β_d	p-value	$Q_{0.05}$	$Q_{0.95}$
short-run	0.09	0.2823	-0.07	0.22	-0.02^{*}	0.0327	-0.31	-0.03
long-run	0.81^{*}	0.0439	-0.66	0.79	-0.19^{*}	0.0439	-1.66	-0.21
p joint hyp.		$\beta_r^{LR} \geq \hat{\beta}_r^{LR} \wedge \beta_d^{LR} \geq \hat{\beta}_d^{LR}$: 0.0439						

$H_0 : \beta_d = 0, \rho = \hat{\rho}$	β_r	p-value	$Q_{0.05}$	$Q_{0.95}$	β_d	p-value	$Q_{0.05}$	$Q_{0.95}$
short-run	0.09	0.2152	0.03	0.33	-0.02	0.5889	-0.22	0.07
long-run	0.81	0.5175	0.20	1.69	-0.19	0.5179	-0.80	0.69
p joint hyp.		$\beta_r^{LR} \leq \hat{\beta}_r^{LR} \wedge \beta_d^{LR} \leq \hat{\beta}_d^{LR}$: 0.5157						

Panel D: Portfolio 4, high IO

$H_0 : \beta_r = 0, \rho = \hat{\rho}$	β_r	p-value	$Q_{0.05}$	$Q_{0.95}$	β_d	p-value	$Q_{0.05}$	$Q_{0.95}$
short-run	0.08	0.4111	-0.10	0.33	-0.03	0.1637	-0.30	0.05
long-run	0.75	0.2156	-1.18	1.26	-0.26	0.2216	-2.18	0.26
p joint hyp.		$\beta_r^{LR} \geq \hat{\beta}_r^{LR} \wedge \beta_d^{LR} \geq \hat{\beta}_d^{LR}$: 0.2156						

$H_0 : \beta_d = 0, \rho = \hat{\rho}$	β_r	p-value	$Q_{0.05}$	$Q_{0.95}$	β_d	p-value	$Q_{0.05}$	$Q_{0.95}$
	β_r	p-value	$Q_{0.05}$	$Q_{0.95}$	β_d	p-value	$Q_{0.05}$	$Q_{0.95}$
short-run	0.08	0.2436	-0.01	0.44	-0.03	0.4466	-0.21	0.14
long-run	0.75	0.3761	-0.11	2.06	-0.26	0.3725	-1.11	1.06
p joint hyp.		$\beta_r^{LR} \leq \hat{\beta}_r^{LR} \wedge \beta_d^{LR} \leq \hat{\beta}_d^{LR}$: 0.3725						

Parameter estimates and simulated p-values for the estimates, yearly regression. Q_p denotes the p-quantile of the simulated distribution. *,** and *** denote significance on the one, five and ten percent level, respectively (based on 5,000 simulations and with respect to the respective null hypothesis, i.e. whether the estimate is consistent with the null. **Stars indicate that it is not**). The rows labeled "$\widehat{\Pr}[\,\cdot\,|H_0]$" show the estimated probability of the noted events occuring given that the respective null hypothesis is true.

Table 5: **Portfolio characteristics**

		1 "low IO"	2	3	4 "high IO"
qtrly return r^i	t.s. mean %	2.04	2.58	3.16	3.27
	t.s. std. %	11.52	9.05	8.28	8.87
	NW t-stat.	2.34	3.29	4.24	4.52
yearly return r^i	t.s. mean %	12.05	12.10	12.38	13.06
	t.s. std. %	22.80	17.28	16.83	17.30
	t-stat.	3.04	4.02	4.22	4.34
qtrly div. growth $\frac{D_{t+1}}{D_t} - 1$	t.s. mean %	0.85	9.12	2.50	1.86
	t.s. std. %	40.90	69.36	16.26	12.44
	NW t-stat.	0.31	1.55	2.43	2.57
	autocorr. qtrly	-0.26	-0.05	-0.21	-0.36
yrly div. growth $\frac{D_{t+1}}{D_t} - 1$	t.s. mean %	-0.31	14.06	6.61	6.16
	t.s. std. %	29.24	81.61	16.66	14.42
	t-stat.	-0.06	0.99	2.28	2.46
	autocorr. yearly	0.01	0.02	0.02	-0.31
price-dividend ratio $\frac{P}{D}$	yearly t.s. mean	84	53	43	50
	yearly t.s. std.	53	31	20	18
market equity (in m \$)	e.w. mean	85	329	1607	3333
market equity inv. percentile	e.w. mean	0.19	0.35	0.56	0.74
book-to-market equity ratio	v.w. mean	0.71	0.67	0.51	0.46
book leverage	e.w. mean	0.62	0.58	0.55	0.55
R&D share of expenses	v.w. mean	0.03	0.03	0.03	0.04
investment in %	e.w. mean	4.30	4.84	3.80	4.13
	v.w. mean	6.19	5.01	3.62	4.10
share of dividend payers	t.s. mean	0.21	0.35	0.48	0.56
cash flow duration	e.w. mean	16.6	16.4	16.2	16.0
Amihud (2002) illiquidity	v.w. mean × 100	0.2470	0.0713	0.0143	0.0031

Portfolio characteristics for the four IO-sorted portfolios (baseline setting). Unless otherwise stated, all statistics are computed on a quarterly frequency. "t.s." indicates moments computed along the time-series, "e.w." and "v.w." denote time-series moments of cross sectional means computed equally (value-) weighted. "std." denotes usual standard deviations. "NW t-stats" denote t-statistics computed with Newey and West (1987) standard errors. The price-dividend ratio is computed annually. "Autocorr." denotes the first order autocorrelation. Market equity inverse percentile is the percentile in the cross-sectional distribution of market equity the stocks in my sample at each quarter. "R&D share of expenses" is the ratio of research and development expenses to total operating expenses. Investment is the relative growth in total assets. "share of dividend payers" is the share of stocks in a given quarter that pays dividends. "cash-flow duration" is Dechow et al. (2004) implied equity duration. Amihud's (2002) illiquidity measure is computed as $\frac{|r_i^{daily}|}{Vol_t}$ where Vol_t is the daily trading volume in Dollars, $r_{i,t}$ is the daily return of individual stocks i.

Table 6: **Simulated $p-$values, small and big portfolios**

Panel A: Small Stocks (bottom 30% by Market Equity)

Panel $A.1$: Portfolio IO<5%

$H_0: \beta_r = 0, \rho = \hat{\rho}$	β_r	p-value	$Q_{0.05}$	$Q_{0.95}$	β_d	p-value	$Q_{0.05}$	$Q_{0.95}$
short-run	0.04	0.3883	-0.13	0.19	-0.34	0.2390	-0.71	-0.23
long-run	0.12	0.3483	-0.37	0.36	-0.88	0.3485	-1.37	-0.64
p joint hyp.		$\beta_r^{LR} \geq \hat{\beta}_r^{LR} \wedge \beta_d^{LR} \geq \hat{\beta}_d^{LR}$: 0.3483						

$H_0: \beta_d = 0, \rho = \hat{\rho}$	β_r	p-value	$Q_{0.05}$	$Q_{0.95}$	β_d	p-value	$Q_{0.05}$	$Q_{0.95}$
	β_r	p-value	$Q_{0.05}$	$Q_{0.95}$	β_d	p-value	$Q_{0.05}$	$Q_{0.95}$
short-run	0.04***	0.0004	0.24	0.57	-0.34	0.0507	-0.34	0.15
long-run	0.12***	0.0006	0.47	1.53	-0.88***	0.0006	-0.53	0.53
p joint hyp.		$\beta_r^{LR} \leq \hat{\beta}_r^{LR} \wedge \beta_d^{LR} \leq \hat{\beta}_d^{LR}$: 0.0006						

Panel $A.2$: Portfolio IO>5%

$H_0: \beta_r = 0, \rho = \hat{\rho}$	β_r	p-value	$Q_{0.05}$	$Q_{0.95}$	β_d	p-value	$Q_{0.05}$	$Q_{0.95}$
short-run	0.18	0.1067	-0.17	0.23	-0.33	0.0714	-0.84	-0.31
long-run	0.35	0.0597	-0.37	0.36	-0.65	0.0603	-1.37	-0.64
p joint hyp.		$\beta_r^{LR} \geq \hat{\beta}_r^{LR} \wedge \beta_d^{LR} \geq \hat{\beta}_d^{LR}$: 0.0597						

$H_0: \beta_d = 0, \rho = \hat{\rho}$	β_r	p-value	$Q_{0.05}$	$Q_{0.95}$	β_d	p-value	$Q_{0.05}$	$Q_{0.95}$
short-run	0.18***	0.0021	0.33	0.76	-0.33	0.0547	-0.34	0.20
long-run	0.35***	0.0043	0.55	1.50	-0.65***	0.0041	-0.45	0.50
p joint hyp.		$\beta_r^{LR} \leq \hat{\beta}_r^{LR} \wedge \beta_d^{LR} \leq \hat{\beta}_d^{LR}$: 0.0041						

Table continues on next page

Table continues from previous page

Panel *B*: Big Stocks (top 30% by Book-to-Market Equity ratio)

Panel *B.1*: Portfolio IO < 30%

$H_0: \beta_r = 0, \rho = \hat{\rho}$	β_r	p-value	$Q_{0.05}$	$Q_{0.95}$	β_d	p-value	$Q_{0.05}$	$Q_{0.95}$
short-run	0.07	0.0923	-0.10	0.09	-0.23	0.1339	-0.67	-0.18
long-run	0.23	0.0642	-0.26	0.26	-0.77	0.0660	-1.26	-0.74
p joint hyp.		$\beta_r^{LR} \geq \hat{\beta}_r^{LR} \wedge \beta_d^{LR} \geq \hat{\beta}_d^{LR}$: 0.0642						

$H_0: \beta_d = 0, \rho = \hat{\rho}$	β_r	p-value	$Q_{0.05}$	$Q_{0.95}$	β_d	p-value	$Q_{0.05}$	$Q_{0.95}$
short-run	0.07^{***}	0.0003	0.20	0.39	-0.23	0.1902	-0.39	0.12
long-run	0.23^{**}	0.0073	0.37	1.59	-0.77^{**}	0.0070	-0.63	0.59
p joint hyp.		$\beta_r^{LR} \leq \hat{\beta}_r^{LR} \wedge \beta_d^{LR} \leq \hat{\beta}_d^{LR}$: 0.0070						

Panel *B.2*: Portfolio IO > 30%

$H_0: \beta_r = 0, \rho = \hat{\rho}$	β_r	p-value	$Q_{0.05}$	$Q_{0.95}$	β_d	p-value	$Q_{0.05}$	$Q_{0.95}$
short-run	0.12	0.2442	-0.09	0.28	0.01	0.0509	-0.29	0.01
long-run	1.13^{*}	0.0429	-0.96	1.08	0.12^{*}	0.0443	-1.96	0.08
p joint hyp.		$\beta_r^{LR} \geq \hat{\beta}_r^{LR} \wedge \beta_d^{LR} \geq \hat{\beta}_d^{LR}$: 0.0429						

$H_0: \beta_d = 0, \rho = \hat{\rho}$	β_r	p-value	$Q_{0.05}$	$Q_{0.95}$	β_d	p-value	$Q_{0.05}$	$Q_{0.95}$
	β_r	p-value	$Q_{0.05}$	$Q_{0.95}$	β_d	p-value	$Q_{0.05}$	$Q_{0.95}$
short-run	0.12	0.3931	0.01	0.40	0.01	0.6622	-0.20	0.11
long-run	1.13	0.6967	0.06	1.86	0.12	0.6900	-0.94	0.86
p joint hyp.		$\beta_r^{LR} \leq \hat{\beta}_r^{LR} \wedge \beta_d^{LR} \leq \hat{\beta}_d^{LR}$: 0.6900						

Parameter estimates and simulated p-values for the estimates, sorted on market equity. Q_p denotes the p-quantile of the simulated distribution. *,** and *** denote significance on the one, five and ten percent level, respectively (based on 5,000 simulations and with respect to the respective null hypothesis, i.e. whether the estimate is consistent with the null. **Stars indicate that it is not**). The rows labeled "$\widehat{\Pr}[\cdot|H_0]$" show the estimated probability of the noted events occuring given that the respective null hypothesis is true.

