Understanding Predictability

Lior Menzly

University of Southern California

Tano Santos

Columbia University and NBER

Pietro Veronesi

University of Chicago, CEPR and NBER

Abstract

We propose a general equilibrium model with multiple securities in which investors' risk preferences and expectations of dividend growth are time varying. While time varying risk preferences induce the standard positive relation between the dividend yield and expected returns, time varying expected dividend growth induces a *negative* relation between them. These offsetting effects reduce the ability of the dividend yield to forecast returns and eliminate its ability to forecast dividend growth, as observed in the data. The model links the predictability of returns to that of dividend growth, suggesting specific changes to standard linear predictive regressions for both. The model's predictions are confirmed empirically.

Acknowledgements

Different versions of the framework advanced in this paper have been presented at workshops at the University of Texas at Austin, NYU, University of Illinois at Urbana-Champaign, McGill, MIT, Columbia University, Northwestern University, University of California at Berkely, Dartmouth College, University of Iowa, and the University of Chicago. We thank Andrew Ang, John Cochrane, George Constantinides, Kent Daniel, Harry DeAngelo, Lars Hansen, John Heaton, Bob Hodrick, Martin Lettau, Toby Moskowitz, Ruy Ribeiro, and Jessica Wachter for their comments.

I. INTRODUCTION

The predictability of stock market returns at long horizons has been at the center of much empirical and theoretical research in finance in the last decade. For the market as a whole, high prices, normalized by dividends for example, forecast low future returns, whereas they do not predict high future dividend growth. This evidence has deep implications for both finance and macroeconomics. In particular, it has been interpreted as additional evidence of a time varying price of risk. Still, the statistical evidence is weak, inconsistent across time periods, and different across different assets. This has led many to challenge the ability of valuation ratios to predict future returns. The behavior of the aggregate market during the 1990s, which saw high price dividend ratios and high returns, has only intensified a debate that remains at the center of current research in asset pricing.[1]

We show that a general equilibrium model in which both investors' preferences for risk and their expectations of future dividend growth are time varying goes a long way in explaining these empirical findings. In particular, while time varying risk preferences induce the standard positive relation between dividend yields and expected returns, time varying expected dividend growth induces a *negative* relation between them in equilibrium. These offsetting effects reduce the ability of the dividend yield to forecast future returns and, essentially, eliminate its ability to forecast future dividend growth. In our model, the extent to which these effects cancel each other depends on the properties of the asset's cash flow process, thereby yielding different predictions across different portfolios. We also show that traditional linear predictive regressions for both returns and

[1] For predictability, see, for instance, Campbell and Shiller (1988), Fama and French (1988), Hodrick (1992), Lamont (1998), and Lettau and Ludvigson (2001). For time varying market price of aggregate risk, see e.g. the discussion in Campbell et al. (1997, page 496), or Campbell and Cochrane (1999, page 206). For an early discussion on the performance of predictability regression during the 1990s see Cochrane (1997, page 10.). For evidence on individual stocks, see Vuolteenaho (2002). On some of the relevant econometric issues see Richardson and Stock (1989) and Stambaugh (1999).

dividend growth should be adjusted in specific ways to fully capture the informational content of prices. The modified forecasting regressions indeed show that both future dividend growth and future returns are strongly predictable, lending strong empirical support to the economic mechanism we propose.

Our model combines two ingredients. First, we introduce a new model to describe the cash flow processes of individual assets that is parsimonious, tractable within a general equilibrium framework, and, most importantly, fits well the properties of cash flow data.[2] Rather than modeling the dividend processes of individual assets, we specify processes for the fraction (share) that each asset contributes to total consumption. The shares are bounded between zero and one and, in addition, they sum up to a value that is always less than or equal to one in every period. We also assume that no asset will dominate the economy, even in the long run. For this reason, we assume that these shares mean revert to a long run value which is strictly less than one. Finally, our cash flow model allows for sources of consumption other than financial assets, like labor income and government transfers, breaking the standard, but unappealing, market clearing condition that equates dividends with consumption. As a second ingredient we use an external habit persistence model, similar to the one of Campbell and Cochrane (1999), which produces time varying preferences for risk and thus a time varying aggregate premium that agents require to hold risky assets.

We first solve for the price dividend ratio of an individual security and show that it is a linear function of a variable that proxies for expected dividend growth, a variable that proxies for investors' aggregate risk tolerance, and, importantly, an interaction term. For a given expected dividend growth, a decrease in risk tolerance increases the equity premia on all assets and decreases their price dividend ratios. That is, the variation in

[2]Our model is closely related to the general equilibrium models of Bossaerts and Green (1990), who also model dividend processes directly, and Brock (1982), though his is in a production economy framework. See also Abel (1999), Bansal, Dittmar, and Lundblad (2001), Bakshi and Chen (1997), Bekaert and Grenadier (2001), and Ang and Liu (2003).

risk preferences induces the standard negative relation between price dividend ratios and expected returns. However, for a given investors' attitude towards risk, an increase in the asset's expected dividend growth yields both an increase in its price dividend ratio and an *increase* in its equity premium. That is, changes in expected dividend growth induce a *positive* relation between price dividend ratios and expected excess returns, contrary to the common wisdom on return predictability. The reason is that an increase in expected dividend growth implies that the asset pays farther ahead in the future, making its price more sensitive to shocks to the aggregate discount rate, that is, to fluctuations in the investors' risk preferences. Since this additional volatility of the asset is perfectly correlated with changes in investors' attitude towards risk, it must be priced, and thus the larger premium. In equilibrium, however, this increase in the premium is not sufficient to offset the increase in the price dividend ratio that stems from a higher expected dividend growth, so the positive relation between price dividend ratios and expected future returns remains, though attenuated.

Notice that it is the combination of time varying expected dividend growth and time varying aggregate risk preferences that generates this effect: were either of the two constant, the effect would vanish. Importantly then, the dividend yield may be at times a poor predictor for future returns, as it may vary opposite to expected excess returns.

Our framework suggests a simple correction to the standard predictability regression to disentangle the conflicting effects that shocks to expected dividend growth and to the preferences for risk have on the expected excess returns of individual assets. We show that the expected excess rate of return of an asset is a linear function not only of its dividend yield, but also of its consumption to price ratio, a result for which we find empirical support for the set of industry portfolios that we use in our empirical tests. In addition, simulations show that the magnitude of the effects generated by our model matches the one in the data.

The predictability of stock returns is only one side of the coin. Fluctuations in price dividend ratios must either predict changes in expected returns, changes in the dividend growth, or both. In the data, aggregate dividend growth though is not predicted by past price dividend ratios of the market portfolio. This is to be expected. We show that an asset's expected dividend growth is indeed linear in its price dividend ratio, but the slope coefficient is itself a function of the variable driving changes in the aggregate preferences for risk. Thus, linear regressions that fail to correct for these changes do not reveal the information on future dividend growth contained in price dividend ratios. As before, the model suggests a straightforward adjustment to the dividend growth predictability regressions: divide the price dividend ratio by the corresponding price consumption ratio, which enters as a control for changes in risk preferences, to obtain the share of dividends to consumption as a forecasting variable for dividend growth. Our empirical tests find considerable support for this simple adjustment, both for the aggregate market and for the set of industry portfolios, and confirm that dividend growth is indeed forecastable – although not by the price dividend ratio. Importantly, these results obtain in simulated data as well.

Finally, we link the predictability of dividend growth to the predictability of returns. We show that those assets characterized by a slow mean reversion of expected dividend growth should have returns that are better predicted by the dividend yield. Instead if the mean reversion is fast, the consumption to price ratio should be a better predictor of returns. We provide empirical evidence to corroborate this prediction.

In addition to the voluminous literature on predictability, surveyed, for instance, in Campbell et al. (1997) and Cochrane (2000), this paper relates to recent empirical work on the relation between valuation ratios, the predictability of returns, and dividend growth. First, our model provides theoretical support for the empirical findings of Vuolteenaho (2002), who shows that for individual assets, changes in expected returns and changes in expected cash flow growth are positively related, specially for small stocks. Second, Lettau and Ludvigson (2003) find that aggregate dividends are forecastable, but not by the price dividend ratio and argue for the existence of common components in expected returns and dividend growth, that they assume exogenously.

Our model shows that this common component should be expected in equilibrium, as positive changes in expected dividend growth naturally increase the asset's riskiness, and thus its required premium.

We introduce the model in Section II. The theoretical and empirical results are contained in Sections III and IV respectively. Section V concludes. Proofs are contained in the Appendix.

II. THE MODEL

II.A Preferences

The economy is composed of a representative consumer who maximizes:

$$E\left[\int_0^\infty u(C_t, X_t, t)\, dt\right] = E\left[\int_0^\infty e^{-\rho t} \log(C_t - X_t)\, dt\right], \tag{1}$$

where X_t denotes the habit level and ρ denotes the subjective discount rate. We assume throughout that habit is *external* to the individual, that is, an individual's habit level is determined by aggregate consumption rather than by his own.[3]

The effect of habit persistence on the agent's attitudes towards risk can be conveniently summarized by the *surplus consumption ratio*, S_t, defined as:

$$S_t = \frac{C_t - X_t}{C_t}. \tag{2}$$

Movements of this surplus produce fluctuations of the local curvature of the utility function, $1/S_t$, and hence they translate naturally into the corresponding variation on the prices and returns of financial assets.[4] Thus, the particular assumptions on the dynamics

[3]On habit persistence and asset pricing see Sundaresan (1989), Constantinides (1990), Abel (1990), Ferson and Constantinides (1991), Detemple and Zapatero (1991), Daniel and Marshall (1997), Chapman (1998), Li (2001), and Wachter (2001). These papers focus on different questions than the present one though.

[4]As in Campbell and Cochrane (1999), we term $-\frac{u_{CC}}{u_C} C_t$ the local curvature of the utility function. However, in our setting the same calculations as in the appendix of Campbell and Cochrane (1999, NBER version) imply a value of the Arrow-Pratt coefficient of relative risk aversion equal to $RRA_t = \overline{Y} + \frac{\rho}{\rho+k}\left(Y_t - \overline{Y}\right)$, where $Y_t = 1/S_t$ as in equation (3). With a slight abuse of terminology then, we refer to Y_t as the degree of risk aversion.

of S_t are of critical importance. Campbell and Cochrane (1999) assume that $\log(S_t)$ follows a mean reverting process with shocks that are conditionally perfectly correlated with innovations in consumption growth. For tractability, we find it convenient to instead impose the stochastic structure on the *inverse* of the surplus consumption ratio, which we denote Y_t,

$$Y_t = \frac{1}{S_t} = \frac{C_t}{C_t - X_t} = \frac{1}{1 - \left(\frac{X_t}{C_t}\right)} > 1. \tag{3}$$

Throughout we refer to Y_t as the *inverse surplus*. Analogously to Campbell and Cochrane (1999), we assume that it follows a mean reverting process, perfectly negatively correlated with innovations in consumption growth

$$dY_t = k\left(\overline{Y} - Y_t\right) dt - \alpha\left(Y_t - \lambda\right)\left(dc_t - E_t\left[dc_t\right]\right), \tag{4}$$

where $\overline{Y}$ is the long run mean of the inverse surplus and k is the speed of the mean reversion. Here $c_t = \log(C_t)$ and we assume that it follows the process:

$$dc_t = \mu_c dt + \sigma_c dB_t^1, \tag{5}$$

where μ_c is the mean consumption growth, $\sigma_c > 0$ is a scalar, and B_t^1 is a Brownian motion.

The parameter $\alpha > 0$ in (4) captures the impact of unexpected consumption growth on the inverse surplus process. A negative innovation to consumption growth, for example, results in an increase in the inverse surplus, or, equivalently, a decrease in the surplus level, capturing the intuition that the consumption level C_t moves further away from a slow moving habit X_t. The parameter $\lambda \geq 1$ ensures a lower bound for the inverse surplus, and an upper bound for the surplus itself. For instance, if the surplus S_t is to live in $[0, .1]$ (as in the calibration of Campbell and Cochrane (1999)) then $Y_t \in [10, \infty)$, and this can be guaranteed by setting $\lambda = 10$. Clearly, we assume that $\overline{Y} > \lambda$.

Finally, in our model, as in Campbell and Cochrane (1999), habit is implicitly defined as $X_t = C_t(1 - S_t)$ and as, a consequence, we must ensure that its joint properties with aggregate consumption conform to economic intuition. Specifically, we impose the restriction that $cov_t(dC_t, dX_t) > 0$ for all S_t to guarantee that positive shocks to aggregate consumption result in a utility loss for the individual agent. For this to be the case it is enough to bound the parameter α by $\overline{\alpha}(\lambda)$, that is,

$$\alpha \leq \overline{\alpha}(\lambda) = (2\lambda - 1) + 2\sqrt{\lambda(\lambda - 1)}, \tag{6}$$

which we assume holds throughout.[5]

II.B The cash flow model

There are n risky financial assets paying a dividend rate, $\{D_t^i\}_{i=1}^n$, in units of a homogeneous and perishable consumption good. We assume that agents total income is made up of these n cash flows, plus other proceeds such as labor income and government transfers. Let D_t^0 be the aggregate income flow that is not financial in nature. Thus, agents' total income is given by $\sum_{i=0}^n D_t^i$. As the consumption good is immediately perishable, in equilibrium total income equals total consumption, and hence, $C_t = \sum_{i=0}^n D_t^i$.

Rather than modeling the process for the dividend rates themselves, we assume that the *share* of consumption that each asset produces,

$$s_t^i = \frac{D_t^i}{C_t} \quad \text{for} \quad i = 1, ..., n \tag{7}$$

evolves according to a mean reverting process of the form[6]

$$ds_t^i = \phi^i\left(\overline{s}^i - s_t^i\right)dt + s_t^i \boldsymbol{\sigma}^i(\mathbf{s}_t)\,d\mathbf{B}_t', \tag{8}$$

where $\mathbf{B}_t = \left(B_t^1, ..., B_t^N\right)$ is a N dimensional row vector of standard Brownian motions, $\overline{s}^i \in [0, 1)$ is asset i's average long-term consumption share, ϕ^i is the speed of mean

[5] It is enough to guarantee that $cov_t(dC_t, dX_t) = C_t^2\left[(1 - S_t) - \alpha S_t(1 - \lambda S_t)\right]\sigma_c^2 > 0$ for $S^{\min} = (1 + \alpha)/2\lambda\alpha$. We thank George Constantinides for pointing this to us.