Table 7: **Simulated $p-$values, growth and value portfolios**

Panel A: Growth Stocks (bottom 30% by Book-to-Market Equity ratio)

Panel $A.1$: Portfolio IO < 30%

$H_0 : \beta_r = 0, \rho = \hat{\rho}$	β_r	p-value	$Q_{0.05}$	$Q_{0.95}$	β_d	p-value	$Q_{0.05}$	$Q_{0.95}$
short-run	0.08	0.2682	-0.16	0.20	-0.39	0.2000	-0.82	-0.28
long-run	0.17	0.2259	-0.36	0.35	-0.83	0.2200	-1.36	-0.65
p joint hyp.		$\beta_r^{LR} \geq \hat{\beta}_r^{LR} \wedge \beta_d^{LR} \geq \hat{\beta}_d^{LR}$: 0.2230						

$H_0 : \beta_d = 0, \rho = \hat{\rho}$	β_r	p-value	$Q_{0.05}$	$Q_{0.95}$	β_d	p-value	$Q_{0.05}$	$Q_{0.95}$
	β_r	p-value	$Q_{0.05}$	$Q_{0.95}$	β_d	p-value	$Q_{0.05}$	$Q_{0.95}$
$Q_{0.05}$	$Q_{0.95}$							
short-run	0.08***	0.0006	0.29	0.67	-0.39*	0.0400	-0.37	0.19
long-run	0.17***	0.0010	0.49	1.54	-0.83***	0.0000	-0.51	0.54
p joint hyp.		$\beta_r^{LR} \leq \hat{\beta}_r^{LR} \wedge \beta_d^{LR} \leq \hat{\beta}_d^{LR}$: 0.0010						

Panel $A.2$: Portfolio IO > 30%

$H_0 : \beta_r = 0, \rho = \hat{\rho}$	β_r	p-value	$Q_{0.05}$	$Q_{0.95}$	β_d	p-value	$Q_{0.05}$	$Q_{0.95}$
short-run	0.19	0.1203	-0.11	0.27	0.02**	0.0100	-0.41	-0.06
long-run	1.09**	0.0083	-0.68	0.76	0.10**	0.0100	-1.68	-0.24
p joint hyp.		$\beta_r^{LR} \geq \hat{\beta}_r^{LR} \wedge \beta_d^{LR} \geq \hat{\beta}_d^{LR}$: 0.0081						

$H_0 : \beta_d = 0, \rho = \hat{\rho}$	β_r	p-value	$Q_{0.05}$	$Q_{0.95}$	β_d	p-value	$Q_{0.05}$	$Q_{0.95}$
short-run	0.19	0.3991	0.06	0.46	0.02	0.7000	-0.24	0.12
long-run	1.09	0.7197	0.28	1.67	0.10	0.7200	-0.72	0.67
p joint hyp.		$\beta_r^{LR} \leq \hat{\beta}_r^{LR} \wedge \beta_d^{LR} \leq \hat{\beta}_d^{LR}$: 0.7197						

Table continues on next page

Table continues from previous page

Panel B: Value Stocks (top 30% by Book-to-Market Equity ratio)

Panel $B.1$: Portfolio IO < 5%

$H_0 : \beta_r = 0, \rho = \hat{\rho}$	β_r	p-value	$Q_{0.05}$	$Q_{0.95}$	β_d	p-value	$Q_{0.05}$	$Q_{0.95}$
short-run	-0.02	0.7154	-0.10	0.12	-0.06	0.3600	-0.40	0.03
long-run	-0.83	0.8956	-1.61	1.64	-2.04	0.9200	-2.61	0.64
p joint hyp.		$\beta_r^{LR} \geq \hat{\beta}_r^{LR} \wedge \beta_d^{LR} \geq \hat{\beta}_d^{LR}$: 0.8956						
$H_0 : \beta_d = 0, \rho = \hat{\rho}$	β_r	p-value	$Q_{0.05}$	$Q_{0.95}$	β_d	p-value	$Q_{0.05}$	$Q_{0.95}$
short-run	-0.02	0.2346	-0.10	0.14	-0.06	0.6100	-0.39	0.05
long-run	-0.83	0.0806	-1.27	2.01	-2.04	0.0600	-2.27	1.01
p joint hyp.		$\beta_r^{LR} \leq \hat{\beta}_r^{LR} \wedge \beta_d^{LR} \leq \hat{\beta}_d^{LR}$: 0.0630						

Panel $B.2$: Portfolio IO > 5%

$H_0 : \beta_r = 0, \rho = \hat{\rho}$	β_r	p-value	$Q_{0.05}$	$Q_{0.95}$	β_d	p-value	$Q_{0.05}$	$Q_{0.95}$
short-run	0.09	0.2016	-0.08	0.19	-0.03	0.0500	-0.36	-0.03
long-run	0.80	0.0522	-0.72	0.81	-0.22	0.0600	-1.72	-0.19
p joint hyp.		$\beta_r^{LR} \geq \hat{\beta}_r^{LR} \wedge \beta_d^{LR} \geq \hat{\beta}_d^{LR}$: 0.0522						
$H_0 : \beta_d = 0, \rho = \hat{\rho}$	β_r	p-value	$Q_{0.05}$	$Q_{0.95}$	β_d	p-value	$Q_{0.05}$	$Q_{0.95}$
	β_r	p-value	$Q_{0.05}$	$Q_{0.95}$	β_d	p-value	$Q_{0.05}$	$Q_{0.95}$
short-run	0.09	0.3036	0.02	0.31	-0.03	0.6000	-0.27	0.08
long-run	0.80	0.5577	0.10	1.79	-0.22	0.5400	-0.90	0.79
p joint hyp.		$\beta_r^{LR} \leq \hat{\beta}_r^{LR} \wedge \beta_d^{LR} \leq \hat{\beta}_d^{LR}$: 0.5422						

Parameter estimates and simulated p-values for the estimates, sorted on book-to-market equity ratio. Q_p denotes the p-quantile of the simulated distribution. *,** and *** denote significance on the one, five and ten percent level, respectively (based on 5,000 simulations and with respect to the respective null hypothesis, i.e. whether the estimate is consistent with the null. **Stars indicate that it is not**). The rows labeled "$\widehat{\Pr}[\cdot|H_0]$" show the estimated probability of the noted events occuring given that the respective null hypothesis is true.

Table 8: **Simulated p-values and R^2, Cross-prediction**

Panel A: Raw returns

	p.-values				
	low IO	2	3	high IO	Mkt.
low IO	0.7139	0.4348	0.6540	0.4476	0.4895
2	0.3430	0.1282	0.2078	0.1391	0.1560
3	0.2583	0.2133	0.2686	0.2112	0.2282
high IO	0.4604	0.3775	0.4753	0.3922	0.4172
Mkt.	0.4360	0.3463	0.4277	0.3585	0.3736

	R^2				
	low IO	2	3	high IO	Mkt.
low IO	0.11	1.36	0.01	1.38	0.82
2	1.00	6.41	7.05	7.56	6.67
3	4.38	7.81	14.08	11.61	11.08
high IO	2.15	6.49	7.81	8.74	8.21
Mkt.	1.63	6.02	7.83	9.36	8.30

Panel B: Excess returns

	p.-values				
	low IO	2	3	high IO	Mkt.
low IO	0.8617	0.7196	0.8602	0.7049	0.7364
2	0.5407	0.3109	0.3981	0.3025	0.329
3	0.3802	0.3466	0.3888	0.3235	0.3403
high IO	0.5825	0.5373	0.6180	0.5379	0.558
Mkt.	0.5523	0.4966	0.5621	0.4898	0.5067

	R^2				
	low IO	2	3	high IO	Mkt.
low IO	1.93	0.07	2.55	0.06	0.24
2	0.01	1.26	1.33	1.79	1.43
3	1.46	2.84	6.51	5.26	4.96
high IO	0.28	1.81	2.28	3.07	2.82
Mkt.	0.11	1.51	2.21	3.35	2.79

Panel C: Dividend growth

	p.-values				
	low IO	2	3	high IO	Mkt.
low IO	0.7630	0.5704	0.3161	0.3073	0.442
2	0.5293	0.6902	0.5269	0.2955	0.4258
3	0.3117	0.6798	0.2003	0.1669	0.2302
high IO	0.3961	0.5577	0.5293	0.3352	0.4237
Mkt.	0.3730	0.5667	0.4027	0.2908	0.3727

	R^2				
	low IO	2	3	high IO	Mkt.
low IO	12.17	0.01	0.63	0.77	0.02
2	0.03	6.39	0.00	1.17	0.15
3	0.03	1.87	0.40	2.1	1.14
high IO	0.15	0.07	0.28	0.54	0.25
Mkt.	0.21	0.32	0.00	1.16	0.02

Simulated $p-$values of the null of $\beta_r^{j,i} = 0$ ($\beta_d^{j,i} = 0$, respectively) for the predictive regressions of log excess returns and log dividend growth on the log dividend yields of various portfolios. The entry in row i and column j is the p-value of the respective null hypothesis in a regression of the return (or dividend growth) of the portfolio in column j on the dividend yield in row i. The portfolios are those of the baseline specification and Mkt. is the CRSP value-weighted market portfolio. The frequency is yearly.

Table 9: **Regression coefficients, high and low valuation**

	β_r^{high}	β_r^{low}	$\Delta\beta_r$
low IO	0.026	0.022	0.004
	(0.60)	(0.82)	(0.42)
2	0.122	0.026	0.096
	(2.90)	(1.09)	(11.71)
3	0.058	0.097	-0.040
	(1.62)	(2.35)	(-4.20)
high IO	0.118	0.101	0.017
	(2.30)	(2.05)	(1.37)

Slope coefficients from predictive regressions of excess returns on the dividend yield sample divided by above (high *dp*) or below (low *dp*) median dividend yield. Quarterly observations, raw returns. Numbers in brackets are standard $t-$statistics computed under the assumption that the estimators for the respective β_r^{high} and β_r^{low} are uncorrelated.

Table 10: **Standard deviation and means of average betas**

	low IO	2	3	high IO
		Std.		
Mkt	1.1603	0.9713	0.8464	0.7547
SMB	1.7248	1.4056	1.1806	1.0370
HML	1.8979	1.5271	1.3143	1.1847
		Mean		
Mkt	0.8665	0.9834	1.063	1.1262
SMB	1.0401	0.8855	0.751	0.6114
HML	0.1887	0.1914	0.1474	0.0892

Time-series means of quarterly cross-sectionally equally-weighted average stock level standard deviation and means of individual stock betas with respect to the Market (Mkt), Small-minus-Big (SMB) and Value-minus-growth (HML) factor as provided by Kenneth French. Betas are estimated on a rolling basis from 24 monthly individual stock returns. Note: Using value-weighted measures shows the expected negative SMB betas for high IO stocks.

Table 11: **Predictive regressions with mutual fund cash holdings (MFCH)**

	$R_{t+1} = \alpha + \beta \cdot MFCH_t$		$R^e_{t+1} = \alpha + \beta \cdot MFCH_t$		$R^e_{t+1} = \alpha + \beta \cdot \Delta MFCH_t$	
	β	R^2	β	R^2	β	R^2
low IO	0.0153 *p.* 0.1345	0.04	0.0082 *p.* 0.3156	0.01	0.0084 *p.* 0.3288	0.00
2	0.0215** *p.* 0.037	0.12	0.0139 *p.* 0.1363	0.05	0.0186 *p.* 0.1244	0.03
3	0.0221* *p.* 0.0957	0.12	0.0146 *p.* 0.1958	0.06	0.0143 *p.* 0.2629	0.01
high IO	0.0311** *p.* 0.0336	0.23	0.0231* *p.* 0.0915	0.14	0.0344** *p.* 0.0406	0.09
Mkt.	0.0267** *p.* 0.0488	0.17	0.019 *p.* 0.1250	0.09	0.0264* *p.* 0.0960	0.05

Slope coefficients of the predictive regression of year $t+1$ (excess) returns on the four IO-sorted portfolios on average mutual fund cash-holdings (MFCH) at the end of year t. Simulated $p-$values $p.$ are computed using 5000 artificial data sets generated under the null of no predictability:

$$R_{t+1} = \bar{R} + \epsilon^r_{t+1}$$
$$X_{t+1} = \rho X_t + \epsilon^X_{t+1},$$

where $\bar{R}$, ρ and $Cov(\epsilon^r, \epsilon^r)$ are as estimated from the actual data. $E[\epsilon^r] = E[\epsilon^x] = 0$. The $MFCH$ sample is from 1979 to 2007 to avoid the issues with the CRSP survivorship-bias free mutual fund data set documented in Chernenko and Sunderam (2016). Equity mutual funds are defined as all funds that in the last two years invested on average at least 75% in common or preferred stocks (in absolute value, i.e. this includes short positions).

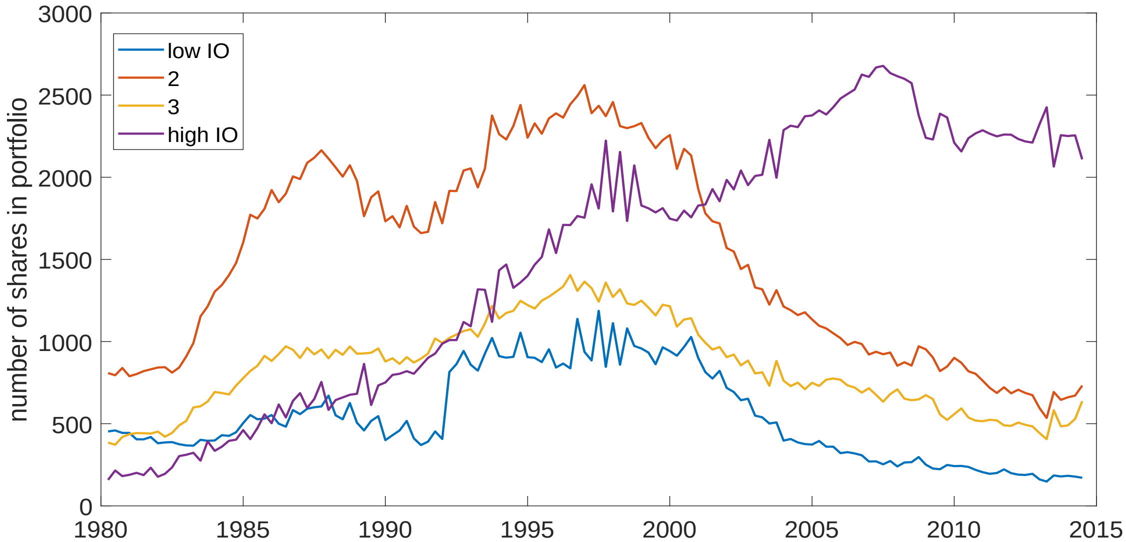


Figure 1: **Number of stocks in each of the four portfolios from the baseline setting.**

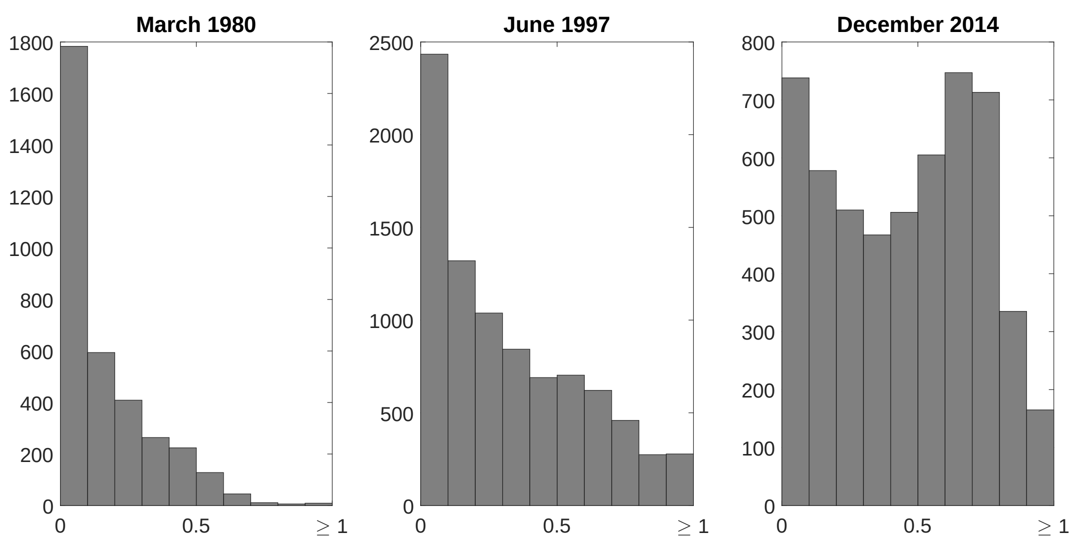


Figure 2: **Distribution of IO share over time**

The figure shows the distribution of IO share across stocks at the beginning (first quarter 1980), in the middle (second quarter 1997) and at the end of the sample (fourth quarter 2014).

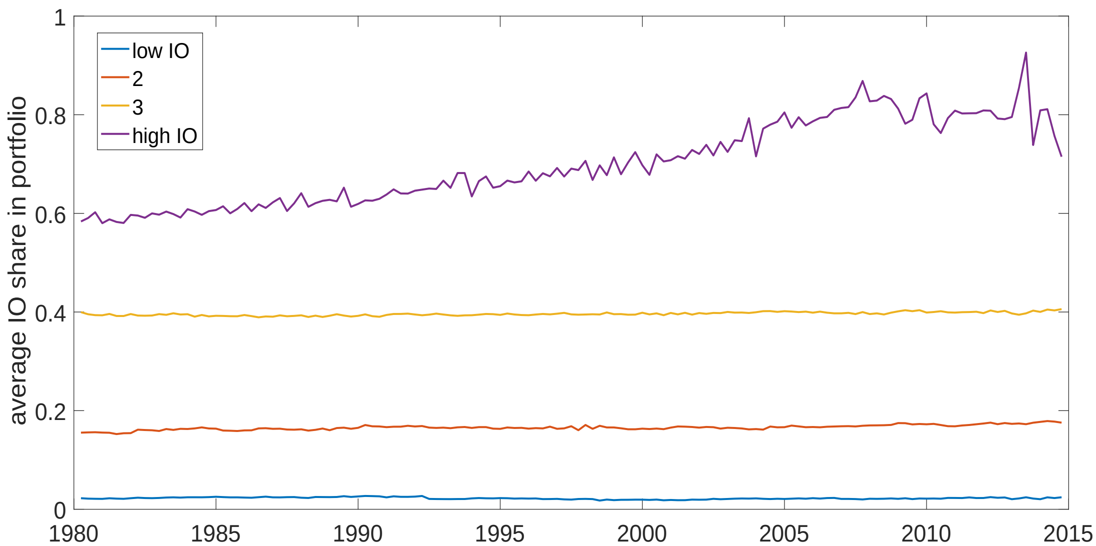
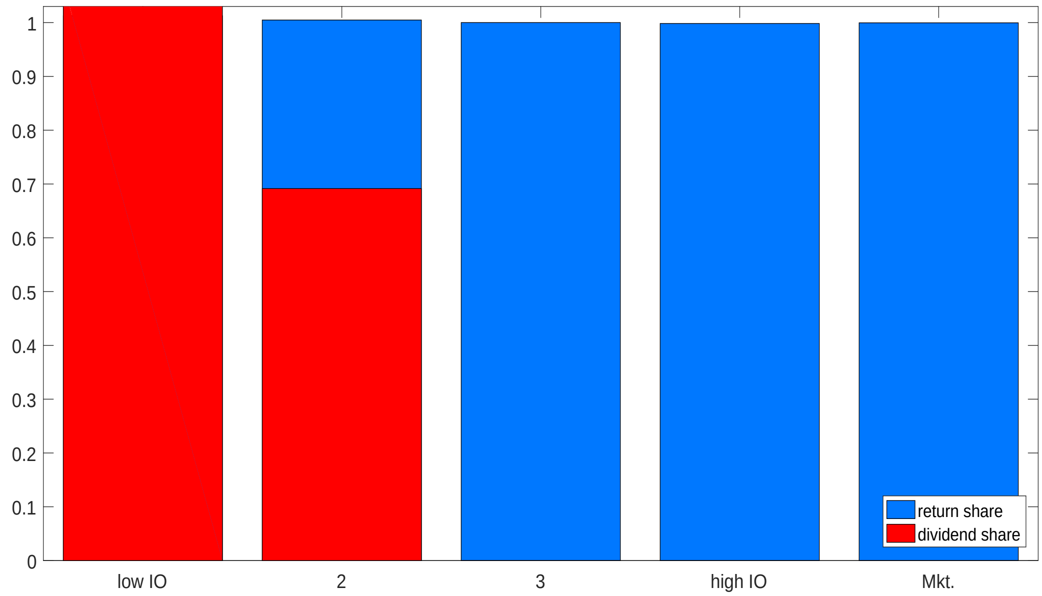


Figure 3: **Average IO in the respective IO-sorted portfolios**

Panel A : Yearly frequency



Panel B : Quarterly frequency

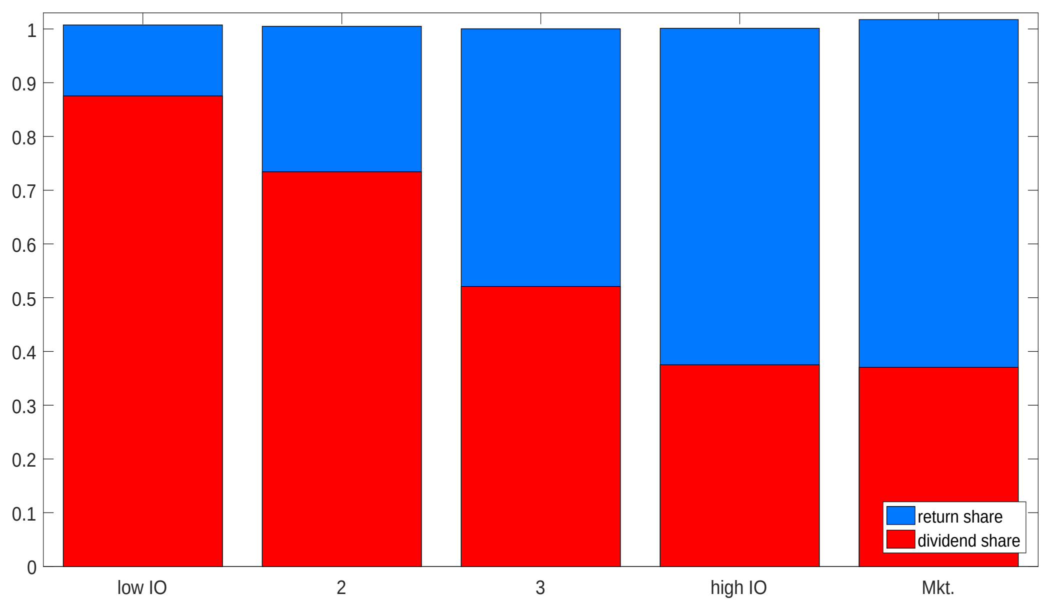


Figure 4: **Shares of variation**

Share of variation due to either dividend growth or returns as computed with long-run coefficients $\beta_r^{LR} = \frac{\beta_r}{1-k_1\rho}$ and $|\beta_d^{LR}| = \frac{|\beta_r|}{1-k_1\rho}$, respectively, for each of the four IO portfolios and the market. The numbers can be found in Table 1.