[6] The process for the alternative source of income, s_t^0, follows immediately from the fact that $s_t^0 = 1 - \sum_{i=1}^n s_t^i$. For the model's implications for the relation of labor income with stock returns see Santos and Veronesi (2001).

reversion, and

$$\boldsymbol{\sigma}^{i}\left(\mathbf{s}_{t}\right)=\mathbf{v}^{i}-\sum_{j=0}^{n} s_{t}^{j} \mathbf{v}^{j}=\left[\sigma_{1}^{i}\left(\mathbf{s}_{t}\right), \sigma_{2}^{i}\left(\mathbf{s}_{t}\right), \cdots, \sigma_{N}^{i}\left(\mathbf{s}_{t}\right)\right] \tag{9}$$

is a N dimensional row vector of volatilities, with $\mathbf{v}^{i}=\left[v_{1}^{i}, v_{2}^{i}, \ldots, v_{N}^{i}\right]$ for $i=0,1,\cdots,n$ a row vector of constants with $N \leq n+1$. $\boldsymbol{\sigma}^{i}\left(\mathbf{s}_{t}\right)$ is parametrically indeterminate, that is, adding a constant vector to all the $\mathbf{v}^{i}$'s leaves the share processes unaltered. For analytical convenience, we then renormalize the constants $\mathbf{v}^{i}$'s, for $i=0,1,...,n$, so that

$$\sum_{j=0}^{n} \overline{s}^{j} \mathbf{v}^{j}=\mathbf{0}. \tag{10}$$

The share process described in (8) has a number of reasonable properties. The functional form of the volatility term (9) arises for *any* homoskedastic dividend growth model. That is, denoting by $\delta_{t}^{i}=\log\left(D_{t}^{i}\right)$, (9) results from any model of the form, $d\delta_{t}^{i}=\mu^{i}\left(\mathbf{D}_{t}\right) dt+\mathbf{v}^{i} d\mathbf{B}_{t}^{\prime}$, as it is immediate to verify by applying Ito's Lemma to $s_{t}^{i}=D_{t}^{i}/(\sum_{j=0}^{n} D_{t}^{j})$. Tighter assumptions, in contrast, are imposed on the drift of (8), and this is the essence of our cash flow model. It is an economically sensible assumption that no asset should dominate the whole economy and for this reason we impose that the process is mean reverting. In addition, we show in the Appendix that, in order to guarantee that dividends are positive, i.e. $s_{t}^{i} \geq 0$, and that total income equals total consumption, that is, $\sum_{i=0}^{n} s_{t}^{i}=1$, it is enough to assume that $\sum_{i=1}^{n} \overline{s}^{i}<1$ and $\phi^{i}>\sum_{j=1}^{n} \overline{s}^{j} \phi^{j}$, which we do throughout.

Finally, the covariance between share and consumption growth is given by,

$$cov_{t}\left(\frac{ds_{t}^{i}}{s_{t}^{i}}, \frac{dC_{t}}{C_{t}}\right)=\theta_{CF}^{i}-\sum_{j=0}^{n} \theta_{CF}^{j} s_{t}^{j} \qquad \text{where} \qquad \theta_{CF}^{i}=v_{1}^{i} \sigma_{c}. \tag{11}$$

The sign of this covariance, which given (10) is largely determined by the sign of θ_{CF}^{i}, characterizes asset i as a good or a bad hedge against adverse consumption shocks. If, say, the covariance is negative, then the asset is a larger fraction of consumption when consumption shrinks and hence it serves as a partial hedge against bad times.

Implications of the share model for cash flows

A straightforward application of Ito's Lemma shows that $\delta_t^i = \log\left(D_t^i\right)$ follows the process[7]

$$
\begin{aligned}
d\delta_t^i &= \mu_D^i\left(\mathbf{s}_t\right)dt + \boldsymbol{\sigma}_D^i\left(\mathbf{s}_t\right)d\mathbf{B}_t' \qquad \text{where} \\
\mu_D^i\left(\mathbf{s}_t\right) &= \mu_c + \phi^i\left(\frac{\overline{s}^i}{s_t^i} - 1\right) - \frac{1}{2}\boldsymbol{\sigma}^i\left(\mathbf{s}_t\right)\boldsymbol{\sigma}^i\left(\mathbf{s}_t\right)', \qquad (12) \\
\boldsymbol{\sigma}_D^i\left(\mathbf{s}_t\right) &= \boldsymbol{\sigma}_c + \boldsymbol{\sigma}^i\left(\mathbf{s}_t\right) \qquad \text{where} \qquad \boldsymbol{\sigma}_c = \left(\sigma_c, 0, ..., 0\right). \qquad (13)
\end{aligned}
$$

The drift of the dividend growth process, (12), is given by that of consumption plus two additional terms. The first term is governed by fluctuations in $\overline{s}^i/s_t^i$ around its long run value of 1. If, for instance, $\overline{s}^i/s_t^i > 1$, asset i experiences high dividend growth in order to "catch up" to its log run share of consumption, $\overline{s}^i$. We call $\overline{s}^i/s_t^i$ the *relative share* and it is our proxy for *expected* dividend growth. The second term in (12) is the usual Jensen's inequality term. The asset then can experience dividend growth rates that are, say, well above those of consumption for long periods of time but eventually it must grow at a rate that is consistent with that of consumption, otherwise the asset will eventually dominate the economy. Modeling this long run connection between consumption and dividends is important particularly if one is interested in the long run predictability of returns or more generally, low frequency moments of the return distribution.

The diffusion component, (13), depends on the vector of share processes, $\mathbf{s}_t$. This heteroskedasticity is a consequence of both the restriction that the shares of the different financial assets add up to a process bounded between zero and one and the assumption that log-consumption is given by (5). We find the magnitude of this heteroskedasticity, however, is negligible compared to the volatility level.

Finally, the covariance between dividend and consumption growth is given by,

$$
cov_t\left(d\delta_t^i, dc_t\right) = \sigma_c^2 + \theta_{CF}^i - \sum_{j=0}^{n} s_t^j \theta_{CF}^j, \qquad (14)
$$

[7]Even if for some asset i we have $\overline{s}^i = 0$, the share itself is strictly positive if it was positive at time $t = 0$. Thus, there are no difficulties in taking the logarithm of s_t^i.

where θ^i_{CF} was defined in (11). The normalization in (10) implies that the unconditional expected covariance between dividend growth and consumption is given by $E\left[cov_t\left(d\delta^i_t, dc_t\right)\right] = \sigma^2_c + \theta^i_{CF}$, as $\sum_{j=0}^n \overline{s}^j \theta^i_{CF} = 0$.

In summary then, our cash flow model rests on two assumptions. First, log dividends and log consumption are cointegrated and their levels are in a fixed long run relation given by $\overline{s}^i$. Second, as shown in (12), the relative share $\overline{s}^i/s^i_t$ should predict future dividend growth. We show in Section IV that there is strong empirical support for both assumptions.

III. PRICES, FUTURE DIVIDEND GROWTH, AND EXPECTED RETURNS

In this section we investigate the way prices reflect information about changes in expected returns, in dividend growth, or in both. We assume the existence of a representative investor. Then if asset g pays $D^g_t = s^g_\tau C_t$ at time τ, the price at time t is:

$$P^g_t = E_t\left[\int_t^\infty e^{-\rho(\tau-t)}\left(\frac{u_c\left(C_\tau - X_\tau\right)}{u_c\left(C_t - X_t\right)}\right) D^g_\tau d\tau\right] = \frac{C_t}{Y_t} E_t\left[\int_t^\infty e^{-\rho(\tau-t)} s^g_\tau Y_\tau d\tau\right]. \quad (15)$$

Closed form solutions obtain whenever the expectation in (15) can be solved, which, under suitable conditions,[8] is identical to solving

$$E_t\left[e^{-\rho(\tau-t)} s^g_\tau Y_\tau\right]. \quad (16)$$

Excess returns are then given by $dR^g_t = (dP^g_t + s^g_t C_t dt)/P^g_t - r_t dt$.

III.A The total wealth portfolio

For the total wealth portfolio $D^g_\tau = C_\tau$, and hence $s^g_\tau = 1$. Then we can show

Proposition 1. (a) The price consumption ratio of the total wealth portfolio is given by:

$$\frac{P^{TW}_t}{C_t} = \frac{1}{\rho}\left(\frac{\rho + k\overline{Y}S_t}{\rho + k}\right). \quad (17)$$

[8] We assume that the process for Y_t is such that $\mathbf{E}_t\left[\int_t^\infty e^{-\rho(\tau-t)} Y_\tau d\tau\right] < \infty$. Thus we can invoke Fubini's Theorem to justify the inversion of integration and expectation implicit in (16).

(b) The process for excess returns on the total wealth portfolio is given by $dR_t^{TW} = \mu_R^{TW}(S_t)\,dt + \sigma_R^{TW}(S_t)\,dB_{1,t}$, where

$$\mu_R^{TW}(S_t) = (1 + \alpha(1 - \lambda S_t))\,\sigma_R^{TW}(S_t)\,\sigma_c \tag{18}$$

$$\sigma_R^{TW}(S_t) = \left[1 + \frac{k\overline{Y}S_t(1 - \lambda S_t)\,\alpha}{k\overline{Y}S_t + \rho}\right]\sigma_c. \tag{19}$$

Expression (17) neatly captures the mechanism embedded in habit persistence models. A positive innovation in aggregate consumption increases S_t, which decreases the local curvature of the utility function, making the investor less "risk averse," and hence the increase in P_t^{TW}/C_t.

The expected excess return and volatility of returns of the total wealth portfolio, (18) and (19), are plotted in figure $1c$. As intuition suggests, for high values of S_t, both $\mu_R^{TW}(S_t)$ and $\sigma_R^{TW}(S_t)$ are decreasing in S_t. However, they are increasing in S_t for very low values of S_t. The reason is that as $S_t \to 0$, its volatility must necessarily vanish in order to prevent S_t from becoming negative, otherwise the marginal utility of consumption could become negative. This results in a lower volatility of returns, and, thus, in a decrease in the expected excess return. To gauge the quantitative importance of the backward bending side of both $\mu_R^{TW}(S_t)$ and $\sigma_R^{TW}(S_t)$, figure $1a$ shows the stationary density of S_t. Notice that this density has a relatively thin left hand side tail and thus this effect does not seem quantitatively important. Finally, the Sharpe ratio, $(1 + \alpha(1 - \lambda S_t))\,\sigma_c$, increases monotonically as S_t drops (figure $1d$).[9]

III.B Individual securities

In this case $s_\tau^g = s_\tau^i$, the share of asset i, and now the expectation in (16) depends on the term $s_\tau^i Y_\tau$. To highlight the main novel implications of our model for predictability we study first the simpler case where all assets have equal cash flow risk, that is, identical covariance between dividend growth and consumption growth. In this case we obtain

[9]This follows from (18) and (19). In the case of log-utility with no habit, that is, $\alpha = 0$ and $Y_t = \overline{Y} = \lambda = 1$, then $S_t = 1/\overline{Y}$ for $t \geq 0$ and $P_t^{TW}/C_t = 1/\rho$, $\mu_R^{TW} = \sigma_c^2$ and $\sigma_R^{TW} = \sigma_c$ (Rubinstein (1976)).

closed form solution for prices. We cover next the case where there are cross-sectional differences in cash flow risk. In this case, we can still obtain closed form solutions for prices in the absence of habit persistence, and an extremely accurate approximate solution when both habit persistence and cross sectional differences in cash flow risk are present.

III.B.1 Price dividend ratios and predictability regressions with equal cash flow risk

Assume that $\theta^i_{CF} = 0$ in (14) so that $cov\left(d\delta^i_t, dc_t\right) = \sigma^2_c$ for all $i = 1, ..., n$.[10] Then we can prove the following proposition.

Proposition 2. Let $\theta^i_{CF} = 0$ for all $i = 0, 1, 2, \cdots, n$. Then (a) the price dividend ratio is given by

$$\frac{P^i_t}{D^i_t} = a^i_0 + a^i_1 S_t + a^i_2 \frac{\overline{s}^i}{s^i_t} + a^i_3 \frac{\overline{s}^i}{s^i_t} S_t, \tag{20}$$

where a^i_1, a^i_1, a^i_2, and a^i_3 are all positive and given in (45) in the Appendix.

(b) The expected excess return of asset i is given by

$$E_t\left[dR^i_t\right] = \left(1 + \alpha\left(1 - \lambda S_t\right)\right)\left[1 + \frac{k\overline{Y} S_t \alpha\left(1 - \lambda S_t\right)}{k\overline{Y} S_t + \rho\left(1 + f\left(\overline{s}^i / s^i_t\right)\right)}\right]\sigma^2_c, \tag{21}$$

where $f(\cdot)$ is given in (48) in the Appendix and is such that $f(1) = 0$ and $f' < 0$.

Equation (20) shows that a high relative share $\overline{s}^i / s^i_t$, that signals high expected dividend growth (see (12)), translates into a high price dividend ratio. The price dividend ratio is a linear function of the surplus consumption ratio S_t as well. Intuitively, a high surplus consumption ratio implies a lower aggregate expected return and, thus, a higher price dividend ratio. Moreover, as captured by the interaction term in (20), a high relative share increases the price dividend ratio more the larger the surplus consumption ratio, S_t, as future cash flows are in this case less heavily discounted.

[10] This is a general result: In a general equilibrium setting, since for all t we have $\Sigma^n_{i=0} D^i_t = C_t$, the only case where all assets have identical cash flow risk is the case where their dividend growth have a covariance with consumption growth that equals the variance of consumption growth itself.