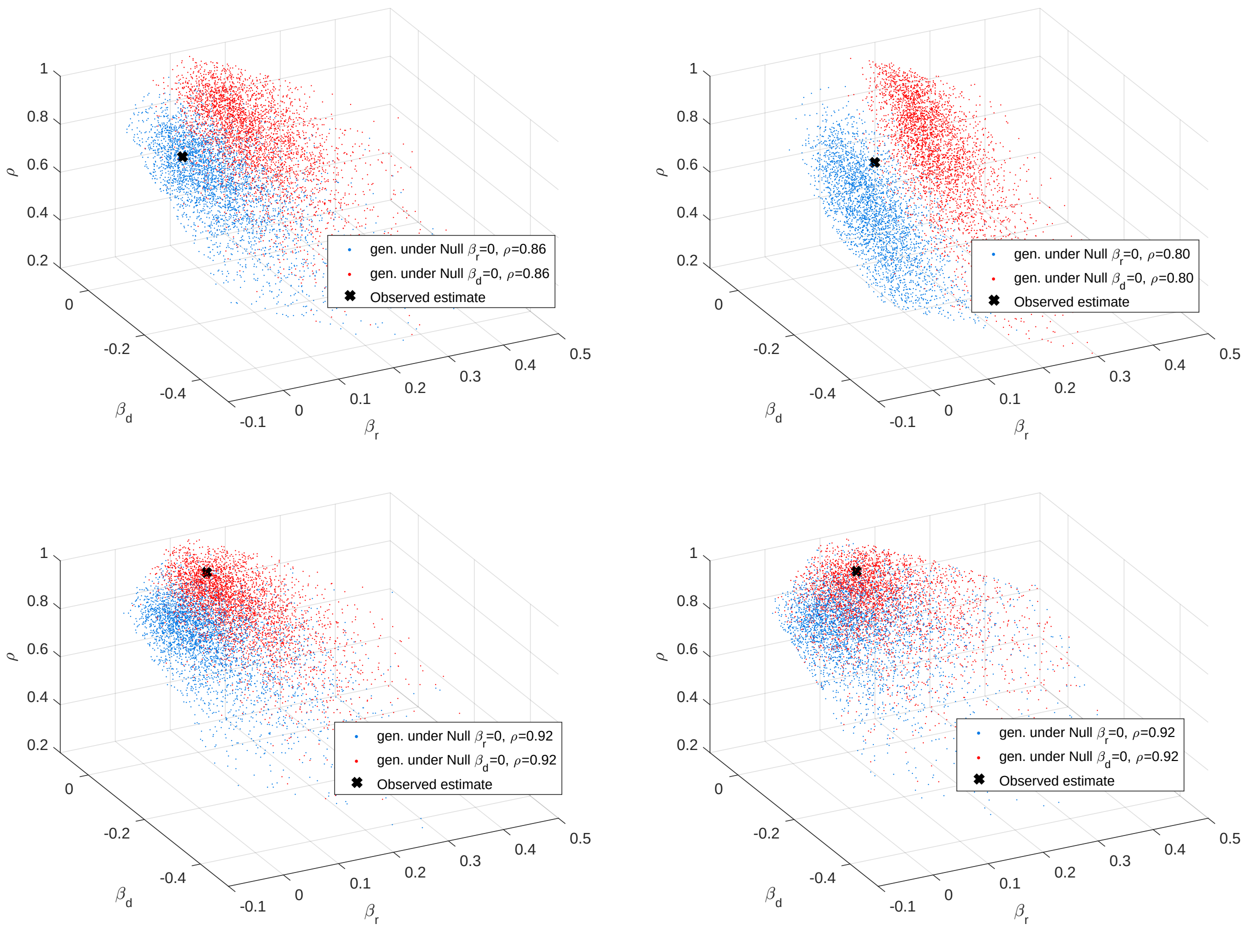
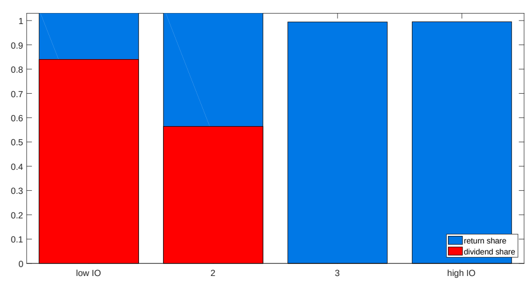


Figure 5: **Simulated slope coefficients under the null hypotheses**

The black crosses indicate the estimates from the actual data. From left to right and top to bottom: first row: 1: low IO, 2. Second row: 3, 4: high IO. For instructive purposes and visibility, the presented results are at quarterly frequency with only 3500 simulated time series.



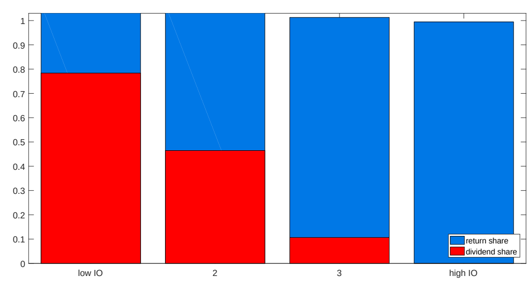


Figure 6: **Share of variation, two- and three- year predictive regressions**

Share of variation due to either dividend growth or returns as computed with long-run coefficients $\beta_r^{LR} = \frac{\beta_r}{1-k_1\rho}$ and $|\beta_d^{LR}| = \frac{|\beta_r|}{1-k_1\rho}$, respectively, for each of the four IO-sorted portfolios with two (left) and three (right) year horizons.

A Log-linearization

Starting from $R_{t+1} = \frac{P_{t+1}+D_{t+1}}{P_t} = \frac{\frac{P_{t+1}}{D_t}+\frac{D_{t+1}}{D_t}}{\frac{P_t}{D_t}} = \frac{\left(1+\frac{P_{t+1}}{D_{t+1}}\right)\frac{D_{t+1}}{D_t}}{\frac{P_t}{D_t}}$, then taking logs and linearizing the middle term around the long run mean of $p-d = \log\left(\frac{\overline{P}}{D}\right) \equiv \overline{pd}$ yields

$$\begin{aligned} r_{t+1} &= \Delta d_{t+1} + \ln\left(1+\exp\left(p_{t+1}-d_{t+1}\right)\right) - pd_t \\ &\approx k_0 + k_1\left(p_{t+1}-d_{t+1}\right) + \Delta d_{t+1} - \left(p_t - d_t\right) \end{aligned}$$

where lower case letters denote logs and $k_0 = \ln\left(1+\exp\left(\overline{pd}\right)\right) - \frac{\overline{P/D}}{1+\overline{P/D}}\overline{pd}$ and $k_1 = \frac{\overline{P/D}}{1+\overline{P/D}}$. Throughout the paper, I use dp and pd as a shorthand notation for $d-p$ and $p-d$, respectively.

B Stambaugh bias

As is shown in Stambaugh (1999), when the error vector u is decomposed into a component that is correlated with v and one that is strictly exogenous,

$$u = \frac{Cov(u,v)}{Var(v)}v + \varepsilon = \frac{\sigma_{u,v}}{\sigma_v^2}v + \varepsilon, \tag{B.1}$$

one gets the result that the bias of the OLS estimator for β_r depends on the bias of the estimator for ρ. In the following, it is useful to adopt matrix notation for ease of exposition. Define $X = [1 \ dp]$, $\beta = [\mu_r \ \beta_r]'$ and $\mathrm{P} = [\alpha \ \rho]'$, then the bias is

$$\hat{\beta} - \beta = \frac{\sigma_{uv}}{\sigma_v^2}(X'X)^{-1}X'v + (X'X)^{-1}X'\varepsilon \tag{B.2}$$

$$= \frac{\sigma_{uv}}{\sigma_v^2}(\hat{\mathrm{P}} - \mathrm{P}) + (X'X)^{-1}X'\varepsilon, \tag{B.3}$$

where hats denote estimated coefficients. By definition, $E[(X'X)^{-1}X'\varepsilon] = 0$. As is well known (see for example Stambaugh (1999)), the usual OLS estimator for P is biased, albeit consistent:

$$E[\hat{\rho} - \rho] = -\frac{1+3\rho}{T}. \tag{B.4}$$

Thus, the resulting bias in the estimator for the slope is

$$E[\hat{\beta}_r - \beta_r] = -\frac{\sigma_{uv}}{\sigma_v^2}\frac{1+3\rho}{T}. \tag{B.5}$$

C Long-run coefficients

To see how the long-run coefficients in Equation (9) are derived, consider Equation (3). Iterating forward yields

$$
\begin{aligned}
r_{t+2} &= \mu_r + \beta_r(\alpha + \rho\ dp_t + v_{t+1}) + u_{t+2}, \\
r_{t+3} &= \mu_r + \beta_r(\alpha + \rho(\alpha + \rho dp_t + v_{t+1}) + v_{t+2}) + u_{t+3} \\
&\vdots \\
r_{t+j} &= const + \beta_r \rho^{j-1} dp_t + \beta_r \sum_{i=1}^{j} \rho^i v_{t+j-i} + u_{t+j}.
\end{aligned}
$$

Using the definition for the infinite *discounted long-run return* first introduced in Campbell and Shiller (1988),

$$
r^{LR} \equiv \sum_{j=1}^{\infty} k_1^{j-1} r_{t+j},
$$

the regression slope coefficient of r^{LR} on dp_t results as

$$
\beta_r^{LR} = \sum_{j=1}^{\infty} k_1^{j-1} \beta_r \rho^{j-1} = \frac{\beta_r}{1 - k_1 \rho},
$$

where the last equality goes through if $\rho < k_1^{-1}$ (which essentially means that the PD-ratio is not explosive). Similarly, one can derive that

$$
\beta_d^{LR} = \frac{\beta_d}{1 - k_1 \rho}.
$$

Moreover, consider the log-linearized return identity (1), demeaned and solved for dp_t

$$
\widetilde{dp}_t = k_1 \widetilde{dp}_{t+1} - \Delta \tilde{d}_{t+1} + \tilde{r}_{t+1}, \tag{C.1}
$$

where tilde denotes demeaned variables. Using that $\widetilde{dp}_{t+1} = k_1 \widetilde{dp}_{t+2} + \widetilde{r}_{t+2} - \Delta \widetilde{d}_{t+2}$, one can iterate the identity forward, i.e.

$$
\begin{aligned}
\widetilde{dp}_t &= k_1(k_1 \widetilde{dp}_{t+2} + \widetilde{r}_{t+2} - \Delta \widetilde{d}_{t+2}) + \widetilde{r}_{t+1} - \Delta \widetilde{d}_{t+1} \\
&= k_1(k_1(k_1 \widetilde{dp}_{t+3} + \widetilde{r}_{t+3} - \Delta \widetilde{d}_{t+3}) + \widetilde{r}_{t+2} - \Delta \widetilde{d}_{t+2}) + \tilde{r}_{t+1} - \Delta \widetilde{d}_{t+1} \\
&= \sum_{j=1}^{3} k_1^{j-1} \widetilde{r}_{t+j} - \sum_{j=1}^{3} k_1^{j-1} \Delta \widetilde{d}_{t+j} + k_1^3 \widetilde{dp}_{t+3},
\end{aligned}
$$

and so on until finally

$$
\widetilde{dp}_t = \sum_{j=1}^{\infty} k_1^{j-1} \tilde{r}_{t+j} - \sum_{j=1}^{\infty} k_1^{j-1} \Delta \widetilde{d}_{t+j}. \tag{C.2}
$$

Computing the covariance with dp_t for the left and right hand side of (C.2), yields that:

$$
\text{Var}(dp_t) = \text{Cov}\left(\sum_{j=1}^{\infty} k_1^{j-1} \widetilde{r}_{t+j}, \widetilde{dp}_t \right) - \text{Cov}\left(\sum_{j=1}^{\infty} k_1^{j-1} \Delta \widetilde{d}_{t+j}, \widetilde{dp}_t \right). \tag{C.3}
$$

where obviously, $\text{Var}(dp_t) = \text{Var}(\widetilde{dp_t})$. Recall the definition of β_j^{LR} as slope coefficient of regressions on dp_t. Dividing by (C.3) by $\text{Var}(dp_t)$ yields

$$1 = \beta_r^{LR} - \beta_d^{LR}, \tag{C.4}$$

where β_d^{LR} will typically be negative. Therefore, the absolute value of the long-run coefficients describe the share of the variance of the dp-ratio that can be attributed to either returns or dividends, providing suitable measures for predictability in the relative sense of attributing variation in the dividend yield to variation in either returns or dividend growth.

D Direct regressions

Rather than inferring long-run coefficients by imposing the structure of the vector autoregression (3) to (5), one can run direct regressions of weighted returns and dividend growth:

$$\sum_{j=1}^{K} k_1^{j-1} r_{t+j} = \mu_r^K + \beta_r^K \; dp_t + u_{t+1} \tag{D.1}$$

$$\sum_{j=1}^{K} k_1^{j-1} \Delta d_{t+j} = \mu_d^K + \beta_d^K \; dp_t + w_{t+1} \tag{D.2}$$

$$k_1^j dp_{t+j} = \alpha^K + \beta_{dp}^K \; dp_t + v_{t+1}. \tag{D.3}$$

It holds that $\beta_r^K - \beta_d^K + \beta_{dp}^K \approx 1$. The coefficients with horizon K are plotted in Figure D below. While these results should be treated with caution due to the even short sample (in order to have the same measure of k_1, I only go from 1980 to 2009 in terms of formation periods), the results from Section 4 are confirmed: Dominance of dividend growth predictability in the low IO portfolio, somewhat mixed results for portfolio 2 and 3, and overwhelming return predictability in portfolio 4. The left figure is somewhat at odds with the characterization of the low IO dividend yield as a fairly low persistence AR(1) process. This is due to the cumulative weighted 4-year return. In particular, this result is driven by the years 1984 and 1988. The considerable effect this has is an unfortunate effect of a short sample:

The stocks that constituted the low IO portfolio in 1984 had fairly low valuations. Four years later, these stocks (that are not to be confused with the stocks in the low IO portfolio in 1988) had huge dividend growth.[19] Hence, this massively drove down the estimate for β_d^4. At the same time, the low valuations in 1984 were supposed to predict high returns in 1988. However, stocks did not do particularly well that year. In particular, the stocks that had made up the low IO portfolio in 1984 actually had very low returns of about -7%, leading to no increase in β_r^4 and consequently to an estimate of a very persistent dp transition over four years.