Expected excess returns, expression (21), are driven by both the surplus consumption ratio S_t and the relative share, $\overline{s}^i/s^i_t$. The properties of $E_t\left[dR^i_t\right]$ with respect to S_t, and the economic intuition behind them, are similar to the those of the total wealth portfolio and in the interest of space we do not repeat them here. Instead, for a given S_t, the expected excess return is increasing in the relative share $\overline{s}^i/s^i_t$, as the function $f\left(\cdot\right)$ in (21) is decreasing in it. The reason is that when $\overline{s}^i/s^i_t$ is high, asset i pays the bulk of its dividend far in the future and thus it is naturally more sensitive to movements in the stochastic discount factor than an otherwise identical asset whose cash flows are paid earlier, and, hence, less heavily discounted. This sensitivity is naturally priced as it correlates with the stochastic discount factor and hence results in higher required premium. Still, in equilibrium, the higher premium is not sufficient to offset the increase in the price dividend ratio stemming from the higher expected dividend growth. A high relative share then translates into both high price dividend ratios *and* high expected excess returns.

Notice then the key implication of the model: time varying risk preferences (S_t) and time varying expected dividend growth $(\overline{s}^i/s^i_t)$ have opposite implications for the relation between expected returns and price dividend ratios. Whereas the first induces the standard negative relation between price dividend ratios and expected excess returns, the second induces a positive relation between them. The next proposition shows how to correct the predictability regressions, for both returns and dividend growth, to account for the complex interaction between changes in expected dividend growth and changes in risk preferences.

Proposition 3. Let $\theta^i_{CF} = 0$ for all $i = 0, 2, \cdots, n$. Then (a) the expected excess rate of return can be written as,

$$E_t\left[dR^i_t\right] = b^i_0\left(S_t\right) + b^i_1\left(S_t\right)\frac{D^i_t}{P^i_t} + b^i_2\left(S_t\right)\frac{C_t}{P^i_t}, \tag{22}$$

where $b^i_j\left(S_t\right)$, $j = 0, 1, 2$, are given in (49) in the Appendix.

(b) The expected log dividend growth can be written as:

$$E_t\left[d\delta^i\right] = m_0^i\left(S_t, \mathbf{s}_t\right) + m_1^i\left(S_t\right)\frac{P_t^i}{D_t^i}, \tag{23}$$

where $m_0^i\left(S_t, \mathbf{s}_t\right)$ is given in (50) in the Appendix, and

$$m_1^i\left(S_t\right) = \frac{\phi^i}{a_2^i + a_3^i S_t}. \tag{24}$$

Expected excess returns, equation (22), are linear in both the dividend yield and the consumption to price ratio. These two predictors are obviously not orthogonal but each "picks" a different side of return predictability. Intuitively, as seen in (20), the price dividend ratio depends directly on the relative share $\overline{s}^i/s_t^i$ and hence it is more sensitive to changes in expected excess returns that result from shocks to the relative share. In contrast, it follows immediately from (20) by multiplying both sides by $s_t^i = D_t^i/C_t$, that the price consumption ratio is linear in s_t^i. Thus, it is relatively less sensitive to changes in $\overline{s}^i/s_t^i$ and relatively more to changes in the surplus S_t. As a consequence the consumption to price ratio captures better variation in $E_t\left[dR_t^i\right]$ that results from fluctuations in the aggregate discount as proxied by S_t.

Figures $2a$ and $2b$ plot $b_1^i\left(S_t\right)$ and $b_2^i\left(S_t\right)$ against S_t.[11] The non linear pattern of these coefficients is inherited from the form of the expected return on the total wealth portfolio, discussed in Section III.A, and it implies that return predictability is stronger during "bad" times, as the overall size of expected returns is higher.[12] Figures $2a$ and $2b$ also shows that a low speed of mean reversion ϕ^i yields a high slope coefficient for the dividend yield, $b_1^i\left(S_t\right)$, and a low coefficient on the consumption to price ratio, $b_2^i\left(S_t\right)$.[13]

[11] Figure 2 reports the coefficients $b_1\left(S\right)$ and $b_2\left(S\right)$ also for various values of the average cash flow risk parameter θ_{CF}^i. Section III.B.2 shows that a result similar to (22) holds also when $\theta_{CF}^i \neq 0$, although approximately.

[12] Still, simulations show that this state dependency has a minor effect on the degree of predictability of returns when both D/P and C/P are included in the regression, as we discuss at length in the empirical section.

[13] That $b_1^i\left(S_t\right)$ is decreasing in ϕ^i holds always whereas, in contrast, that $b_2^i\left(S_t\right)$ is an increasing function of ϕ^i only holds for parameter values that are empirically relevant, but not in general.

Indeed, the lower the ϕ^i, the more it takes for s_t^i to revert to $\overline{s}^i$, which implies a greater long term covariance between expected excess returns and dividend yields, as they both move slowly together with $\overline{s}^i/s_t^i$ and thus the higher coefficient on D_t^i/P_t^i, and the lower one on C_t/P_t^i.

Part (b) of Proposition 3 shows that the ability of the price dividend ratio to predict future dividend growth is mediated by the surplus consumption ratio. In particular, the size of the coefficient, $m_1^i(S_t)$, depends inversely on S_t. If, say, S_t increases but the relative share stays constant, the price dividend ratio will increase as well but the coefficient $m_1^i(S_t)$ will decrease to offset the change in P_t^i/D_t^i and leave the forecast of future dividend growth unaltered. Figure $2c$ and $2d$ plot both $m_0^i(\mathbf{s}_t, S_t)$, for $\mathbf{s}_t = \overline{\mathbf{s}}$, and $m_1^i(S_t)$ as a function of S_t for two different values of the speed of mean reversion ϕ^i. A smaller mean reversion coefficient ϕ^i leads, not surprisingly, to a lower slope coefficient on the price dividend ratio, as the same happens to the relative share $\overline{s}^i/s_t^i$ itself.

To uncover the information on future dividend growth that the price dividend ratio contains we can then proxy the denominator of $m_1^i(S_t)$ by P_t^i/C_t, which, as already argued, is relatively more sensitive to changes in S_t. Given that $s_t^i = D_t^i/C_t$ though, (23) collapses simply to a version of equation (12). That is, dividend growth should be predicted by the asset's inverse share, $1/s_t^i$, which is, in the context of our model, the only variable needed to do so.

III.B.2 Price dividend ratios and predictability regressions with heterogeneous cash flow risk

Differences in cash flow risk are a key component in determining cross-sectional differences in asset pricing. We study first the case without habit persistence to better isolate the effect that cash flow risk has on expected returns.

Proposition 4. Let $\alpha = 0$ and $Y_t = \overline{Y} = \lambda = 1$, that is, $u(C_t, X_t, t) = u(C_t, t) =$

$e^{-\rho t}\log(C_t)$. Then (a) the price dividend ratio of the asset i is

$$\frac{P_t^i}{D_t^i} = \frac{1}{\rho+\phi^i}\left(1+\frac{\phi^i}{\rho}\left(\frac{\overline{s}^i}{s_t^i}\right)\right) \tag{25}$$

(b) The expected excess return of asset i is given by

$$E_t\left[dR_t^i\right] = \sigma_c^2 + \frac{\theta_{CF}^i - \sum_{j=0}^n \theta_{CF}^j s_t^j}{1+\frac{\phi^i}{\rho}\left(\frac{\overline{s}^i}{s_t^i}\right)} \tag{26}$$

(c) The expected excess return of asset i, $E_t\left[dR_t^i\right]$ can be written as

$$E_t\left[dR_t^i\right] = b_0 + b_1^i(\mathbf{s}_t)\frac{D_t^i}{P_t^i} \qquad \text{where} \tag{27}$$

$$b_0 = \sigma_c^2 \quad \text{and} \quad b_1^i(\mathbf{s}_t) = \frac{\theta_{CF}^i - \sum_{j=0}^n \theta_{CF}^j s_t^j}{\rho+\phi^i} \tag{28}$$

In this version of the model, the price dividend ratio and expected dividend growth are perfectly correlated, and, as before, a high relative share results in a high price dividend ratio.[14] $\overline{s}^i/s_t^i$ also affects the level of expected excess return of asset i. The intuition is straightforward. Notice first that the numerator in (26) is simply the covariance between share and consumption growth (see equation (11)). Then, if this covariance is, say, positive, a low relative share translates into a high expected excess return, as now the asset, which pays when consumption is high, is a large fraction of consumption and thus risky. This effect is weighted by the ratio ϕ^i/ρ. A shock to s_t^i has a relatively higher percentage impact on the price of the asset the lower the ratio ϕ^i/ρ, because either the shock to the share is more persistent (low ϕ^i) or it has a higher weight in the pricing function (high ρ).

The implications for predictability, (27), are now immediate. If, for example, the covariance between share and consumption growth is positive, a high dividend yield should predict high future returns. As in Section III.B.1, and for the same reason, a

[14]Notice that in the limit, as $\phi^i \to \infty$, we have that $s_t^i \approx \overline{s}^i$ for all t, and the Gordon model obtains, with constant dividend growth and a constant price dividend ratio $P_t^i/D_t^i = \rho^{-1}$.

low speed of mean reversion ϕ^i will induce a large coefficient $b_1^i(\mathbf{s}_t)$ on the dividend yield. In addition, the higher the covariance between share and consumption growth, as approximated by θ_{CF}^i, the higher is the slope coefficient $b_1^i(\mathbf{s}_t)$, as in this case expected returns are higher and display a larger variation over time, while the dividend yield does not because, as can be seen in (25), it does not depend on the parameter θ_{CF}^i.[15]

Closed form solutions in contrast are not available when we combine differences in cash flow risk with habit persistence. Still, the appendix shows that an accurate approximation of the price dividend ratio is given by:

$$\frac{P_t^i}{D_t^i} \approx \widehat{a}_0^i + \widehat{a}_1^i S_t + \widehat{a}_2^i \frac{\overline{s}^i}{s_t^i} + \widehat{a}_3^i \frac{\overline{s}^i}{s_t^i} S_t, \tag{29}$$

where $\widehat{a}_0^i$, $\widehat{a}_1^i$, $\widehat{a}_2^i$, and $\widehat{a}_3^i$, are given in (52) in the Appendix. Numerical computations show that the approximation error of using (29) is less than 0.1%. (29) has identical functional form as (20) in section III.B.1. Now, though the coefficients depend on the average cash flow risk parameter θ_{CF}^i, and, generally, a high θ_{CF}^i tends to decrease P_t^i/D_t^i.[16] It follows from (29) that the predictability equations hold approximately as well, that is,

$$E_t\left[dR_t^i\right] \approx \widehat{b}_0^i(S_t) + \widehat{b}_1^i(S_t)\frac{D_t^i}{P_t^i} + \widehat{b}_2^i(S_t)\frac{C_t}{P_t^i}. \tag{30}$$

$$E\left[d\delta_t^i\right] \approx \widehat{m}_0^i(S_t,\mathbf{s}_t) + \widehat{m}_1^i(S_t)\frac{P_t^i}{D_t^i} \tag{31}$$

where $\widehat{b}_j^i(S_t)$, $j = 0,1,2$, $\widehat{m}_0(S_t,\mathbf{s}_t)$ and $\widehat{m}_1(S_t)$ are given in the Appendix and are comparable to their counterparts in Section III.B.1. Now, of course, $\widehat{b}_j^i(S_t)$, $j = 0,1,2$

[15] The cash flow risk parameter θ_{CF}^i does not enter into the pricing equation (25). Under logarithmic preferences substitution and wealth effects exactly balance each other out: a high θ_{CF}^i yields a higher discount in its price due to the substitution effect but a higher expected return, which induces a wealth effect.

[16] Briefly, the essential technical problem is that the drift rate of dq_t^i , where $q_\tau^i = s_\tau^i Y_\tau$, contains the quadratic Ito term $(Y_t - \lambda) s_t^i \sum_{j=1}^n s_t^j \theta_{CF}^j$, which complicates the computation of the expectation in (15). Condition (10) however suggests that this term should be "small." Our estimates in Section IV.B indeed confirm that its value is negligible and thus we approximate q_t^i with a process $\widehat{q}_t^i$ defined as $d\widehat{q}_t^i = dq_t^i - (Y_t - \lambda) s_t^i \sum_{j=1}^n s_t^j \theta_{CF}^j dt$.

depend on θ_{CF}^i as well. Figures $2a$ and $2b$ show though that cross sectional differences in θ_{CF}^i generate small cross sectional differences in the predictability coefficients, $\widehat{b}_1^i\left(S_t\right)$ and $\widehat{b}_2^i\left(S_t\right)$. Finally, figure $2d$ shows that $\widehat{m}_1\left(S_t\right)$ is increasing in θ_{CF}^i. A high θ_{CF}^i implies that the price dividend ratio is more sensitive to cash flow risk, which is correlated with expected dividend growth, that is, the relative share $\overline{s}^i/s_t^i$. Movements in the latter are then more correlated with movements in the price dividend ratio and, thus the increase in $\widehat{m}_1\left(S_t\right)$.

IV. EMPIRICAL EVALUATION

IV.A Data, construction of the cash flow series, and cointegration tests

Quarterly dividends, returns, market equity and other financial series are obtained from the CRSP database, for the sample period 1947-2001. We focus our empirical exercises on a set of twenty value-weighted industry portfolios. The use of portfolios has become a relatively standard procedure in the financial literature in order to mitigate the residual variance of returns. In our case, the use of portfolios enables us to obtain relatively smooth cash flow data that are a-priory consistent with the underlying model for cash flows put forward in this paper (equation (8)). More explicitly, our general equilibrium model focuses on assets that are predicted to last forever. For this reason, we choose to concentrate our empirical exercises on a very coarse definition of industries – the first two SIC codes[17] – which are likely to generate cash flows for a very long time. The secular trend of an industry, if any, can then be interpreted in our framework as a convergence towards its long-term mean $\overline{s}^i$, which is assumed to be known by agents, a strong assumption that we discuss further in the conclusions.