One can come up with explanations for why this may have been the case. However, with one datapoint only, these remain speculations. Unfortunately, in a short time series like the one at hand, rare events like this can have sizable effects on estimates.

[19] One reason for this is the October 1987 crash and working under the assumption of reinvested dividends, for a thorough discussion of this effect, see Chen (2009).

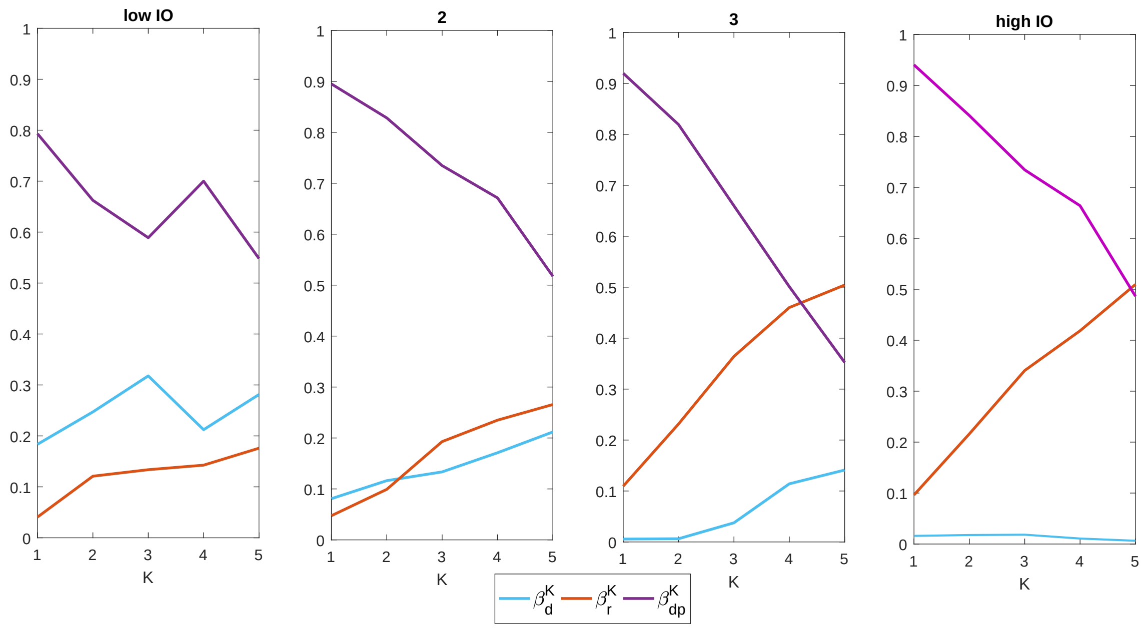


Figure 7: **Multi-period coefficients, direct estimate**

Multi-period regression coefficients as in Equations D.1 to D.3. The formation period sample is from 1980 to 2009.

E Model

E.1 Model solution

Unit IES stochastic differential utility implies a constant wealth-consumption ratio. Together with the aforementioned cash flow dynamics, it follows that the maximization problem of any agent in the model is:

$$\max_{\delta_t^{\xi,\phi}} V_t$$

$$\text{s.t.}\quad dW_t = W_t\phi_t\left(\mu_{c,t}dt + \sigma_c dB_{\delta,t} + \frac{\delta_t^{\xi}}{W_t}\right) + W_t(1-\phi_t)r_t - c_t + W_{t^-}(e^{-\xi L}-1)dN_t \tag{E.1}$$

$$(\phi - 1)\mathbb{1}_h \leq 0 \tag{E.2}$$

with $L = 0$ if the optimizing agent is the household. $\mathbb{1}_h$ is an indicator function that is one if the optimizing agent is the household. The unit IES implies that the wealth-consumption ratio is constant, i.e. $\frac{\delta_t^{\xi}}{W_t} = \bar{c}$. This results in the following Hamilton-Jacobi-Bellman (HJB) equation:

$$\begin{aligned}\sup_{c_t,\phi_t}\Bigg\{ & f(c_t,J) + J_W W_t\left(\phi_t\mu_{c,t} + \phi_t\bar{c} + (1-\phi_t)r_t\right) - J_W c_t + \frac{1}{2}J_{WW}W_t^2\phi_t^2\sigma_c^2 \\ & + \lambda_t\Big[J\left((1+\phi_t(\exp(-L\xi)-1))W_t,\mu_t,\lambda_t\right) - J\left(W,\mu_t,\lambda_t\right)\Big] \\ & + J_\mu\kappa_\mu(\bar{\mu}-\mu_t) + \frac{1}{2}J_{\mu\mu}\sigma_\mu^2 + J_\lambda\kappa(\bar{\lambda}-\lambda_t) + \frac{1}{2}J_{\lambda\lambda}\lambda_t\sigma_\lambda^2 + \eta_t(1-\phi_t)\mathbb{1}_h\Bigg\} = 0\end{aligned} \tag{E.3}$$

where for the household, it holds that $L = 0$ and where η_t is the Lagrange multiplier associated with the leverage constraint of the household. The first order conditions (FOCs) read:

$$f_c(c,V) = \beta(1-\gamma)V_t c_t^{-1} = J_W \tag{E.4}$$

$$J_W W_t(\mu_{c,t} + \bar{c} - r_t - \tilde{\eta}_t) + \phi J_{WW}W_t^2\sigma_c^2 + \lambda_t J_W W_t(e^{-\xi L}-1) = 0, \tag{E.5}$$

where $\tilde{\eta}_t = \frac{\eta_t}{J_W W_t}\mathbb{1}_h$. A plausible guess for the value function is

$$J(W_t,\mu_t,\lambda_t) = \frac{W_t^{1-\gamma}}{1-\gamma}\, g\left(\lambda_t,\mu_t\right) = \frac{W_t^{1-\gamma}}{1-\gamma}\exp\left(A_0 + A_1\lambda_t + A_2\mu_t\right). \tag{E.6}$$

It is verified in Appendix E.3. With $\gamma > 1$, it holds that $A_1 < 0$ and $A_2 > 0$. Plugging the guess into the FOCs gives the following policy functions (see Appendix E.2)

$$c_t = \beta W_t \tag{E.7}$$

$$\phi_i = \frac{\mu_{c,t} + \beta + \lambda_t e^{\gamma\xi L}(e^{-\xi L}-1) - r_t}{\gamma\sigma_c^2} \tag{E.8}$$

$$\phi_h = \frac{\mu_{c,t} + \beta - \tilde{\eta}_t - r_t}{\gamma\sigma_c^2}, \tag{E.9}$$

where the subscripts i and h denote the policy functions of the institution and the household, respectively. Market clearing on the consumption good and asset market, in conjunction with the leverage constraint requires

$$\phi_h = \phi_i = 1. \tag{E.10}$$

In equilibrium, the institution and the household will hold measures of the stock market proportional to their relative wealth. Solving for the market clearing shadow interest rate yields

$$r_t = \mu_{c,t} + \beta - \gamma\sigma_c^2 + \lambda_t e^{\gamma\xi L}(e^{-\xi L} - 1) \text{ and } \tilde{\eta}_t = -\lambda_t e^{\gamma\xi L}(e^{-\xi L} - 1). \tag{E.11}$$

Hence, the shadow rate for the household is $r_t^h = r_t + \tilde{\eta}_t = \mu_t + \beta - \gamma\sigma_\delta^2$. The state price density for the institution is given by

$$\pi_t = \exp\left(\int_0^t f_V(C_s, V_s)ds\right) f_c(C_t, V_t), \tag{E.12}$$

Using that in equilibrium, $V_t = J(\frac{C_t}{\beta}, \lambda_t, \mu_t)$, and applying Ito's lemma, one gets that the dynamics of π_t are given by

$$\frac{d\pi_t}{\pi_{t-}} = \mu_\pi dt - \gamma\sigma_\delta dB_t^\delta + A_2\sigma_\mu dB_t^\mu + \sqrt{\lambda_t}A_1\sigma_\lambda dB_t^\lambda + \left(e^{\xi\gamma L} - 1\right)dN_t, \tag{E.13}$$

where $A_2 < 0$ and $A_1 > 0$ (if $L > 0$, see Appendix E.3). Because the negative of the expected change in the pricing kernel must equal the risk-free rate, it holds that

$$\mu_\pi = -r_t - \lambda_t\left(e^{\gamma\xi L} - 1\right). \tag{E.14}$$

Time-varying λ in the fourth term on the right hand side of Equation (E.13) will generate time-varying covariation between the pricing kernel of the institution and the returns on any asset that loads on B_t^λ. It remains to be shown that returns of institutionally held stocks load on B_t^λ. I do so by computing the price-dividend ratio of these stocks.

The price of the dividend claim is then given by:

$$P_t = E_t\left[\int_t^\infty \frac{\pi_s}{\pi_t}\tilde{\delta}_s\, ds\right] = \int_t^\infty E_t\left[\frac{\pi_s}{\pi_t}\tilde{\delta}_s\right] ds = \int_t^\infty H_{s,t}\, ds. \tag{E.15}$$

Here, $H_{s,t}$ can be interpreted as the time t price of a 'zero coupon equity' claim. In Appendix E.4, I show that $H_{s,t}$ is of the form

$$H_{s,t} = \tilde{\delta}_t \exp\left(B_0(\tau) + B_1(\tau)\lambda_t + B_2(\tau)\mu_t\right) = \tilde{\delta}_t f(\tau; \mu_t, \lambda_t), \tag{E.16}$$

where $B_2 > 0$ and, if $L > 0$, $B_1 < 0$ with with $\tau = s - t$. So in the case of the institution, a higher redemption intensity λ leads to a higher dividend yield. For the household, $L = 0$ and $B_1 = 0$, i.e. the dividend yield is not affected by λ. B_2 is positive, independent of L, so dividend yields decrease in mean dividend growth μ, no matter if a stock is held by the household or the institution. Integrating over the $H_{s,t}$ and dividing by time t dividend level δ_t then yields the price-dividend ratio. Note that empirically, an econometrician in the model economy would not observe

$$F_t(\lambda_t, \mu_t) = \int_t^\infty f(\tau; \mu_t, \lambda_t)\, d\tau \tag{E.17}$$

as price-dividend ratio but rather

$$\tilde{F}_t(\lambda_t; \mu_t; \delta_{s_i}, i = 1, \ldots, N_t) = F_t(\lambda_t, \mu_t) \prod_{0 < s_i \leq t} \left(e^{-L}\right), \tag{E.18}$$

where s_i denotes the time of the i^{th} jump. In expectation, the size of the divergence between F and $\tilde{F}$ depends on the the intensity λ and the loss parameter L. Importantly, it holds that F inherits the properties from B_1

and B_2 that $\frac{\partial F}{\partial \lambda} < 0$ and $\frac{\partial F}{\partial \mu} > 0$. Moreover, because the market price of innovations to λ is negative and the size of innovations depends on λ itself, ceteris paribus, a high dividend yield predicts higher risk premia:

$$r_t^j - r_t = \text{const.} - \underbrace{A_1 \frac{F'_\lambda}{F} \sigma_\lambda^2}_{<0} \lambda_t + \lambda_t \underbrace{\left[(e^{\gamma \xi L} - 1)(1 - e^{-L}) \right]}_{>0}. \tag{E.19}$$

Were the same stocks held by the household, excess returns would not be predictable by the dividend yield, because for the households, $L = A_1 = 0$ and the dividend yield varies only with changes in expected dividend growth. Consequently, for sufficiently large L, sufficiently volatile λ and sufficiently weak time-variation in μ_t, the dividend yield of the household's portfolio predicts dividend growth and that of the institutionally held portfolio predicts returns. In line with the evidence on the predictive power of MFCH from Section 6, λ positively predicts returns in the model.

Finally, when stocks held by the institution have higher expected returns, then they are undervalued from the household sector's perspective. Hence, the household would want to buy these stocks. If the household bought those stocks, their expected returns would drop to the level of household-held stocks because they are only higher while they compensate the institution for the institution-specific risks it faces. In other words, the individual investors would trade against and "correct the misvaluation". The leverage constraint of the household implies that the household cannot raise sufficient capital to trade away the perceived undervaluation. The only way for the household to raise funds in order to buy the undervalued institutionally-held stock would be to sell the stocks in its portfolio to the institution who is the only potential buyer in the economy but who would not buy the claims at the price that the household would ask for. Put differently, other than with respect to their ownership, all stocks in the economy are identical and thus - given the leverage constraint - the agents in the economy do not exchange them.