[17] The two digit SIC grouping are similar to those employed by Moskowitz and Grinblatt (1999). The SIC codes are obtained from CRSP, which reports the time-series of industry classification codes. Although the COMPUSTAT classification is considered to be more accurate, the series is modified only from 1994, which leads to a survival bias.

A detailed description of the construction of the cash flow series is contained in the appendix and it follows Hansen, Heaton and Li (2003) and Bansal, Dittmar, and Lundblad (2002). Our cash flow series includes both dividends as well as share repurchases, where our construction of the repurchases series is similar to that of Jagannathan, Stephens and Weisbach (2000) and Grullon and Michaely (2002). With some abuse of terminology we use the expressions "cash flow" and "dividend" interchangeably throughout. Finally consumption is defined as per capita consumption of non durables plus services, seasonally adjusted, and it is obtained from NIPA. We use the personal consumption expenditures deflator (PCE) to convert nominal quantities into real quantities, whenever necessary.

From dividends and consumption, we construct our key cash flow variable $s_t^i = D_t^i / C_t$. Our framework assumes that for every asset i, $\log(D_t^i) - \log(C_t) = \log(s_t^i)$, which is stationary, and hence that log dividend and log consumption are cointegrated series with cointegrating vector $[1, -1]$. In the last column of Panel C in Table 1 we test the null hypothesis of no cointegration by using the cointegration test developed by Horvath and Watson (1995), a test that explicitly exploits knowledge of the cointegrating vector under the alternative hypothesis, as in our case. First, as can be seen in the very last line of the panel, log dividends and log consumption are indeed cointegrated for the case of the market portfolio, as we can reject the null of no cointegration at the 1% level. At the individual industry level, we can reject the null of no cointegration for twelve industries at the 5 % level – seven of which at the 1% level. Failure to reject no cointegration for some industries is to be expected even within our model, as while the share s_t^i slowly converges to its long run value $\overline{s}^i$, it may entail a trend in the data due to small sample.

IV.B Choice of parameters

As in Campbell and Cochrane (1999), we choose preference parameters to match basic moments of the market portfolio. We use the stationary distribution of the process

Y_t, given in equation (40) in the Appendix, to calibrate these parameters to the market expected excess rate of return, its Sharpe ratio and the mean risk free rate.[18] In addition, simulations show that a high value of λ is necessary to match the long run predictability of stock returns at reasonable levels and for this reason we set $\lambda = 20$. This level implies a lower bound for the local curvature of the utility function that matches that of Campbell and Cochrane (1999). We also set $\overline{Y} = 34$ to match the steady state value of the instantaneous local curvature of the utility function in Campbell and Cochrane (1999).[19] Table 1 Panel A lists the values of the remaining parameters. Panel B contains the moments obtained for the aggregate portfolio.

As for the share process (8), we estimate the speed of mean reversion ϕ^i and the long term mean $\overline{s}^i$ by applying time series linear regressions to their discretized version. In order to estimate the parameter θ^i_{CF}'s, we appeal to the relation (14), which, under condition (10), implies that $\theta^i_{CF} = E\left[cov_t\left(d\delta^i_t, dc_t\right)\right] - var\left(dc_t\right)$. Since $E_t\left[dc_t\right]$ is constant, we therefore simply find that $\theta^i_{CF} = cov\left(d\delta^i_t, dc_t\right) - var\left(dc_t\right)$. The results of the estimation are contained in Panel C. Given the central role that the speed of mean reversion parameter ϕ^i plays in the interpretation of our results below, we have ordered the industries by the decreasing size of the estimated ϕ^i.

IV.C The model's pricing ability

Table 2 reports the results of both cross sectional and time series regression of the

[18]We match the unconditional moments for the total wealth portfolio with the average values for the *market* portfolio in the data. Clearly the two portfolios are different, but simulations show that their unconditional moments are similar. The availability of an analytical formula for the stationary density of Y_t favors the use of the total wealth portfolio for calibration purposes.

[19]This value of $\overline{Y}$ is slightly higher than the one in Campbell and Cochrane (1999). The steady state surplus consumption ratio these authors is $\overline{S} = .0676$. Given that they work with $\gamma = 2$ rather than with $\gamma = 1$, their steady state local curvature is $\frac{\gamma}{\overline{S}} \approx 29.6$. The difference is due to the fact that our results match the numerical average in our generated Campbell and Cochrane (1999) surplus consumption series.

form

$$\ln\left(\frac{P_t^i}{D_t^i}\right) = \alpha_0^i + \alpha_1^i \ln\left(\widehat{PD}_t^i\right) + \varepsilon_t^i, \tag{32}$$

where $\widehat{PD}_t^i$ is the model implied price dividend ratio, obtained by feeding into formula (29) the realized time series of the relative share and the surplus consumption ratio for the parameter values contained in Table 1. We use two proxies for the surplus consumption ratio S_t, which is not observed. The first is obtained from innovations in real consumption growth, $\widehat{\Delta B}_t = \frac{1}{\sigma_c}\left(\Delta c_{t+1} - E_t\left[\Delta c_{t+1}\right]\right)$. We use this series to approximate the Brownian motion appearing in the process for the inverse surplus Y_t in (4) and thus compute $S_t = 1/Y_t$ for all t. The second proxy relies on approximating the total wealth price consumption ratio, P_t^{TW}/C_t, by the price dividend ratio of the market, P_t^{Mkt}/D_t^{Mkt}. Inverting formula (17), we can then extract S_t from P_t^{Mkt}/D_t^{Mkt} for all t.

Panel A of Table 2 reports the result of Fama-MacBeth cross sectional regressions, where $\ln\left(P_t^i/D_t^i\right)$ is regressed, period by period, on a constant and the model implied $\ln\left(\widehat{PD}_t^i\right)$. The table reports the average intercept and slope, together with their standard errors, which are Newey-West adjusted for heteroskedasticity and autocorrelation. Note first that, independently of whether S_t is extracted from consumption growth or from the market price dividend ratio, the average slope coefficient, α_1, is .55 and significantly different from zero at the 5% level. Still, we can reject that $\alpha_0 = 0$ and $\alpha_1 = 1$. The model then is able to generate a substantial component of the cross sectional dispersion of the price dividend ratios observed in the data, although not completely.

Panel B reports the results of time series regressions, industry by industry, as in (32). When S_t is extracted from real consumption, the slope coefficient α_1^i is significantly different from zero for most industries, namely, fifteen out of twenty. In other words, even when only cash flows and consumption data are used in the fitting exercise, and in particular no information from asset prices contaminates the inference, the model shows that relative shares and surplus combine to yield a good description of the dynamics of asset prices. When we use information from the price dividend ratio of the market portfolio to compute S_t, the results improve substantially, and essentially all regression coefficients are different from zero. Moreover, in this case we cannot reject the null that $\alpha_0^i = 0$ and $\alpha_1^i = 1$ for fourteen of the twenty industries in our sample. Figure 3 provides a visual impression of the results contained in Table 2 Panel B by plotting the log price dividend ratios of the market portfolio and three industries, Financials, Paper, and Utilities (solid-dot line) together with the corresponding model predictions (dot line). These three industries were chosen because they provided the best, the median and worst fit in terms of the R^2 of the time series regression (32).

IV.D The predictability of dividend growth

Our cash flow model implies that the relative share forecasts future dividend growth. To test this implication we run,

$$\Delta d_{t,t+\tau}^i = \beta_0^i + \beta_X^i X_t + \varepsilon_{t,t+\tau}^i \qquad \text{for} \qquad \tau = 1,\ 4 \text{ and } 7 \text{ years}, \tag{33}$$

where X_t is either $\overline{s}^i/s_t^i$, P_t^i/D_t^i, or, in a multivariate regression, both of them.

Panel A of Table 3 shows the results for the market portfolio. The relative share is a strongly significant predictor of dividend growth at the aggregate level with R^2s equal to 5%, 31% and 41% for the 1, 4, and 7 years horizon respectively. Instead the price dividend ratio is never significant and, moreover, it enters with the wrong sign – a standard result in the predictability literature. There is then substantial predictability of dividend growth at the aggregate level but the price dividend ratio is not capturing it.[20]

Turning to individual industries, Panel B of Table 3 reports the results of a pooled regression (33), with fixed industry effects, for the same three horizons. The relative

[20] Ang and Bekaert (2002) find that dividend *and* earnings yields predict cash flow growth in a cross section of five countries at short horizons and Ang (2002) finds predictability using dividend yields in the US for the sample period 1927-2000. Lettau and Ludvigson (2003) find predictability in dividend growth at longer horizons at the aggregate level using a variable that measures deviations of consumption from a stable relation with dividends from human and non human (financial) wealth. Ribeiro (2002) finds that the ratio of dividends to labor income contain also information about future dividend growth.

share $\overline{s}^i/s^i_t$ is again the strongest predictor of future dividend growth, with a R^2's equal to 6%, 20% and 24% for the three horizons, respectively. Now, however, also the price dividend ratio enters significantly for the one and four year horizon, although its predictive power is much weaker: the R^2 never exceeds 10% across horizons. In a multivariate regression, the relative share remains the strongest predictor.[21]

Panel C of Table 3 reports the predictive regression results for individual industries. In the interest of space we report only the four year regression. As it can be seen, and in line with the results of Panel B, the relative share is a significant predictor of future dividend growth for fifteen of the twenty industries, with R^2s above 30% in ten cases. Among the industries for which the relative share does not forecast future dividend growth are those for which the null of no cointegration could not be rejected (see Table 1), like Mining, Food, Utilities, and Manufacturing, so the lack of forecasting ability of the relative share is perhaps less surprising. The price dividend ratio in contrast is only significant for four of the twenty industries. As before, the relative share remains the strongest predictor of the two in the multivariate regression.[22]

Price dividend ratios and future dividend growth

The price dividend ratio is, at best, a weak predictor of future dividend growth. Is this weak forecasting ability of the price dividend ratio to be expected even in the presence of substantial dividend growth predictability? To address this question, we run regression (33) in simulated data and compare the results to those in Table 3. For

[21]Our cash flow model implies that the relative share $\overline{s}^i/s^i_t$ should be the only predictor of future dividend growth. Thus, the finding that the price dividend ratio P^i_t/D^i_t still remains significant in the multivariate regression can be taken as a rejection of the cash flow model. Still price dividend ratios are likely to be less affected by measurement error than dividend consumption ratios and this can account for this result.

[22]For the one and seven year horizons, the relative share is significant for ten and fifteen industries, respectively. Instead, the price dividend ratio is significant for only eight and four industries for these two horizons. In addition neither the significance of the relative share nor the R^2s were affected by the inclusion of lagged dividend growth.

this purpose we obtain 40,000 quarters of artificial data by generating dividend and consumption shocks for the processes described in Section II using the parameter values in Table 1.

The results are contained in Panel A of Table 4, for the one and four year horizon. The price dividend ratio is a substantially worse predictor of future dividend growth than the relative share $\overline{s}^i/s_t^i$, as the former is always associated with a much smaller R^2 than the latter. Indeed, for ten out of twenty industries, the R^2 associated with the price dividend ratio is less than half that associated with $\overline{s}^i/s_t^i$. To put it differently, although the model entails a substantial predictability of future dividend growth across industries, the price dividend ratio fails to capture it.

We can use the simulation results also to compare the magnitudes of the predictive regression coefficients obtained in the data, Panel C of Table 3, with those in simulations which we may take as population values. The magnitude of the coefficients is similar. For the four year regression, we can reject that the coefficients in Table 3 are different from their population counterparts for only five out of the twenty industries. In addition, and as in the data, the R^2 increases with the horizon, although it is slightly lower than the one obtained in the data, with values that range from 30% to 0% at the four year horizon. Finally, in results not reported here, for the one year horizon we can never reject the null that the coefficients on $\overline{s}^i/s_t^i$ are equal in the data and in simulations.

To see whether the model captures the observed cross sectional differences in dividend growth predictability, we also regress cross-sectionally the coefficients obtained from the predictive regressions in the empirical data on their population counterparts. That is we run

$$\widehat{\beta}_X^{Data,i} = a_0 + a_1 \widehat{\beta}_X^{Simulation,i} + \varepsilon^i \qquad \text{where} \qquad X_t \in \left\{ \overline{s}^i/s_t^i, P_t^i/D_t^i \right\}. \tag{34}$$

This exercise allows for a simple metric to evaluate the model: the intercept, a_0, should equal zero and the slope coefficient, a_1, should equal one. The results are in Panel B of Table 4.

Start with the case of the relative share. For both the one year and four year horizon, the slope a_1 is significantly different from zero and positive, and it is not significantly different from one. The intercept a_0 is statistically different from zero at the one year horizon but it is not at the four year horizon. That is, the cross sectional variation in the predictability of future dividend growth implied by the model matches the data well. Indeed, at the four year horizon, even regressing the predictive regression R^2 in the data on its population counterpart results in a slope coefficient a_1 that is positive and significantly different from zero, but not significantly different from one. For the one year horizon, instead, the slope a_1 for the R^2 regression is positive but insignificant. Similar results hold for the price dividend ratio.

A noticeable difference between the results in Table 3 and Table 4 is that the level of the R^2s in the data is generally higher than in the simulations, which may be the result of small sample problems. To address this issue we simulate 1000 samples of artificial data of 54 years each. On each sample, we performed the same predictive regressions, thereby obtaining the distribution of the regression coefficients as well as that of the R^2's. The results for the four year ahead predictive regressions case are plotted in figure $4a$.[23]

In each of the subpanels we report the 90% confidence intervals for the coefficient on the relative share, the price dividend ratio, and the R^2 (solid lines) along with its estimated value from the data (stars) for each of the industries in our sample. Recall that the industries are ordered according to the decreasing size of the speed of mean reversion coefficient, ϕ^i. The first plot in figure $4a$ shows that indeed the confidence bands display the predicted downward sloping pattern, and, moreover, the data coefficients nicely fit in these bands for all but two cases. As for the R^2, all but one of the R^2's from the data lie inside the 90% confidence interval obtained from simulations. Notice though that, in general, the data R^2s are biased upwards compared to their counterparts in the Table 4.

[23]Results for the one year predictive regression are similar.

As for the price dividend ratios, for all but five industries, the regression coefficient in the data fits into the confidence bands obtained from simulations. In addition, from the last subpanel in the first row, the predictive power of the price dividend ratio is smaller than the one of the relative share, as already discussed.

IV.E The predictability of stock returns

Our model provides for a role for both the dividend yield and the consumption to price ratio in the return predictability regressions. To test this implication we run

$$r^i_{t,t+\tau} = \beta^i_0 + \beta^i_{D/P}\left(\frac{D^i_t}{P^i_t}\right) + \beta^i_{C/P}\left(\frac{C_t}{P^i_t}\right) + \varepsilon^i_{t,t+\tau}, \tag{35}$$

where $r^i_{t,t+\tau}$ denotes the cumulative log excess return on asset i between t and $t+\tau$, and where $\tau = 1$, 4 and 7 years.

Table 5 Panel A reports the result of running (35) for the aggregate market portfolio. The first three columns show the familiar pattern regarding return predictability: Both the regression coefficient and R^2 increase with the forecasting horizon. Interestingly, the predictability of future return is rather strong notwithstanding the inclusion of the 1990's in the sample, which was known to have reduced significantly the predictive ability of the dividend yield. The reason is that our dividend series accounts for repurchases, which has mitigated the drop in the dividend yield in the 1990's. The multivariate regression instead shows that the consumption to price ratio has no predictive power for future returns. This result is consistent with our model however, as the speed of mean reversion of the relative share of the market is quite low, only $\phi^i = .07$ as shown in Table 1, thereby yielding a stronger role for the dividend yield than for the consumption to price ratio in the predictability regression. Simulations, contained in Table 6 and discussed below, confirm this finding.

Table 5 Panel B reports the result of a pooled predictive regression with fixed industry effects. The dividend yield is still a strong predictor of future returns. Now, however, the consumption to price ratio also enters significantly in the multivariate regression, showing that at the individual industry level, a second predictor is necessary to uncover the information that prices contain about future returns. This finding is consistent with our model, since as shown in Table 1, many industries have a faster mean reversion of the share process than the market and thus the stronger the role of C_t/P_t^i for individual industries.

This intuition is confirmed when we look at individual industries. Panel C reports the four year ahead predictive regression for each industry in our sample. The dividend yield alone predicts future returns for nine of the twenty industries, especially for those cases where the speed of mean reversion is low, as the model requires. When we add the consumption to price ratio in a multivariate regression, the number of industries for which returns are predictable increases to thirteen out of twenty. In many occasions, the adjusted R^2 jumps up significantly, from levels such as 3% (Mining) and 8% (Petroleum), to 37% and 48%, respectively. In addition, and consistently with our model, most of the cases where the consumption to price ratio is significant occur for industries that have a strong mean reverting share ϕ^i.[24]

Finally, the theoretical results in (30), and figure 2a, suggest that the regression coefficients $\beta_{D/P}^i$ and $\beta_{C/P}^i$ should also depend on the surplus consumption ratio. Yet, simulations of the model, reported next, show that D_t^i/P_t^i and C_t/P_t^i capture most of the information on dividend growth and the aggregate discount, which is at the source of the predictability result. That is, while adding the consumption to price ratio may even double the (adjusted) R^2 in a predictive regression, the inclusion of the valuation ratios interacted with S_t, for instance, increases the adjusted R^2 by less than 1%. This is also confirmed in the data: adding S_t improves little the predictive power of (35). We omit these additional results in the interest of space.

Simulation evidence

Table 6 Panel A reports the results of predictive regressions (35) on a 40,000

[24]For the one and seven year ahead forecasting regressions, the dividend yield is significant for fifteen and fourteen industries, respectively. In a multivariate regression, C_t/P_t^i is significant in five and ten industries for the two horizons, respectively.

quarters sample of artificial data. As it was the case in the data, the predictability of the market portfolio return is unaffected by the consumption to price ratio. At the industry level in contrast, the consumption to price ratio adds substantially to the predictability of the dividend yield, especially for industries with a high speed of mean reversion ϕ^i. For instance, for a number of industries, the (adjusted) R^2 goes from the 7-8% range to the 14-15% range.

The regression coefficients obtained from data are of the same order of magnitude as the simulated ones, both for the dividend yield and the consumption to price ratio, although the data show a much higher dispersion of the coefficients. For the four year horizon, for all but three industries we cannot reject the hypothesis that the coefficient on D_t^i/P_t^i is the same in the data and in simulations. In the multivariate regression we can reject that the coefficient on C_t/P_t^i in the data equals its population counterpart for six cases.[25]

Panel B of Table 6 shows that the model captures the cross-sectional differences in predictability, as the predictive regression coefficients from the data nicely line up with the population ones obtained in simulations, for both one and four years. For instance, a cross-sectional regression of the univariate four year predictability regression coefficient $\beta_{D/P}^i$ obtained from the data (Table 5) on its population counterpart, similar to (34) in the previous section, results in an intercept that is statistically not different from zero and a slope coefficient that is strongly different from zero and not different from one. The results concerning $\beta_{C/P}$ are very similar and we omit a discussion in the interest of space.

As it was the case in the dividend growth predictability regressions, the results related to R^2's are the weakest. The slope coefficients have the right sign for the four year ahead predictive regression results – that is, those industries that display more

[25] In results not reported here, for the one year regression, for all but one industry we cannot reject the null that the coefficients in the data and the simulations are the same. For the coefficient on C_t/P_t^i in the multivariate regression, we can only reject the same null for three industries.

return predictability according to the model do as well in the data – but they are not significantly different from zero. Indeed from Panel A of Table 6 we see that the R^2's are lower in simulation than the corresponding ones in empirical data. There are two reasons for this difference. The first is that we were unable to match the volatility of returns to its empirical counterpart in the calibration. As shown in Table 1, the volatility of returns in our model is 23% versus the 16% observed in the historical sample. Although population values for average predictability can be obtained by resorting to the long sample, the R^2 would clearly be lower because of the higher volatility implied by our model.

The second reason is that the R^2 is biased upwards relative to the population one due to the highly persistent nature of both the dividend yield and the consumption to price ratio. To investigate this potential small sample problem we turn to the 1000 samples of 54 years each introduced in Section IV.D. The results are reported in figure 4*b*. There we show the data estimates (stars) together with the 90% confidence bands (solid lines) obtained out of the 1000 samples, for $\beta^i_{D/P}$, $\beta^i_{C/P}$, and R^2. The two slope coefficients are within the 90% confidence bands for nineteen and fourteen of the twenty industries, respectively. As shown in the last panel the simulated confidence bands for the R^2s are rather wide and only five out of the twenty R^2s in the data do not fall within these bands.

IV.F The connection between dividend growth and return predictability

Where do the cross sectional differences in the ability of the dividend yield or the consumption to price ratio to predict future returns come from? We showed in Section III that a high speed of mean reversion ϕ^i should be associated with a low (regression) coefficient of the dividend yield, and a high coefficient of the consumption to price ratio. In contrast, differences in the cash flow risk parameter θ^i_{CF} had little impact on $\beta^i_{D/P}$ and $\beta^i_{C/P}$. To check whether these predictions are met in the data, Table 7 reports the result of the following linear cross-sectional regressions

$$\beta^i_{D/P} = \alpha_0 + \alpha_1 \phi^i + \alpha_2 \theta^i_{CF} + \varepsilon^i \qquad (36)$$

$$\beta^i_{C/P} = \alpha_0 + \alpha_1 \phi^i + \alpha_2 \theta^i_{CF} + \varepsilon^i \qquad (37)$$

where $\beta^i_{D/P}$ and $\beta^i_{C/P}$ are the regression coefficients obtained from the return predictive regression (35), and ϕ^i and θ^i_{CF} are the characteristics of the industry portfolio cash flow process in Table 1. We run these two cross-sectional regressions for each of the three predictive horizons, $\tau = 1, 4$, and 7 years.

Panel A of Table 7 shows that for regression (36) the slope coefficient α_1 is negative, as the model predicts, and statistically significant for all three horizons. In addition, α_1 becomes more negative with horizon τ, as the dispersion of $\beta^i_{D/P}$ increases with it. The coefficient on θ^i_{CF}, α_2, is in contrast never significant for any horizon. As already suggested by figure 2, cross sectional variation in θ^i_{CF} has little impact on the cross sectional variation of $\beta^i_{D/P}$. For the regression (37), Panel A shows that $\alpha_1 > 0$ and $\alpha_2 < 0$ for all horizons, though never significantly different from zero.

Panel B of Table 7 reports the same cross-sectional regression results but now for simulated data. We also report standard errors for the simulated data, as the cross-sectional regression is made with only a twenty portfolios, and thus the coefficients from simulations are less likely to constitute "population values". The results are similar to those observed in empirical data: the coefficient α_1 in regression (36) is negative, significantly different from zero, and of similar magnitude for all horizons. All the other coefficients, in either (36) and (37), are not significantly different from zero. The effect of ϕ^i and θ^i_{CF} on $\beta^i_{C/P}$ is limited even in simulations. In contrast, the model successfully reproduces the link between return predictability and the underlying cash flow properties of the asset, captured by the strong negative relationship between ϕ^i and $\beta^i_{D/P}$.

V. CONCLUSIONS

Prices, which are observable, contain information about expected returns and expected future cash flows, which are not. In this paper we combine a novel model of cash flows with the stochastic discount factor model introduced by Campbell and Cochrane (1999) that allows for a complete characterization of individual price dividend ratios and expected excess returns. The model produces several new insights about the relationship between valuation ratios, expected returns and expected future cash flows. First, while changes in risk preferences generate the traditional negative relation between price dividend ratios and expected returns, the variation in expected dividend growth generates a positive one. Intuitively, a higher expected dividend growth implies both a higher price dividend ratio and a riskier asset as its cash flows are expected farther in the future. The asset's duration is longer and thus the higher expected returns. Second, these offsetting effects weaken the ability of the dividend yield to forecast returns, and essentially eliminate its ability to predict future dividend growth. Third, predictive regressions for returns should include both the dividend yield and the consumption to price ratio to disentangle the effect that changes in risk preferences and expected dividend growth have on prices and returns. Fourth, there is a direct link between the predictability of future cash flows – by the share of dividends over consumption in our model – and the ability of the dividend yield to predict future returns. We use simulations to quantify the magnitude of these empirical predictions, and find that they are met with substantial support in the data.

We argued that the cash flow model introduced in this paper has a number of attractive properties like tractability and plausibility. A strong assumption though is that agents have perfect knowledge of the long term average size of each industry. This assumption can be relaxed, however, to assume that agents *learn* about this long run relative size over time, thereby implying a time varying expected long term average size $\overline{s}_t^i = E_t\left[\overline{s}^i\right]$. To see the effect of this extension intuitively, consider the effect of a positive dividend shock that places the current share above its long run mean. When the long run mean $\overline{s}^i$ is known, this shock signals poor dividend growth in the future as the share mean reverts to $\overline{s}^i$. For this reason the price dividend ratio decreases in response to a positive dividend shock. In contrast, in the presence of learning effects, the agent may actually update upwards his expectation $\overline{s}^i_t = E_t\left[\overline{s}^i\right]$. Hence, the price may react more strongly to positive news about cash flows, inducing possibly a positive correlation between the price dividend ratio and current dividend shocks.

As both Campbell (2000, page 1529) and Cochrane (2001, page 11) have recently emphasized, focusing on prices and not only on returns should be the ultimate object of interest in the field of asset pricing. Prices contain useful information about unobserved agents' expectation on returns and future cash flows. For instance, much has been made recently of the role of conditioning information in tests of the cross section of stock returns. The selection of variables that are useful as proxies for the agents' information set may be greatly helped by looking at whether they can capture the time series and cross sectional variation observed in price dividend ratios. This paper makes progress in this direction by emphasizing an empirically relevant model that allows for closed form solutions of the present value relations.

APPENDIX

(A) Cash flow data

The cash flow series that we use in this paper includes both dividend distributions as well as share repurchases where our construction of the latter follows closely that of Jagannathan et al. (2000) and Grullon and Michaely (2002).[26] Repurchases are defined as total expenditure on the purchase of common and preferred stocks (Compustat data item #115) minus any reduction in value (redemption value, Compustat data item #56) of the net number of preferred stocks outstanding (not available for banks, utilities, and insurance companies.) Our empirical exercise focuses on a set of twenty industry portfolios, identified by their first two SIC codes. Following Fama and French (1997), we form the portfolios in July of every year by assigning firms to them according to their SIC classification. Time series of value weighted returns and cash flows are then calculated for each of the portfolios.