E.2 Policy functions

The first order condition (FOC) for consumption (E.4) with the guess for the value function implies:

$$\beta(1-\gamma) \frac{W_t^{1-\gamma}}{1-\gamma} g c_t^{-1} = W_t^{-\gamma} g \tag{E.20}$$

$$\beta(1-\gamma) \frac{W_t^{-\gamma}}{1-\gamma} g \left(\frac{c_t}{W_t} \right)^{-1} = W_t^{-\gamma} g \tag{E.21}$$

$$c_t = \beta W_t. \tag{E.22}$$

E.3 Value function

Verifying the guess (E.6) for the value function (E.3). Plugging the optimal policy functions derived from the FOCs and market clearing into (E.3) yields

$$\begin{aligned} &\Big\{ f(c_t, J) + J_W W_t \mu_{c,t} + \frac{1}{2} J_{WW} W_t^2 \sigma_c^2 + \lambda_t \left[J(W_t e^{-L\xi}, \mu_t, \lambda_t) - J(W, \mu_t, \lambda_t) \right] \\ &+ J_\mu \kappa_\mu (\bar{\mu} - \mu_t) + \frac{1}{2} J_{\mu\mu} \sigma_\mu^2 + J_\lambda \kappa (\bar{\lambda} - \lambda_t) + \frac{1}{2} J_{\lambda\lambda} \lambda_t \sigma_\lambda^2 \Big\} = 0. \end{aligned} \tag{E.23}$$

Plugging in the optimal consumption rule $c = \beta W$ and the guess for the functional form (E.6) into (24) gives

$$f(c_t, V_t) = \beta(1-\gamma)\frac{W^{1-\gamma}g}{1-\gamma}\ln(\beta W_t) - \beta\frac{W^{1-\gamma}g}{1-\gamma}\ln\left((1-\gamma)\frac{W^{1-\gamma}g}{1-\gamma}\right) \tag{E.24}$$

$$= \beta W^{1-\gamma}g\left(\ln\beta - \frac{\ln g}{1-\gamma}\right). \tag{E.25}$$

Plugging in the guess for the functional form (E.6) and (E.25) into (E.23) gives:

$$\beta W^{1-\gamma}g\left(\ln\beta - \frac{\ln g}{1-\gamma}\right) + W_t^{-\gamma}W_t g\mu_{1,t} - \frac{\gamma}{2}W_t^{-\gamma-1}W_t^2 g\sigma_\delta^2 + \lambda_t\left[J(W_te^{-L\xi}, \mu_t, \lambda_t) - J(W, \mu_t, \lambda_t)\right]$$
$$+ \frac{W^{1-\gamma}}{1-\gamma}g'_\mu\kappa_\mu(\bar{\mu} - \mu_t) + \frac{1}{2}\frac{W^{1-\gamma}}{1-\gamma}g''_{\mu\mu}\sigma_\mu^2 + \frac{W^{1-\gamma}}{1-\gamma}g'_\lambda\kappa(\bar{\lambda} - \lambda_t) + \frac{1}{2}\frac{W^{1-\gamma}}{1-\gamma}g''_{\lambda\lambda}\lambda_t\sigma_\lambda^2 = 0. \tag{E.26}$$

Note that

$$J(W_te^{-L\xi}, \mu_t, \lambda_t) - J(W_t, \mu_t, \lambda_t) \tag{E.27}$$

$$= (W_te^{-L\xi})^{1-\gamma}\frac{1}{1-\gamma}g - W_t^{1-\gamma}\frac{1}{1-\gamma}g \tag{E.28}$$

$$= W_t^{1-\gamma}\frac{g}{1-\gamma}\left(e^{-L\xi(1-\gamma)} - 1\right). \tag{E.29}$$

This implies that

$$\beta g\left(\ln\beta - \frac{\ln g}{1-\gamma}\right) + \mu_{c,t}g - \frac{\gamma}{2}g\sigma_\delta^2 + \lambda_t\frac{g}{1-\gamma}\left(e^{-L\xi(1-\gamma)} - 1\right)$$
$$+ \frac{1}{1-\gamma}g'_\mu\kappa_\mu(\bar{\mu} - \mu_t) + \frac{1}{2}\frac{1}{1-\gamma}g''_{\mu\mu}\sigma_\mu^2 + \frac{1}{1-\gamma}g'_\lambda\kappa(\bar{\lambda} - \lambda_t) + \frac{1}{2}\frac{1}{1-\gamma}g''_{\lambda\lambda}\lambda_t\sigma_\lambda^2 = 0. \tag{E.30}$$

Plugging in the affine guess for g in (E.6) and dividing by g yields

$$\beta\left(\ln\beta - \frac{A_0 + A_1\lambda_t + A_2\mu_t}{1-\gamma}\right) + \mu_{c,t} - \frac{\gamma}{2}\sigma_\delta^2 + \lambda_t\frac{1}{1-\gamma}\left(e^{-L\xi(1-\gamma)} - 1\right)$$
$$+ \frac{1}{1-\gamma}A_2\kappa_\mu(\bar{\mu} - \mu_t) + \frac{1}{2}\frac{1}{1-\gamma}A_2^2\sigma_\mu^2 + \frac{1}{1-\gamma}A_1\kappa(\bar{\lambda} - \lambda_t) + \frac{1}{2}\frac{1}{1-\gamma}A_1^2\lambda_t\sigma_\lambda^2 = 0. \tag{E.31}$$

Collecting terms in λ_t, μ_t and constants yields a system of equations in A_0, A_1 and A_2

$$\beta\ln\beta - \frac{\beta}{1-\gamma}A_0 - \frac{\gamma}{2}\sigma_\delta + \frac{1}{1-\gamma}A_2\kappa_\mu\bar{\mu} + \frac{1}{2}\frac{1}{1-\gamma}A_2^2\sigma_\mu^2 + \frac{1}{1-\gamma}A_1\kappa\bar{\lambda} = 0 \tag{E.32}$$

$$\xi - \frac{\beta}{1-\gamma}A_2 - \frac{1}{1-\gamma}A_2\kappa_\mu = 0 \tag{E.33}$$

$$-\frac{\beta}{1-\gamma}A_1 + \frac{1}{1-\gamma}\left(e^{-L\xi(1-\gamma)} - 1\right) - \frac{1}{1-\gamma}A_1\kappa + \frac{1}{2}\frac{1}{1-\gamma}A_1^2\sigma_\lambda^2 = 0. \tag{E.34}$$

The solutions are

$$A_2 = \frac{\xi - \gamma}{\kappa_\mu + \beta} \tag{E.35}$$

$$A_1 = \frac{\kappa + \beta}{\sigma_\lambda^2} \pm \sqrt{\frac{(\kappa+\beta)^2}{\sigma_\lambda^4} - 2\frac{e^{-L\xi(1-\gamma)} - 1}{\sigma_\lambda^2}} \tag{E.36}$$

$$A_0 = \frac{1-\gamma}{\beta}\left(\beta \ln \beta - \frac{\gamma}{2}\sigma_\delta + \frac{1}{1-\gamma}A_2\kappa_\mu\bar{\mu} + \frac{1}{2}\frac{1}{1-\gamma}A_2^2\sigma_\mu^2 + \frac{1}{1-\gamma}A_1\kappa\bar{\lambda}\right). \tag{E.37}$$

In (E.36), the term under the square root must be nonnegative, i.e. $(\kappa+\beta)^2 \leq 2\left(e^{-L\xi(1-\gamma)} - 1\right)\sigma_\lambda^2$. The only reasonable choice for (E.36) is $A_1 = \frac{\kappa+\beta}{\sigma_\lambda^2} - \sqrt{\frac{(\kappa+\beta)^2}{\sigma_\lambda^4} - 2\frac{e^{-L\xi(1-\gamma)}-1}{\sigma_\lambda^2}} > 0$ for $\gamma > 1$. As in Wachter (2013), the thought experiment that if $L = 0$ (which is the case for the household), the effect of λ should be zero, yields that the sign of the square-root term should be negative. The term under the square-root most be nonnegative for a solution A_1 to exist. This places a joint constraint on all primitive parameters involved.

E.4 Dividend claims

'Zero coupon' equity claims $H_{s,t}$ as in Equation E.15 are conjectured to be of the form

$$H_{s,\tau} = H(\tilde{\delta}_t, \lambda_t, \mu_t, \tau) = \delta_t \exp\left(B_0(\tau) + B_1(\tau)\lambda_t + B_2(\tau)\mu_t\right) \tag{E.38}$$

with $\tau = s - t$. Generically, dynamics are given by

$$dH = H_t\left(\mu_H\ dt + \sigma_{H,\delta}dB_t^\delta + \sigma_{H,\lambda}dB_t^\lambda + \sigma_{H,\mu}dB_t^\mu + \left(e^{-L} - 1\right)dN_t\right). \tag{E.39}$$

Note that

$$\begin{aligned}\pi_t H_t = \pi_0 H_0 &+ \int_0^t \pi_s H_s\left(\mu_H + \mu_\pi - \gamma\sigma_\delta\sigma_{H,\delta} + A_2\sigma_\mu\sigma_{H,\mu} + \sqrt{\lambda_s}A_1\sigma_\lambda\sigma_{H,\lambda}\right)\ ds \\ &+ \int_0^t \pi_s H_s \sigma_{H,\delta} dB_s^\delta + \int_0^t \pi_s H_s \sigma_{H,\mu} dB_s^\mu + \int_0^t \pi_s H_s \sigma_{H,\lambda} dB_s^\lambda + \sum_{0<s_i\leq t}\left(\pi_{s_i}H_{s_i} - \pi_{s_{i^-}}H_{s_{i^-}}\right)\end{aligned} \tag{E.40}$$

is a martingale.[20] This allows me to compute identifying restrictions for μ_H and $\sigma_{H,\delta}$, $\sigma_{H,\mu}$, $\sigma_{H,\lambda}$. First however, the jump terms have to be computed. Note that the diffusion processes related to the wealth-consumption and price-dividend ratios are not affected by the jump and therefore cancel out. Thus,

$$\frac{\pi_t H_t - \pi_{t^-}H_{t^-}}{\pi_t H_t} = \frac{1}{C_{t^-}\tilde{\delta}_{t^-}}\left[\left(C_{t^-}e^{-L\xi}\right)^{-\gamma}\tilde{\delta}_{t^-}e^{-L} - C_{t^-}^{-\gamma}\tilde{\delta}_{t^-}\right] = e^{L(\gamma\xi-1)} - 1. \tag{E.41}$$

[20]Here, s_i denotes the time of the i^{th} jump.