Given the aggregate cash flow data for dividends and repurchases as well as the series for returns, we then follow a procedure similar to Hansen, Heaton, and Li (2002) and Bansal, Dittmar, and Lundblad (2002), to construct cash flow series that are consistent with an initial investment in each industry, as our model requires. To understand the logic, it is useful to concentrate on an individual firm. The methodology naturally extends to industries as well as the market as a whole. Consider then a firm i. Let p_t^i and N_t^i be the price per share and the number of shares outstanding at the beginning of period t, and let $\widetilde{D}_{t,t+1}^i$ and $\widetilde{F}_{t,t+1}^i$ be the total amount of dividends and repurchases made between t and $t+1$. If an investor purchases n_t shares at the beginning of time t for $P_t^i = n_t p_t^i$, the total amount of cash flows obtained at time $t+1$ is then

$$D_{t,t+1}^i = \frac{n_t}{N_t^i}\left(\widetilde{D}_{t+1}^i + \widetilde{F}_{t,t+1}^i\right) = \left(d_{t+1}^i + f_{t+1}^i\right) P_t^i, \tag{38}$$

where, denoting the market capitalization by $V_t^i = N_t p_t^i$, we denote $d_{t+1}^i = \widetilde{D}_{t,t+1}^i / V_t^i$, and $f_{t+1}^i = \widetilde{F}_{t,t+1}^i / V_t^i$. Assuming that repurchases occur at the end of period t, the

[26] We thank Harry DeAngelo for suggestions and insightful comments in constructing the data.

total amount of shares outstanding after repurchases (but before new issues) is $N_{t+1}^i = N_t^i - \widetilde{F}_{t,t+1}^i / p_{t+1}^i$. At time $t+1$, the investor then holds $n_{t+1} = n_t - (n_t/N_t^i)\left(N_t^i - N_{t+1}^i\right)$ shares. This implies that for every t,

$$P_{t+1}^i = n_{t+1} p_{t+1}^i = n_t p_t^i \left(\frac{p_{t+1}^i}{p_t^i} - \frac{\left(N_t^i - N_{t+1}^i\right) p_{t+1}^i}{N_t^i p_t^i} \right) = P_t^i \left(h_{t+1}^i - f_{t+1}^i\right) \tag{39}$$

where $h_{t+1}^i = p_{t+1}^i / p_t^i$ is the capital gain from an investment at time t. We apply equations (38) and (39) recursively to compute the cash flows of value-weighted industry portfolios as well as of the market portfolio. We smooth the dividend series to correct for seasonalities by using a trailing four quarter average. We finally assume that the total initial investment made at time $t = 0$ (1947) equals the total market capitalization at that time, divided by the population M_t, $P_0^j = V_0^j / M_0$. The flow of dividends in (38) therefore can be interpreted as the cash flows to a typical (representative) investor in the economy who invested in 1947.

What is the corresponding consumption C_t that flows to this representative investor? We simply assume that it is given by the per-capita aggregate consumption $C_t = \widetilde{C}_t / M_t$, where $\widetilde{C}_t$ is defined as total consumption expenditures of non-durables plus services, obtained from the National Income and Product Accounts (NIPA) for the period 1947-2001. The definition $C_t = \widetilde{C}_t / M_t$ has three appealing features: First, it is the natural assumption in an economy where population grows at the same rate as the economy. Second, it is not contaminated by variables that are related to prices, as other plausible definitions would have it. Third, data strongly endorse it, as we find that $\log(C_t)$ is in fact strongly cointegrated with $\log\left(D_t^A\right)$, where D_t^A denote the aggregate dividends constructed according to the procedure above (see Table 1).[27]

Whenever we need to convert nominal quantities into real ones, we use the personal

[27]Our empirical findings turn out to be rather insensitive to the rescaling of both dividends and consumption. Using simply the total amount of dividends paid every period t from each industry, and using total market capitalization as price level P_t, for instance, yields very similar results. Other types of rescaling also had minor effects on the results in this paper.

consumption expenditures deflator (PCE), also obtained from the NIPA tables.

(B) The calibration of preference parameters

The stationary density for the process Y in equation (4) depends only on three parameters, $\overline{Y}$, λ and $b = k/\left(\alpha^2\sigma_c^2\right)$ and it is given by

$$\psi(Y) = \frac{e^{-2b\frac{\overline{Y}-\lambda}{(Y-\lambda)}} \times (Y-\lambda)^{-2b-2}}{\int_\lambda^\infty e^{-2b\frac{\overline{Y}-\lambda}{(y-\lambda)}} \times (y-\lambda)^{-2b-2}\,dy}. \tag{40}$$

We use equation (40) to compute the unconditional moments of aggregate variables. We match these unconditional moments to their sample counterparts in the calibration described in Section V. Specifically, we choose the parameters $\overline{Y}$, λ, k and α to match the following moments:

$$E\left[dR_t^{TW}\right] \quad E\left[r^f(Y)\right] \quad E\left[\sigma_{r_f}^2(Y)\right] \quad E\left[\frac{P}{C}(Y)\right] \quad \frac{E\left[dR_t^{TW}\right]}{\sqrt{E\left[\left(dR_t^{TW}\right)^2\right]}},$$

to the corresponding sample counterparts. For instance, for the Sharpe ratio, we make

$$\frac{E\left[dR_t^{TW}\right]}{\sqrt{E\left[\left(dR_t^{TW}\right)^2\right]}} = \frac{\int_\lambda^\infty \mu^{TW}(Y)\,\psi(Y)\,dY}{\sqrt{\int_\lambda^\infty \|\sigma^{TW}(Y)\|^2\psi(Y)\,dY}} = \text{Data.}$$

(C) The share process

We prove the claim made in Section II.B that $s_t^i \geq 0$ and $\sum_{i=0}^n s_t^i = 1$ for all t. The process described in equation (8) is a special case of a more general VAR model with

$$ds_t^i = \left(\sum_{j=0}^n s_t^j \lambda_{ji}\right) dt + s_t^i \left(\mathbf{v}^i - \mathbf{s}_t'\mathbf{v}\right) d\mathbf{B}_t'$$

where $\mathbf{v}$ is a $N \times N$ matrix $\mathbf{v} = \left[\mathbf{v}^0, ..., \mathbf{v}^n\right]'$, (I) $\lambda_{ji} > 0$ and (II) $\lambda_{ii} = -\sum_{k\neq i}\lambda_{ik}$. A solution to this system of stochastic differential equations is known to exist (see Liptser and Shyriaiev (1977, Ch. 9)). In addition, $\Pr\left(s_\tau^i > 0\right) = 1$ for all finite τ, and it is easy to see that if $\sum_{i=0}^n s_0^i = 1$, then $\sum_{i=0}^n s_t^i = 1$ for all $t > 0$, as $d\left(\sum_{i=0}^n s_t^i\right) = 0dt + 0dB_t = 0$. Finally, we have to show that (8) can be obtained as a special case for a choice of the parameters λ_{ji}. We concentrate then only on the drift. First, for $i = 1, .., n$, we can impose the condition $s_t^0 = 1 - \sum_{j=1}^n s_t^j$, so that $\sum_{j=0}^n s_t^j \lambda_{ji} = s_t^0 \lambda_{0i} + \sum_{j=1}^n s_t^j \lambda_{ji} = \lambda_{0i} + \sum_{j=1}^n s_t^j (\lambda_{ji} - \lambda_{0i})$. Thus, imposing the restriction $\lambda_{0i} = \phi^i \overline{s}^i$ and $\lambda_{ji} - \lambda_{0i} = 0$ for $j \neq i, 0$ we find that the drift of share ds_t^i is

$$\sum_{j=0}^n s_t^j \lambda_{ji} = \phi^i \overline{s}^i + s_t^i (\lambda_{ii} - \lambda_{0i}) = \phi^i \overline{s}^i + s_t^i \left(\lambda_{ii} - \phi^i \overline{s}^i\right)$$

The next step is to show that we can define $\lambda_{ii} - \phi^i \overline{s}^i = -\phi^i$ without violating conditions (I) and (II) above. Imposing (II), we have $\lambda_{ii} = -\sum_{k \neq i} \lambda_{ik} = -\lambda_{i0} - \sum_{k \neq i,0} \lambda_{ik}$. Since the above restrictions hold for each $k = 1, ..., n$, we have $\lambda_{ik} = \lambda_{0k} = \phi^k \overline{s}^k > 0$. Thus, $\sum_{k \neq i,0} \lambda_{ik} = \sum_{k \neq i,0} \phi^k \overline{s}^k$ which implies that

$$\lambda_{ii} - \phi^i \overline{s}^i = -\phi^i \text{ iff } \lambda_{i0} = \phi^i - \sum_{k \neq 0} \phi^k \overline{s}^k$$

We notice that λ_{i0} was still a free parameter so far. Clearly, condition (I) is then also satisfied if $\lambda_{i0} > 0$ for all i, which occurs if for all $i = 1, ..., n$ $\phi^i > \sum_{k=1}^n \phi^k \overline{s}^k$ for all $i = 1, ..., n$.

(D) Proofs

We assume throughout that $E\left[\int_0^\infty e^{-\rho\tau} Y_\tau d\tau\right] < \infty$. Define for convenience $\boldsymbol{\sigma}_c = (\sigma_c, 0, ..., 0)$. From (4), an application of Ito's Lemma yields the process for $S_t = 1/Y_t$:

$$dS_t/S_t = \left(k\left(1 - \overline{Y} S_t\right) + (1 - \lambda S_t)^2 \alpha^2 \sigma_c^2\right) dt + (1 - \lambda S_t) \alpha \boldsymbol{\sigma}_c d\mathbf{B}_t'. \tag{41}$$

The pricing kernel is given by

$$m_t = u_c(C_t, X_t, t) = e^{-\phi t} Y_t / C_t \tag{42}$$

and it follows the process $dm_t/m_t = -r_t dt + \boldsymbol{\sigma}_m d\mathbf{B}_t'$ where

$$r_t = \phi + \mu_c - \sigma_c^2 + k\left(1 - \overline{Y} S_t\right) - \alpha (1 - \lambda S_t) \sigma_c^2, \tag{43}$$

$$\boldsymbol{\sigma}_m = -(1 + \alpha (1 - S_t \lambda)) \boldsymbol{\sigma}_c. \tag{44}$$

Proof of Proposition 1 and 2: Proposition 1 is a special case of the result in Proposition 2, for $s_t^i = 1$ for all t. (a) From (15) with $s_t^g = s_t^i$, and the fact that $s_{t+\tau}^i Y_{t+\tau} < Y_{t+\tau}$ almost surely, we can use Fubini Theorem, and invert the order of integration, $P_t^i = C_t/Y_t \int_t^\infty E_t\left[s_{t+\tau}^i Y_{t+\tau}\right] d\tau$. The assumption $E_t\left[ds_t^i dc_t\right] = 0$ implies that $E_t\left[ds_t^i dY_t\right] = 0$, and thus[28]

$$E_t\left[s_{t+\tau}^i Y_{t+\tau}\right] = E_t\left[s_{t+\tau}^i\right] E_t\left[Y_{t+\tau}\right] = \left(\overline{s}^i + \left(s_t^i - \overline{s}^i\right) e^{-\phi^i \tau}\right)\left(\overline{Y} + \left(Y_t - \overline{Y}\right) e^{-k\tau}\right)$$

Substituting this expression in the pricing formula, tedious calculus derivations yield equation (20) with

$$a_0^i = \frac{1}{\rho + k + \phi^i},\; a_1^i = a_0^i \frac{k\overline{Y}}{\rho + \phi^i},\; a_2^i = a_0^i \frac{\phi^i}{\rho + k},\;\; a_3^i = a_0^i \frac{\phi^i k\overline{Y}}{\rho}\left(\frac{1}{\rho + k} + \frac{1}{\rho + \phi^i}\right) \quad (45)$$

Part (b): An application of Ito's lemma to (20), together with (41) yields $dP_t^i/P_t^i = \mu_{P,t} dt + \boldsymbol{\sigma}_{P,t}^i d\mathbf{B}_t$ where

$$\boldsymbol{\sigma}_{P,t}^i = \boldsymbol{\sigma}_c + \frac{\left(\left(a_0 + a_1 S_t\right) s_t^i \boldsymbol{\sigma}_i\left(\mathbf{s}\right) + \left(a_3 \overline{s}^i + a_1^i s_t^i\right) S_t \left(1 - \lambda S_t\right) \alpha \boldsymbol{\sigma}_c\right)}{\left(s_t^i a_0^i + a_1^i S_t s_t^i + a_2^i \overline{s}^i + a_3^i \overline{s}^i S_t\right)} \quad (46)$$

The excess return process $dR_t^i = \left(dP_t^i + D_t^i dt\right)/P_t^i - r_t dt$ has the same diffusion as (46), $\boldsymbol{\sigma}_{R,t}^i = \boldsymbol{\sigma}_{P,t}^i$. Equilibrium requires that $E\left[dR_t^i\right] = -\boldsymbol{\sigma}_{R,t}^i \boldsymbol{\sigma}_m'$, where the stochastic discount factor is in (42). Since $E\left[ds_t^i dc_t\right] = 0$ for all i, we have $\boldsymbol{\sigma}_i\left(\mathbf{s}\right)\boldsymbol{\sigma}_m' = 0$, which implies

$$E\left[dR_t^i\right] = \left(1 + \alpha\left(1 - S_t \lambda\right)\right)\left(1 + \frac{\left(a_3 \overline{s}^i + a_1^i s_t^i\right) S_t \left(1 - \lambda S_t\right) \alpha}{s_t^i a_0^i + a_1^i S_t s_t^i + a_2^i \overline{s}^i + a_3^i \overline{s}^i S_t}\right)\sigma_c^2 \quad (47)$$

Finally, substituting for a_j^i's from in (45), a little algebra shows that

$$\frac{a_3 \overline{s}^i + a_1^i s_t^i}{s_t^i a_0^i + a_2^i \overline{s}^i + \left(a_1^i s_t^i + a_3^i \overline{s}^i\right) S_t} = \frac{k\overline{Y}}{\rho\left(f\left(\frac{\overline{s}^i}{s_t^i}\right) + 1\right) + k\overline{Y} S_t}$$

[28]It can be checked that $E_t\left[Y_\tau\right] = \overline{Y} + \left(Y_t - \overline{Y}\right) e^{-k(\tau - t)}$ (see Karatzas and Shreve (1991), equation 6.34, page 361.)

where

$$f\left(\frac{\overline{s}^i}{s^i_t}\right)=\frac{\left(1+\frac{\phi^i}{\rho+k}\frac{\overline{s}^i}{s^i_t}\right)}{\left(\phi^i\left(\frac{1}{\rho+k}+\frac{1}{\rho+\phi^i}\right)\frac{\overline{s}^i}{s^i_t}+\frac{\rho}{\rho+\phi^i}\right)}-1 \tag{48}$$

In addition, it is immediate to see that $f(1)=0$ and $f'(x)<0$ for all x.■

Proof of Proposition 3: Part (a). From (47) and (20) we have

$$E\left[dR^i_t\right]=\left(1+\alpha\left(1-S_t\lambda\right)\right)\left(1+\frac{C_t}{P^i_t}\left(a_3\overline{s}^i+a^i_1 s^i_t\right)S_t\left(1-\lambda S_t\right)\alpha\right)\sigma^2_c$$