Adding and subtracting a compensation term related to the expected change in (E.40) due to jumps to (E.40) gives

$$\pi_t H_t = \pi_0 H_0 + \int_0^t \pi_s H_s \left(\mu_H + \mu_\pi - \gamma\sigma_\delta\sigma_{H,\delta} + A_2\sigma_\mu\sigma_{H,\mu} + \sqrt{\lambda_s}A_1\sigma_\lambda\sigma_{H,\lambda} + \lambda_s(e^{L(\gamma\xi-1)}-1)\right)\, ds$$
$$+ \int_0^t \pi_s H_s\sigma_{H,\delta}dB_s^\delta + \int_0^t \pi_s H_s\sigma_{H,\mu}dB_s^\mu + \int_0^t \pi_s H_s\sigma_{H,\lambda}dB_s^\lambda$$
$$+ \sum_{0<s_i\le t}\left(\pi_{s_i}H_{s_i} - \pi_{s_{i^-}}H_{s_{i^-}}\right) - \int_0^t \pi_s H_s\lambda_s(e^{L(\gamma\xi-1)}-1)ds. \tag{E.42}$$

Because the πH is a martingale, it must be that

$$0 = \mu_H + \mu_\pi - \gamma\sigma_\delta\sigma_{H,\delta} + A_2\sigma_\mu\sigma_{H,\mu} + \sqrt{\lambda_s}A_1\sigma_\lambda\sigma_{H,\lambda} + \lambda_s(e^{L(\gamma\xi-1)}-1). \tag{E.43}$$

Moreover, by applying Ito's lemma to the guess about the functional form of H, we get that

$$\mu_H = \frac{1}{H}\left(H_D\mu + H_\lambda\kappa(\bar\lambda-\lambda_t) + H_\mu\kappa_\mu\left(\bar\mu-\mu_t\right) - \frac{\partial H}{\partial\tau} + \frac{1}{2}H_{\lambda\lambda}\sigma_\lambda^2\lambda_t + \frac{1}{2}H_{\mu\mu}\sigma_\mu^2\right)$$
$$= \mu_t + B_1(\tau)\kappa(\bar\lambda-\lambda_t) + B_2(\tau)\kappa_\mu(\bar\mu-\mu_t) - B_0'(\tau) - B_1'(\tau)\lambda_t - B_2'(\tau)\mu_t + \frac{1}{2}B_1(\tau)^2\sigma_\lambda^2\lambda_t + \frac{1}{2}B_2(\tau)^2\sigma_\mu^2\mu_t, \tag{E.44}$$

$$\sigma_{H,\delta} = \sigma_\delta, \tag{E.45}$$

$$\sigma_{H,\lambda} = B_1(\tau)\sigma_\lambda\sqrt{\lambda_t}, \tag{E.46}$$

and

$$\sigma_{H,\mu} = B_2(\tau)\sigma_\mu. \tag{E.47}$$

Substituting the expressions gives

$$0 = \mu_t + B_1(\tau)\kappa(\bar\lambda-\lambda_t) + B_2(\tau)\kappa_\mu(\bar\mu-\mu_t) - B_0'(\tau) - B_1'(\tau)\lambda_t - B_2'(\tau)\mu_t + \frac{1}{2}B_1(\tau)^2\sigma_\lambda^2\lambda_t$$
$$+ \frac{1}{2}B_2(\tau)^2\sigma_\mu^2 - \mu_{c,t} - \beta + \gamma\sigma_c^2 - \left(\lambda_t e^{\gamma\xi L}(e^{-\xi L}-1) + \lambda_t\left(e^{\xi\gamma L}-1\right)\right)$$
$$- \gamma\sigma_\delta\sigma_\delta + A_2\sigma_\mu B_2(\tau)\sigma_\mu + \sqrt{\lambda_t}A_1\sigma_\lambda B_1(\tau)\sigma_\lambda\sqrt{\lambda_t} + \lambda_t(e^{L(\gamma\xi-1)}-1). \tag{E.48}$$

Collecting constants and expressions in λ_t and μ_t gives

$$B_0'(\tau) = B_1(\tau)\kappa\bar\lambda + B_2(\tau)\kappa_\mu\bar\mu + \beta + \gamma\sigma_\delta^2 - \gamma\sigma_\delta^2\xi^2 + \frac{1}{2}B_2(\tau)^2 + A_2B_2(\tau)\sigma_\mu^2 \tag{E.49}$$

$$B_1'(\tau) = B_1(\tau)\left(\kappa + A_2\sigma_\lambda^2\right) + \frac{1}{2}B_1(\tau)^2\sigma_\lambda^2 + e^{-L\xi(1/\xi-\gamma)} - e^{-L\xi(1-\gamma)} \tag{E.50}$$

$$B_2'(\tau) = -B_2(\tau)\kappa_\mu - \xi. \tag{E.51}$$

Obviously, $B_0(\tau) = B_1(\tau) = B_2(\tau) = 0$. The solutions to (E.49) to (E.51) are then given by:

$$B_2(\tau) = \frac{\xi}{\kappa_\mu} - \frac{\xi}{\kappa_\mu}\exp(-\kappa_\mu \tau), \tag{E.52}$$

$$B_1(\tau) = -\frac{2\left(e^{-L\xi(1/\xi-\gamma)} - e^{-L\xi(1-\gamma)}\right)(1-e^{-C\tau})}{(C + A_1\sigma_\lambda^2 - \kappa)(1-e^{-C\tau}) - 2C}, \tag{E.53}$$

$$\begin{aligned} B_0(\tau) = {} & \frac{1}{\kappa_\mu}\left(A_2\xi\sigma_\mu^2 + \kappa_\mu\beta + \kappa_\mu\gamma(\sigma_c^2 - \sigma_\delta) - \frac{2\kappa_\mu\left(e^{-L\xi(1/\xi-\gamma)} - e^{-L\xi(1-\gamma)}\right)}{C + A_1\sigma_\lambda^2 - \kappa} + \frac{\xi^2}{2\kappa_\mu}\right)\tau \\ & + \frac{1}{\kappa_\mu}\left(\xi^2\left(\frac{1}{4}e^{2\kappa_\mu\tau} - e^{\kappa_\mu\tau}\right)\right) + \frac{1}{\kappa_\mu^2}A_2\xi e^{\kappa_\mu\tau} + \frac{3}{4\kappa_\mu}\xi^2 - \frac{1}{\kappa_\mu^2}A_2\xi, \end{aligned} \tag{E.54}$$

where $C = \sqrt{(A_1\sigma_\lambda^2 - \kappa)^2 + 2(e^{-L\xi(1-\gamma)} - e^{-L\xi(1/\xi-\gamma)})\sigma_\lambda^2}$. Note that $B_1(\tau) < 0$ for $\tau > 0$, which ensures that the dividend yield is positively related to the outflow risk λ, whereas $B_2(\tau) > 0$ for $\tau > 0$ which ensures that the level of dividend growth is negatively related to the dividend yield. The price of a dividend claim is then given by

$$P_t = \int_t^\infty H_{s,t}\, ds = \tilde{\delta}_t \int_t^\infty \exp\left(B_0(\tau) + B_1(\tau)\lambda_t + B_2(\tau)\mu_t\right)\, ds = \tilde{\delta}_t F(\lambda_t, \mu_t). \tag{E.55}$$

E.5 Returns

Prices P_t then have dynamics given by

$$\frac{dP_t}{P_t} = \mu_{P,t} + \frac{F'_\lambda}{F}\sqrt{\lambda_t}\sigma_\lambda dB_t^\lambda + \frac{F'_\mu}{F}\sigma_\mu dB_t^\mu + \sigma_\delta dB_t^\delta + \left(e^{-L} - 1\right) dN_t \tag{E.56}$$

Applying the same logic as in E.40 to

$$\begin{aligned} \pi_t P_t + \int_0^t \pi_s \delta_s ds = {} & \int_0^t \pi_s P_s \left(\mu_{P,s} + \mu_{\pi,s} + \frac{\tilde{\delta}_s}{P_s} + A_2\frac{F'_\mu}{F}\sigma_\mu^2 + A_1\frac{F'_\lambda}{F}\lambda_t\sigma_\lambda^2 + \lambda_t(e^{L(\gamma\xi-1)} - 1)\right) ds \\ & + \int_0^t \pi_s P_s \frac{F'_\lambda}{F}\sqrt{\lambda_t}\sigma_\lambda dB_s^\lambda + \int_0^t \pi_s P_s \frac{F'_\mu}{F}\sigma_\mu dB_s^\mu + \int_0^t \pi_s P_s \sigma_\delta dB_s^\delta \\ & - \int_0^t \pi_s P_s \gamma\sigma_\delta dB_t^\delta + \int_0^t \pi_s P_s A_2\sigma_\mu dB_s^\mu + \int_0^t \pi_s P_s \sqrt{\lambda_t} A_1\sigma_\lambda dB_s^\lambda \\ & + \sum_{0 < s_i \le t}\left(\pi_{s_i}P_{s_i} - \pi_{s_{i-}}P_{s_{i-}}\right) - \int_0^t \pi_s P_s \lambda_s (e^{L(\gamma\xi-1)} - 1) ds. \end{aligned} \tag{E.57}$$

i.e. that the drift term of the expression on the right hand side must be zero, one gets that the expected return on a dividend claim j is given by:

$$r_t^j = \mu_{P,t} + \frac{\tilde{\delta}_t}{P_t} + \lambda_t\left(e^{-L} - 1\right) \tag{E.58}$$

$$\begin{aligned} & = \mu_{c,t} + \beta - \gamma\sigma_c^2 + \lambda_t e^{\gamma\xi L}\left(e^{-\xi L} - 1\right) + \lambda_t(e^{\gamma\xi L} - 1) - \lambda_t(e^{L(\gamma\xi-1)} - 1) + \lambda_t(e^{-L} - 1) \\ & \quad - A_2\frac{F'_\mu}{F}\sigma_\mu^2 - A_1\frac{F'_\lambda}{F}\sigma_\lambda^2\lambda_t \end{aligned} \tag{E.59}$$

and the expected excess return by

$$r_t^j - r_t = \lambda_t(e^{\gamma\xi L} - 1) - \lambda_t(e^{L(\gamma\xi-1)} - 1) + \lambda_t(e^{-L} - 1) \\ - \underbrace{A_2\frac{F'_\mu}{F}\sigma_\mu^2}_{<0} - \underbrace{A_1\frac{F'_\lambda}{F}\sigma_\lambda^2}_{<0}\lambda_t \tag{E.60}$$

$$= \quad - \underbrace{A_2\frac{F'_\mu}{F}\sigma_\mu^2}_{<0} - \underbrace{A_1\frac{F'_\lambda}{F}\sigma_\lambda^2}_{<0}\lambda_t + \lambda_t\underbrace{\left[(e^{\gamma\xi L} - 1)(1 - e^{-L})\right]}_{>0}. \tag{E.61}$$

F Additional tables

Table F.1: **Predictive regressions, quartile portfolios**

Panel A: Raw returns and dividend growth

		1^{st} quartile	2	3	4^{th} quartile
r_{t+1}	β_r	0.06**	0.07***	0.12**	0.07**
		(2.08)	(4.01)	(2.43)	(2.19)
	R^2	2.55	6.48	12.86	5.14
	β_r^{LR}	0.26***	0.53***	0.85***	0.94***
		(3.21)	(3.2)	(4.58)	(2.73)
Δd_{t+1}	β_d	-0.17	-0.06*	-0.02	-0.01
		(-1.45)	(-1.66)	(-1.15)	(-0.24)
	R^2	4.44	4.68	0.92	0.04
	β_d^{LR}	-0.76***	-0.47***	-0.16	-0.09
		(-2.78)	(-2.39)	(-1.38)	(-0.26)

Panel B: Excess returns and excess dividend growth

		1^{st} quartile	2	3	4^{th} quartile
$r_{t+1} - r_{f,t+1}$	β_r	0.02	0.04**	0.08	0.03
		(0.52)	(2.21)	(1.37)	(0.7)
	R^2	2.55	6.48	12.86	5.14
	β_r^{LR}	0.08	0.28**	0.55***	0.36
		(0.86)	(1.8)	(2.81)	(0.96)
$\Delta d_{t+1} - r_{f,t+1}$	β_d	-0.22*	-0.09**	-0.07**	-0.05
		(-1.8)	(-2.38)	(-2.26)	(-1.49)
	R^2	4.44	4.68	0.92	0.04
	β_d^{LR}	-0.95***	-0.72***	-0.46***	-0.67**
		(-3.31)	(-3.24)	(-3.46)	(-1.75)

Panel C: Dividend yield autoregression

		low IO	2	3	high IO
dp_{t+1}	ρ	0.79***	0.89***	0.88***	0.95***
		(5.46)	(21.19)	(18.77)	(38.45)

Predictive regression and dp-ratio autoregression results. Numbers in brackets are NW-t-statistics with 10 lags. ***, ** and * for one-period slope estimates indicate significance at the ten, five and one percent level, respectively. For long-run coefficients, stars refer to the respective significance levels of one-sided tests computed according to the delta method.