Using the definition $D^i_t=s^i_t C_t$, and defining $\sigma_S(S_t)=S_t(1-\lambda S_t)\alpha$, equation (22) follows with

$$b_0(S_t)=(1+\alpha(1-S_t\lambda))\sigma^2_c;\ b_1(S_t)=b_0(S)\,a^1_i\sigma_S(S_t);\ b_2(S_t)=b_0(S_t)\,\sigma_S(S_t)\,a_3\overline{s}^i \tag{49}$$

Part (b) follows immediately from collecting $\overline{s}^i/s^i_t$ in the price dividend ratio formula (20) and substituting it into (12), where we obtain in addition to $m_1(S_t)$ as in (24), the intercept

$$m^i_0(S_t,\mathbf{s}_t)=\mu_c-\phi^i-\phi^i\frac{(a^i_0+a^i_1S_t)}{(a^i_2+a^i_3S_t)}-\frac{1}{2}\boldsymbol{\sigma}^i(\mathbf{s}_t)\,\boldsymbol{\sigma}^i(\mathbf{s}_t)'. \tag{50}$$

■

Proof of Proposition 4: Part (a). As in the proof of Proposition 2, from $P^i_t=C_t\int_t^\infty e^{-\rho t}E_t\left[s^i_{t+\tau}\right]d\tau$ and $E_t\left[s^i_{t+\tau}\right]=\overline{s}^i+\left(s^i_t-\overline{s}^i\right)e^{-\phi^i\tau}$, the result (25) follows. Part (b). From Ito's Lemma applied to (25) we have $dP^i_t/P^i_t=\boldsymbol{\sigma}^i_{P,t}d\mathbf{B}_t+o(dt)$ where $o(dt)$ collects all the "dt" terms, and where $\boldsymbol{\sigma}^i_{P,t}=\boldsymbol{\sigma}_c+\left(1+\frac{\phi^i}{\rho}\frac{\overline{s}^i}{s^i_t}\right)^{-1}\boldsymbol{\sigma}(s_t)$. As in part (b) of Proposition 2, from $E\left[dR^i_t\right]=-\boldsymbol{\sigma}^i_{R,t}\boldsymbol{\sigma}'_m$, the fact that $\boldsymbol{\sigma}^i_{R,t}=\boldsymbol{\sigma}^i_{P,t}$ and that with no habit $\boldsymbol{\sigma}_m=\boldsymbol{\sigma}_c$, the result follows from $\boldsymbol{\sigma}(s_t)\,\boldsymbol{\sigma}'_c=\theta^i_{CF}-\sum_{j=0}^n\theta^j_{CF}s^j_t$. Part (c) follows trivially from (a) and (b) upon substitution. ■

Sketch of the derivation of general approximate pricing formulas

Let $q^i_t=s^i_tY_t$. As in the proof of Proposition 2, we can write $P^i_t=C_t/Y_t\int_0^\infty e^{-\rho\tau}E_t\left[q^i_{t+\tau}\right]d\tau$. Using Ito's lemma

$$dq^i_t=\left\{\phi^i\left(Y_t\overline{s}^i-q^i_t\right)+k\left(s^i_t\overline{Y}-q^i_t\right)-\alpha\left(Y_t-\lambda\right)s^i_t\theta^i_{CF}+\left(Y_t-\lambda\right)s^i_t\sum_{j=1}^n s^j_t\theta^j_{CF}\right\}dt$$

$$+\left\{q_t^i\left(\mathbf{v}_i-\sum_{j=1}^n s_t^j\mathbf{v}_j\right)-s_t^i\left(Y_t-\lambda\right)\alpha\boldsymbol{\sigma}_c\right\}d\mathbf{B}_t'$$

We now use (10) and approximate the term $\sum_{j=1}^n s_t^j\theta_{CF}^j$ by $E\left[\sum_{j=1}^n s_t^j\theta_{CF}^j\right]=\sum_{j=1}^n\overline{s}^j\theta_{CF}^j=0$. Thus, the pricing is based on the approximating process:

$$\begin{aligned}dq_t^i &= \left\{\phi^i\overline{s}^iY_t+\left(k\overline{Y}+\alpha\lambda\theta_{CF}^i\right)s_t^i-\left(\phi^i+k+\alpha\theta_{CF}^i\right)q_t^i\right\}dt\\&\quad+\left\{q_t^i\left(\mathbf{v}_i-\sum_{j=1}^n s_t^j\mathbf{v}_j\right)-s_t^i\left(Y_t-\lambda\right)\alpha\boldsymbol{\sigma}_c\right\}d\mathbf{B}_t'\end{aligned}$$

Define $\mathbf{Z}_t=\left(Y_t,q_t^i,s_t^i\right)'$. This evolves according to the system $d\mathbf{Z}_t=\left(\mathbf{A}_0+\mathbf{A}_1\mathbf{Z}_t\right)dt+\boldsymbol{\Sigma}\left(Y_t,\mathbf{s}_t\right)d\mathbf{B}_t'$ where $\mathbf{A}_0=\left(k\overline{Y},0,\phi^i\overline{s}^i\right)'$,

$$\mathbf{A}_1=\begin{pmatrix}-k & 0 & 0\\ \phi^i\overline{s}^i & -\left(\phi^i+k+\alpha\theta_{CF}^i\right) & k\overline{Y}+\lambda\alpha\theta_{CF}^i\\ 0 & 0 & -\phi^i\end{pmatrix}$$

and $\boldsymbol{\Sigma}\left(Y_t,\mathbf{s}_t\right)$ is a $3\times n$ appropriate matrix. We assume throughout that $\left(\phi^i+k+\alpha\theta_{CF}^i\right)>0$. Since $\mathbf{A}_1$ admits 3 distinct real eigenvalues, we have $E_t\left[\mathbf{Z}_{t+\tau}\right]=\boldsymbol{\Psi}\left(\tau\right)\mathbf{Z}_t+\int_t^\tau\boldsymbol{\Psi}\left(\tau-s\right)\mathbf{A}_0ds$, with $\boldsymbol{\Psi}\left(\tau\right)=\mathbf{U}\exp\left(\boldsymbol{\Lambda}\cdot\tau\right)\mathbf{U}^{-1}$, where $\exp\left(\boldsymbol{\Lambda}\cdot\tau\right)$ is the diagonal matrix with $\left[\exp\left(\boldsymbol{\Lambda}\cdot\tau\right)\right]_{ii}=e^{\lambda_i\tau}$, where λ_i's are the eigenvalues of $\mathbf{A}_1$, and $\mathbf{U}$ is the matrix of associated eigenvectors. Direct computation shows that

$$\boldsymbol{\Psi}\left(\tau\right)=\begin{pmatrix}e^{-k\tau} & 0 & 0\\ \frac{\phi^i\overline{s}_i}{\phi^i+\alpha\theta_{CF}^i}\left(e^{-k\tau}-e^{-\left(k+\phi^i+\alpha\theta_{CF}^i\right)\tau}\right) & e^{-\left(k+\phi^i+\alpha\theta_{CF}^i\right)\tau} & \frac{\left(k\overline{Y}+\lambda\alpha\theta_{CF}^i\right)}{k+\alpha\theta_{CF}^i}\left(e^{-\phi^i\tau}-e^{-\left(k+\phi^i+\alpha\theta_{CF}^i\right)\tau}\right)\\ 0 & 0 & e^{-\phi^i\tau}\end{pmatrix}$$

Therefore, $E_t\left[q_{t+\tau}^i\right]=e_2E_t\left[\mathbf{Z}_{t+\tau}\right]=e_2\boldsymbol{\Psi}\left(\tau\right)\mathbf{Z}_t+\int_0^\tau e_2\boldsymbol{\Psi}\left(\tau-s\right)\mathbf{A}_0ds$, where $e_2=(0,1,0)$. At this point, it is just a matter of tedious algebra to show that the two terms in

$$E_t\left[\int_t^\infty e^{-\rho(\tau-t)}q_\tau^i d\tau\right]=\int_0^\infty e^{-\rho\tau}e_2\boldsymbol{\Psi}\left(\tau\right)\mathbf{Z}_td\tau+\int_0^\infty e^{-\rho\tau}\int_0^\tau e_2\boldsymbol{\Psi}\left(\tau-s\right)\mathbf{A}_0dsd\tau \quad (51)$$

are given by

$$\int_0^\infty e^{-\rho\tau}\int_0^\tau e_2\mathbf{\Psi}(\tau-s)\,\mathbf{A}_0 ds d\tau = \widehat{a}_0^i\phi^i\overline{s}^i\left\{\frac{k\overline{Y}}{\rho(\rho+k)}+\frac{\left(k\overline{Y}+\lambda\,\alpha\theta_{CF}^i\right)}{\rho\left(\rho+\phi^i\right)}\right\}$$

$$\int_0^\infty e^{-\rho\tau}e_2\mathbf{\Psi}(\tau)\,\mathbf{Z}_t d\tau = \widehat{a}_0^i\left\{\frac{\phi^i\,\overline{s}_i}{(\rho+k)}Y_t+q_t^i+\frac{\left(k\overline{Y}+\lambda\,\alpha\theta_{CF}^i\right)}{(\rho+\kappa)}s_t^i\right\}$$

where $\widehat{a}_0^i=\left(\rho+k+\phi^i+\alpha\theta_{CF}^i\right)^{-1}$. Substituting into (51) we find (29) with

$$\widehat{a}_1^i=\widehat{a}_0^i\frac{k\overline{Y}+\lambda\,\alpha\theta_{CF}^i}{\left(\rho+\phi^i\right)};\ \widehat{a}_2^i=\widehat{a}_0^i\frac{\phi^i}{\rho+k};\ \widehat{a}_3^i=\widehat{a}_0^i\frac{\phi^i}{\rho}\left(\frac{k\overline{Y}}{\rho+k}+\frac{k\overline{Y}+\lambda\alpha\,\theta_{CF}^i}{\rho+\phi^i}\right) \tag{52}$$

We finally compute the formula for expected returns. From (29) the diffusion of the capital gain return process dP_t^i/P_t^i is as in equation (46), with the a_j^i's substituted with $\widehat{a}_j^i$'s. The same argument implies that expected returns must be given by $E\left[dR_t^i\right]=-\boldsymbol{\sigma}_{R,t}^i\boldsymbol{\sigma}_m'$. Computing this product explicitly, using the fact that $P_t^i/C_t=s_t^i\widehat{a}_0^i+\widehat{a}_1^iS_ts_t^i+\widehat{a}_2^i\overline{s}^i+\widehat{a}_3^i\overline{s}^iS_t$, and the definition $D_t^i=s_t^iC_t$, we then find

$$E\left[dR_t^i\right]=(1+\alpha(1-S_t\lambda))\left(\sigma^2+\frac{D_t^i}{P_t^i}\left(\widehat{a}_0^i+\widehat{a}_1^iS_t\right)\boldsymbol{\sigma}_i(\mathbf{s})\,\boldsymbol{\sigma}_c'+\left(\widehat{a}_3^i\overline{s}^i\frac{C_t}{P_t^i}+\widehat{a}_1^i\frac{D_t^i}{P_t^i}\right)S_t(1-\lambda S_t)\,\alpha\sigma^2\right)$$

Finally, we use the approximation $\sum_{j=1}^n s_t^j\theta_{CF}^j\approx 0$ to approximate $\boldsymbol{\sigma}_i(\mathbf{s})\,\boldsymbol{\sigma}_c=\theta_{CF}^i-\sum_{j=1}^n s_t^j\theta_{CF}^j\approx\theta_{CF}^i$ to obtain equation (30) with $\widehat{b}_0(S_t)=b_0(S_t)$ in (49), and

$$\widehat{b}_1(S_t)=\widehat{b}_0(S_t)\left(\left(\widehat{a}_0^i+\widehat{a}_1^iS_t\right)\frac{\theta_{CF}^i}{\sigma_c^2}+\widehat{a}_1^i\sigma_S(S)\right);\ \widehat{b}_2(S_t)=\widehat{b}_0(S_t)\,\widehat{a}_3^i\overline{s}^i\sigma_S(S)\,, \tag{53}$$

where recall that $\sigma_S(S)=S_t(1-\lambda_tS_t)\,\alpha$. The derivation regarding the predictability of dividend growth is the same as in Proposition 3.■

References

Abel, Andrew B. "Asset Prices Under Habit Formation and Catching Up with the Jones." American Economic Review 80 (1990): 38-42.

Abel, Andrew B. "Risk Premia and Term Premia in General Equilibrium." Journal of Monetary Economics 43 (1999): 3-33.

Ang, Andrew "Characterizing the ability of Dividend Yields to Predict Future Dividends in Log-Linear Present Value Models." Manuscript. Columbia University, Graduate School of Business, 2002.

Ang, Andrew and Geert Bekaert "Stock Return Predictability: Is it There?" Manuscript. Columbia University, Graduate School of Business, 2001.

Ang, Andrew and Jun Liu "How to Discount Cashflows with Time Varying Expected Returns." Manuscript. Columbia University, Graduate School of Business, 2003.

Bakshi, Gurdip S. and Zhiwu Chen "An Alternative valuation model for contingent claims," Journal of Financial Economics 44 (1997): 123-165.

Bansal, Ravi, Robert F. Dittmar, and Christian T. Lundblad "Consumption, Dividends, and the Cross Section of Stock Returns," Manuscript. Duke University, 2002.

Bekaert, Geert and Stephen Grenadier "Stock and Bond Pricing in an Affine Economy." Manuscript. Columbia University, Graduate School of Business, 2001.

Bossaerts, Peter and Richard C. Green "A General Equilibrium Model of Changing Risk Premia: Theory and Tests," Review of Financial Studies 2 (1989): 467-493.

Brock, William A. "Asset Prices in a Production Economy," in John J. McCall ed.: The Economics of Information and Uncertainty. The University of Chicago Press, Chicago, IL, 1982.