Table F.2: **Simulated $p-$values, quarterly regression**

Panel A: Portfolio 1, low IO

$H_0 : \beta_r = 0, \rho = \hat{\rho}$	β_r	p-value	$Q_{0.05}$	$Q_{0.95}$	β_d	p-value	$Q_{0.05}$	$Q_{0.95}$
short-run	0.01	0.2003	-0.02	0.03	-0.11	0.2057	-0.22	-0.08
long-run	0.12	0.1297	-0.17	0.16	-0.89	0.1507	-1.17	-0.84
p joint hyp.		$\beta_r^{LR} \geq \hat{\beta}_r^{LR} \wedge \beta_d^{LR} \geq \hat{\beta}_d^{LR}$: 0.1297						

$H_0 : \beta_d = 0, \rho = \hat{\rho}$	β_r	p-value	$Q_{0.05}$	$Q_{0.95}$	β_d	p-value	$Q_{0.05}$	$Q_{0.95}$
short-run	0.03**	0.0081	-0.02	0.02	-0.10	0.0862	-0.24	-0.09
long-run	0.25***	0.0007	-0.12	0.13	-0.75***	0.0011	-1.12	-0.87
p joint hyp.		$\beta_r^{LR} \leq \hat{\beta}_r^{LR} \wedge \beta_d^{LR} \leq \hat{\beta}_d^{LR}$: 0.0007						

Panel B: Portfolio 2

$H_0 : \beta_r = 0, \rho = \hat{\rho}$	β_r	p-value	$Q_{0.05}$	$Q_{0.95}$	β_d	p-value	$Q_{0.05}$	$Q_{0.95}$
short-run	0.03***	0.0000	0.12	0.15	-0.10	0.0696	-0.11	0.04
long-run	0.25***	0.0000	0.54	1.43	-0.75***	0.0000	-0.46	0.43
p joint hyp.		$\beta_r^{LR} \geq \hat{\beta}_r^{LR} \wedge \beta_d^{LR} \geq \hat{\beta}_d^{LR}$: 0.0000						

$H_0 : \beta_d = 0, \rho = \hat{\rho}$	β_r	p-value	$Q_{0.05}$	$Q_{0.95}$	β_d	p-value	$Q_{0.05}$	$Q_{0.95}$
short-run	0.03***	0.0000	0.12	0.15	-0.10**	0.0110	-0.07	0.03
long-run	0.25***	0.0000	0.67	1.32	-0.75***	0.0000	-0.33	0.32
p joint hyp.		$\beta_r^{LR} \leq \hat{\beta}_r^{LR} \wedge \beta_d^{LR} \leq \hat{\beta}_d^{LR}$: 0.0000						

Table continues on next page

Table continues from previous page

Panel C: Portfolio 3

$H_0 : \beta_r = 0, \rho = \hat{\rho}$	β_r	p-value	$Q_{0.05}$	$Q_{0.95}$	β_d	p-value	$Q_{0.05}$	$Q_{0.95}$
short-run	0.04*	0.0490	-0.02	0.04	-0.04^{**}	0.0124	-0.16	-0.05
long-run	0.48***	0.0043	-0.29	0.31	-0.52^{***}	0.0043	-1.29	-0.69
p joint hyp.		$\beta_r^{LR} \geq \hat{\beta}_r^{LR} \wedge \beta_d^{LR} \geq \hat{\beta}_d^{LR}$: 0.0043						

$H_0 : \beta_d = 0, \rho = \hat{\rho}$	β_r	p-value	$Q_{0.05}$	$Q_{0.95}$	β_d	p-value	$Q_{0.05}$	$Q_{0.95}$
short-run	0.04***	0.0043	0.05	0.11	-0.04	0.2487	-0.08	0.02
long-run	0.48	0.0576	0.47	1.45	-0.52	0.0576	-0.53	0.45
p joint hyp.		$\beta_r^{LR} \leq \hat{\beta}_r^{LR} \wedge \beta_d^{LR} \leq \hat{\beta}_d^{LR}$: 0.0576						

Panel D: Portfolio 4, high IO

$H_0 : \beta_r = 0, \rho = \hat{\rho}$	β_r	p-value	$Q_{0.05}$	$Q_{0.95}$	β_d	p-value	$Q_{0.05}$	$Q_{0.95}$
short-run	0.05	0.0563	-0.03	0.05	-0.03^{***}	0.0043	-0.15	-0.05
long-run	0.62***	0.0027	-0.37	0.39	-0.38^{***}	0.0029	-1.37	-0.61
p joint hyp.		$\beta_r^{LR} \geq \hat{\beta}_r^{LR} \wedge \beta_d^{LR} \geq \hat{\beta}_d^{LR}$: 0.0027						

$H_0 : \beta_d = 0, \rho = \hat{\rho}$	β_r	p-value	$Q_{0.05}$	$Q_{0.95}$	β_d	p-value	$Q_{0.05}$	$Q_{0.95}$
short-run	0.05*	0.0430	0.05	0.12	-0.03	0.3120	-0.08	0.03
long-run	0.62	0.1689	0.47	1.48	-0.38	0.1659	-0.53	0.48
p joint hyp.		$\beta_r^{LR} \leq \hat{\beta}_r^{LR} \wedge \beta_d^{LR} \leq \hat{\beta}_d^{LR}$: 0.1659						

Parameter estimates and simulated p-values for the estimates. Q_p denotes the p-quantile of the simulated distribution. *,** and *** denote significance on the one, five and ten percent level, respectively (based on 5,000 simulations and with respect to the respective null hypothesis, i.e. whether the estimate is consistent with the null. **Stars indicate that it is not**). The rows labeled "$\widehat{\Pr}[\,\cdot\,|H_0]$" show the estimated probability of the noted events occuring given that the respective null hypothesis is true.

Table F.3: **Simulated $p-$values, quarterly regression, excess returns**

Panel A: Portfolio 1, low IO

$H_0: \beta_r = 0, \rho = \hat{\rho}$	β_r	p-value	$Q_{0.05}$	$Q_{0.95}$	β_d	p-value	$Q_{0.05}$	$Q_{0.95}$
short-run	0.01	0.2420	-0.02	0.03	-0.11	0.2242	-0.22	-0.08
long-run	0.10	0.1747	-0.17	0.16	-0.91	0.2055	-1.17	-0.84
p joint hyp.		$\beta_r^{LR} \geq \hat{\beta}_r^{LR} \wedge \beta_d^{LR} \geq \hat{\beta}_d^{LR}$: 0.1747						
$H_0: \beta_d = 0, \rho = \hat{\rho}$	β_r	p-value	$Q_{0.05}$	$Q_{0.95}$	β_d	p-value	$Q_{0.05}$	$Q_{0.95}$
short-run	0.01^{***}	0.0000	0.10	0.15	-0.11^{*}	0.0434	-0.11	0.04
long-run	0.10^{***}	0.0000	0.53	1.44	-0.91^{***}	0.0000	-0.47	0.44
p joint hyp.		$\beta_r^{LR} \leq \hat{\beta}_r^{LR} \wedge \beta_d^{LR} \leq \hat{\beta}_d^{LR}$: 0.0000						

Panel B: Portfolio 2

$H_0: \beta_r = 0, \rho = \hat{\rho}$	β_r	p-value	$Q_{0.05}$	$Q_{0.95}$	β_d	p-value	$Q_{0.05}$	$Q_{0.95}$
short-run	0.03^{**}	0.0126	-0.02	0.02	-0.10	0.0983	-0.24	-0.09
long-run	0.23^{***}	0.0020	-0.12	0.13	-0.77^{***}	0.0024	-1.12	-0.87
p joint hyp.		$\beta_r^{LR} \geq \hat{\beta}_r^{LR} \wedge \beta_d^{LR} \geq \hat{\beta}_d^{LR}$: 0.0002						
$H_0: \beta_d = 0, \rho = \hat{\rho}$	β_r	p-value	$Q_{0.05}$	$Q_{0.95}$	β_d	p-value	$Q_{0.05}$	$Q_{0.95}$
short-run	0.03^{***}	0.0000	0.12	0.15	-0.10	0.0650	-0.11	0.04
long-run	0.23^{***}	0.0000	0.54	1.43	-0.77^{***}	0.0000	-0.46	0.43
p joint hyp.		$\beta_r^{LR} \leq \hat{\beta}_r^{LR} \wedge \beta_d^{LR} \leq \hat{\beta}_d^{LR}$: 0.0000						

Table continues on next page

Table continues from previous page

Panel C: Portfolio 3

$H_0 : \beta_r = 0, \rho = \hat{\rho}$	β_r	p-value	$Q_{0.05}$	$Q_{0.95}$	β_d	p-value	$Q_{0.05}$	$Q_{0.95}$
short-run	0.03	0.0620	-0.02	0.04	-0.04^{**}	0.0199	-0.16	-0.05
long-run	0.45^{**}	0.0063	-0.29	0.31	-0.55^{**}	0.0066	-1.29	-0.69
p joint hyp.		$\beta_r^{LR} \geq \hat{\beta}_r^{LR} \wedge \beta_d^{LR} \geq \hat{\beta}_d^{LR}$: 0.0063						

$H_0 : \beta_d = 0, \rho = \hat{\rho}$	β_r	p-value	$Q_{0.05}$	$Q_{0.95}$	β_d	p-value	$Q_{0.05}$	$Q_{0.95}$
short-run	0.03^{***}	0.0031	0.05	0.11	-0.04	0.2280	-0.08	0.02
long-run	0.45^{*}	0.0417	0.47	1.46	-0.55^{*}	0.0416	-0.53	0.46
p joint hyp.		$\beta_r^{LR} \leq \hat{\beta}_r^{LR} \wedge \beta_d^{LR} \leq \hat{\beta}_d^{LR}$: 0.0416						

Panel D: Portfolio 4, high IO

$H_0 : \beta_r = 0, \rho = \hat{\rho}$	β_r	p-value	$Q_{0.05}$	$Q_{0.95}$	β_d	p-value	$Q_{0.05}$	$Q_{0.95}$
short-run	0.04	0.0690	-0.03	0.05	-0.03^{**}	0.0070	-0.15	-0.05
long-run	0.58^{**}	0.0051	-0.38	0.39	-0.42^{**}	0.0053	-1.38	-0.61
p joint hyp.		$\beta_r^{LR} \geq \hat{\beta}_r^{LR} \wedge \beta_d^{LR} \geq \hat{\beta}_d^{LR}$: 0.0051						

$H_0 : \beta_d = 0, \rho = \hat{\rho}$	β_r	p-value	$Q_{0.05}$	$Q_{0.95}$	β_d	p-value	$Q_{0.05}$	$Q_{0.95}$
short-run	0.04^{*}	0.0296	0.05	0.12	-0.03	0.2788	-0.08	0.03
long-run	0.58	0.1259	0.48	1.48	-0.42	0.1227	-0.52	0.48
p joint hyp.		$\beta_r^{LR} \leq \hat{\beta}_r^{LR} \wedge \beta_d^{LR} \leq \hat{\beta}_d^{LR}$: 0.1227						

Parameter estimates and simulated p-values for the estimates. Q_p denotes the p-quantile of the simulated distribution. *,** and *** denote significance on the one, five and ten percent level, respectively (based on 5,000 simulations and with respect to the respective null hypothesis, i.e. whether the estimate is consistent with the null. **Stars indicate that it is not**). The rows labeled "$\widehat{\Pr}[\cdot|H_0]$" show the estimated probability of the noted events occuring given that the respective null hypothesis is true.

Table F.4: **Sorting variables descriptives**

	low BM	high BM	low ME	high ME
	BM		ME in m $	
IO $< 30\%$	0.23			3332
IO $\geq 30\%$	0.25			5337
IO $< 5\%$		2.25	17	
IO $\geq 5\%$		1.36	25	
	Average number of stocks			
IO $< 30\%$	407			276
IO $\geq 30\%$	659			1275
IO $< 5\%$		156	353	
IO $\geq 5\%$		909	670	

Descriptive statistics on the sorting variables of the double sorts. BM is the book-to-market equity ratio. ME is market equity. low (high), BM (ME) are the bottom (top) 30% of stocks in the sample. All values are equal weighted averages.

Table F.5: **Predictive regressions of earnings growth and using the price-earnings ratio**

	Panel A:		Panel B:		Panel C:	
	$\Delta E_{t+1} = \alpha + \beta \cdot DP_t$		$\Delta E_{t+1} = \alpha + \beta \cdot EP_t$		$R_{t+1} = \alpha + \beta \cdot EP_t$	
	β	R^2	β	R^2	β	R^2
low IO	-2.7994		-6.627		-0.4754	
	p. 0.4936	0.00	*p.* 0.1548	0.04	*p.* 0.8244	0.02
2	-8.4709		-0.0973		-0.1797	
	p. 0.3933	0.03	*p.* 0.4735	0.00	*p.* 0.8189	0.02
3	-4.9606		-0.5448		0.8265	
	p. 0.2866	0.01	*p.* 0.4161	0.00	*p.* 0.1073	0.05
high IO	6.6171		2.2675		2.2625	
	p. 0.7194	0.01	*p.* 0.7071	0.00	*p.* 0.114	0.08

Slope estimates of predictive regressions of earnings growth $\Delta E_{t+1} = \frac{E_{t+1}}{E_t}$ and returns R on the earnings-price ratio EP and on the dividend yield DP for each of the four portfolios. Simulated $p-$values $p.$ are computed using 5000 articfical data sets generated under the null of no predictability:

$$Y_{t+1} = \bar{Y} + \epsilon^Y_{t+1}$$
$$X_{t+1} = \rho X_t + \epsilon^X_{t+1},$$

where $\bar{Y}$, ρ and $Cov(\epsilon^Y, \epsilon^r)$ are as estimated from the data. $E[\epsilon^Y] = E[\epsilon^x] = 0$.