Campbell, John Y. "Asset Pricing at the Millenium," Journal of Finance 55 no. 4 (2000): 1515-1568.

Campbell, John Y. and John H. Cochrane "By Force of Habit: A Consumption Based Explanation of Aggregate Stock Market Behavior." Journal of Political Economy

107 (1999): 205-251.

Campbell, John Y., Andrew W. Lo, and A. Craig MacKinlay. The Econometrics of Financial Markets. Princeton University Press, Princeton, New Jersey, 1997.

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

Chapman, David A. "Habit Formation and Aggregate Consumption." Econometrica 66 (September, 1998): 1223-1230.

Cochrane, John H. "Where is the market going? Uncertain facts and novel theories," Economic Perspectives, Federal Reserve Bank of Chicago, (November-December 1997).

Cochrane, John H. Asset Pricing, Princeton University Press, Princeton, NJ, 2000.

Constantinides, George "Habit Formation: A Resolution of the Equity Premium Puzzle," Journal of Political Economy 98 (1990): 519-543.

Daniel, Kent and David Marshall "The Equity Premium Puzzle and the Risk-Free Rate Puzzle at Long Horizons." Macroeconomic Dynamics, 1 (1997): 452-484.

Detemple, Jerome B. and Fernando Zapatero "Asset Prices in an Exchange Economy with Habit Formation." Econometrica 59 (1991): 1633-1657.

Fama, Eugene F. and Kenneth R. French "Permanent and Temporary Components of Stock Prices." Journal of Political Economy 96 (1988): 246-273.

Ferson, Wayne and George Constantinides "Habit Persistence and Durability in Aggregate Consumption: Empirical Tests." Journal of Financial Economics 29 (1991): 199-240.

Grinblatt, Mark and Toby Moskowitz "Do Industries Explain Momentum." Journal of Finance 54 (1999): 1249-1290.

Grullon, Gustavo, and Roni Michaely "Dividends, Share Repurchases, and the Substitution Hypothesis." Journal of Finance 57 (2002): 1649-1684.

Hansen, Lars P., John C. Heaton, and Nan Li "Intangible Risk?" Manuscript.

University of Chicago, 2002.

Hodrick, Robert “Dividend Yields and Expected Stock Returns: Alternative Procedures for Inference and Measurement.” Review of Financial Studies 5 (1992): 357-386.

Horvath Michael T.K. and Mark W. Watson “Testing for Cointegration when Some of the Cointegrating Vectors are Known.” Econometric Theory 11 (1995): 952-984.

Jagannathan, Murray, Clifford P. Stephens, and Michael S. Weisbach “Financial Flexibility and the Choice Between Dividends and Stock Repurchases.” Journal of Financial Economics 57 (2000): 355-384.

Karatzas, Ioannis and Steven E. Shreve. Brownian Motion and Stochastic Calculus, 2nd ed., Springer-Verlag, New York, 1991.

Lamont, Owen “Earnings and Expected Returns.” Journal of Finance 53 (October 1998): 1563-1587.

Lettau, Martin and Sydney Ludvigson “Consumption, Aggregate Wealth, and Expected Stock Returns.” Journal of Finance 56 (June 2001): 815-849.

Lettau, Martin and Sydney Ludvigson “Expected Stock Returns and Expected Dividend Growth,” Manuscript. New York University, Stern School of Business, 2003.

Li, Yuming “Expected Returns and Habit Persistence,” Review of Financial Studies 14 (2001): 861-899.

Liptser S. Robert, and Albert N. Shiryaev. Statistics of Random Processes I, Springer-Verlag, New York, 1977.

Lucas, Robert E. “Asset Prices in an Exchange Economy.” Econometrica 46 (1978): 1429-1445.

Ribeiro, Ruy “Predictable Dividends and Returns,” Manuscript. university of Chicago, Graduate School of Business, University of Chicago, 2003.

Richardson, Mathew and James Stock “Drawing Inferences from Statistics Based on Multi-Year Asset Returns.” Journal of Financial Economics 25 (1989): 323-348.

Rubinstein, Mark “The Valuation of Uncertain Income Streams and the Price of Options.” Bell Journal of Economics 7 (1976): 407-425.

Santos, Tano, and Pietro Veronesi "Labor Income and Predictable Stock Returns," Manuscript. University of Chicago, Graduate School of Business, 2001.

Stambaugh, Robert "Predictive Regressions." Journal of Financial Economics 54 (1999): 375-421.

Sundaresan, Suresh "Intertemporal Dependent Preferences and the Volatility of Consumption and Wealth." Review of Financial Studies 2 (1989): 73-88.

Vuolteenaho, Tuomo "What Drives Firm-Level Stock Returns?" Journal of Finance 57 (2002): 233-264.

Wachter, Jessica "Habit Formation and Returns on Bonds and Stocks." Manuscript. New York University, Stern School of Business, 2001.

Table 1
Model parameters and moments of aggregate quantities

Panel A: Preference parameters and consumption parameters

ρ	$\bar{Y}$	k	λ	α	μ_C	σ_C
0.04	33.97	0.16	20.00	79.39	0.02	0.01

Panel B: Aggregate Moments

	$E(R)$	$Vol(R)$	$E(r_f)$	$Vol(r_f)$	Ave(PC/100)	SR
Data	0.07	0.16	0.01	0.01	0.30	0.46
Model	0.07	0.23	0.01	0.04	0.30	0.31

Panel C: Share Process

Industry	$\overline{s}^i$	ϕ^i	θ (x1000)	Cov($dc, d\delta$) (x1000)	HW Test LR
Constr.	0.04	0.52	-0.12	0.05	14.26 *
Railroads	0.09	0.20	-0.47	-0.30	12.91 **
Retail	0.04	0.20	-0.09	0.08	22.95 *
Petroleum	0.52	0.16	-0.20	-0.02	12.29 **
Mining	0.05	0.16	-0.33	-0.15	3.93
Elect.Eq.	0.09	0.14	-0.21	-0.03	10.21 **
Apparel	0.01	0.12	-0.16	0.01	6.18
Machinery	0.12	0.11	-0.10	0.08	19.92 *
Paper	0.05	0.11	-0.19	-0.02	14.13 *
Other Transp.	0.01	0.09	-0.06	0.11	11.03 **
Dept.Stores	0.09	0.09	-0.03	0.15	14.83 *
Transp.Eq.	0.25	0.08	0.27	0.44	7.91
Manufact.	0.05	0.06	-0.13	0.05	3.63
Other	0.17	0.06	-0.08	0.09	11.85 **
Fab.Metals	0.03	0.05	-0.17	0.00	21.13 *
Financial	0.05	0.04	-0.02	0.15	27.30 *
Chemical	0.29	0.03	-0.14	0.03	9.33 ***
Prim.Metals	0.12	0.01	-0.32	-0.14	1.19
Utilities	0.10	0.00	-0.06	0.11	4.85
Food	0.15	0.00	-0.09	0.09	5.30
Mkt.Ptfl.	2.22	0.07	-0.10	0.07	16.70 *

Notes to Table 1: *Panel A*: Annualized preference and consumption process parameters chosen to calibrate the mean average excess returns, the average price consumption ratio, the average risk free rate and its volatility, and the Sharpe ratio of the market portfolio. *Panel B:* Expected excess return of the market portfolio, $E(R)$, standard deviation of returns of the market portfolio, $Vol(R)$, expected risk free rate, $E(r_f)$, standard deviation of the risk free rate, $Vol(r_f)$, average price consumption ratio, Ave(PC/100), and Sharpe ratio of the market portfolio, SR. *Panel C:* Estimates of the long run mean, $\overline{s}^i$, and the speed of mean reversion ϕ^i, cash flow risk, θ^i_{CF}, and covariance between dividend growth and consumption growth, $cov(d\delta^i_t, dc_t)$ for each industry. Industries are ordered, in this and subsequent tables, according to the parameter ϕ^i. The last column describes the likelihood ratio statistic for testing the null of no cointegration versus the alternative of cointegration with the prespecified cointegrating vector as described in Horvath and Watson (1995). Entries marked with *, **, and *** denotes those industries for which the null of no cointegration can be rejected at the 1%, 5 % and 10 %, respectively. All entries in the table are in annual units.

TABLE 2
Price dividend ratios

Panel A: All Industries

	Surplus from consumption					Surplus from P^{mkt}/D^{mkt}				
	α_0	α_1	se(α_0)	se(α_1)	R^2	α_0	α_1	se(α_0)	se(α_1)	R^2
	1.43	0.55	0.30	0.09	0.20	1.45	0.55	0.30	0.09	0.20

Panel B: Individual Industries

	Surplus from consumption					Surplus from P^{mkt}/D^{mkt}				
	α_0	α_1	se(α_0)	se(α_1)	R^2	α_0	α_1	se(α_0)	se(α_1)	R^2
Construction	1.88	0.38	0.66	0.18	0.17	1.28	0.57	0.43	0.12	0.49
Railroads	1.18	0.54	0.35	0.11	0.42	0.50	0.75	0.46	0.14	0.59
Retail	2.27	0.36	0.61	0.17	0.11	0.80	0.82	0.39	0.12	0.55
Petroleum	1.55	0.44	0.46	0.13	0.21	0.20	0.86	0.29	0.09	0.76
Mining	1.39	0.62	0.83	0.22	0.33	1.31	0.65	0.48	0.13	0.57
Electric Eq.	2.53	0.28	0.52	0.15	0.04	-0.78	1.31	0.41	0.13	0.67
Apparel	1.36	0.52	0.35	0.11	0.32	1.27	0.55	0.30	0.10	0.49
Machinery	1.63	0.57	0.41	0.13	0.16	0.73	0.85	0.41	0.13	0.53
Paper	2.19	0.33	0.39	0.12	0.11	0.19	0.96	0.48	0.15	0.60
Other Transport.	0.47	0.90	0.51	0.15	0.48	0.27	0.98	0.27	0.08	0.68
Dept. Stores	1.37	0.63	0.64	0.19	0.19	-0.30	1.15	0.31	0.09	0.73
Transport Eq.	1.65	0.38	0.62	0.18	0.14	0.85	0.64	0.26	0.07	0.58
Manufacturing	2.68	0.31	0.65	0.21	0.04	-0.00	1.13	0.55	0.18	0.47
Other	2.11	0.35	0.58	0.18	0.10	0.16	0.95	0.44	0.14	0.69
Fab. Metals	1.79	0.44	0.56	0.17	0.13	-0.17	1.04	0.32	0.10	0.79
Financial	2.28	0.29	0.41	0.12	0.09	0.05	0.99	0.20	0.06	0.84
Chemical	2.47	0.27	0.39	0.13	0.08	-0.03	1.04	0.24	0.07	0.81
Prim. Metals	1.28	0.52	0.41	0.12	0.28	0.63	0.72	0.39	0.12	0.61
Utilities	1.93	0.27	0.58	0.18	0.06	0.88	0.60	0.73	0.23	0.26
Food	2.44	0.24	0.61	0.18	0.06	0.45	0.86	0.41	0.13	0.65

Notes to Table 2: Panel A: Parameter estimates of the cross sectional regression

$$\ln\left(\frac{P_t^i}{D_t^i}\right) = \alpha_0 + \alpha_1 \ln\left(\widehat{PD}_t^i\right) + \varepsilon_t^i,$$

where $\widehat{PD}_t^i$ is the price dividend ratio implied by the model. That is, given the parameter values for the cash flow processes, preference and consumption contained in Table 1, we feed our pricing formula, equation (29), the observed consumption and dividend realizations to obtain $\widehat{PD}_t^i$. The results are reported both for the case where the surplus consumption ratio is obtained from consumption shocks and also for the case where the total wealth portfolio is approximated by the market portfolio and expression (17) is used to obtain a time series of S_t. Standard errors are denoted by $se(\alpha_0)$ and $se(\alpha_1)$ for the intercept and the slope coefficients, respectively. We use the Fama-MacBeth methodology, and standard errors are adjusted for heteroskedasticity and autocorrelation. Adjusted R^2 are also reported. Panel B: Parameter estimates of analogous time series regressions, industry by industry. Standard errors are adjusted for heteroskedasticity and autocorrelation. The sample period is 1947-2001 and the data is quarterly.

TABLE 3
The predictability of dividend growth

	Relative Share			PD Ratio			Multivariate Regression				
Panel A: Market Portfolio											
Horizon	$\frac{\bar{s}^i}{s^i_t}$	se($\frac{\bar{s}^i}{s^i_t}$)	R^2	$\frac{P^i_t}{D^i_t}$	se($\frac{P^i_t}{D^i_t}$)	R^2	$\frac{\bar{s}^i}{s^i_t}$	$\frac{P^i_t}{D^i_t}$	se($\frac{\bar{s}^i}{s^i_t}$)	se($\frac{P^i_t}{D^i_t}$)	R^2
1 year	0.11 *	0.05	0.05	-0.43	1.57	0.00	0.11 *	-0.09	0.05	1.57	0.04
4 year	0.68 *	0.26	0.31	-0.75	7.51	0.00	0.68 *	0.62	0.27	6.22	0.30
7 year	0.92 *	0.26	0.41	-6.83	8.28	0.05	0.89 *	-4.10	0.32	6.59	0.41
Panel B: Pooled Regressions											
Horizon	$\frac{\bar{s}^i}{s^i_t}$	se($\frac{\bar{s}^i}{s^i_t}$)	R^2	$\frac{P^i_t}{D^i_t}$	se($\frac{P^i_t}{D^i_t}$)	R^2	$\frac{\bar{s}^i}{s^i_t}$	$\frac{P^i_t}{D^i_t}$	se($\frac{\bar{s}^i}{s^i_t}$)	se($\frac{P^i_t}{D^i_t}$)	R^2
1 year	0.09 *	0.02	0.06	3.69 *	0.76	0.04	0.08 *	2.82 *	0.02	0.63	0.08
4 year	0.28 *	0.06	0.20	8.38 *	2.50	0.10	0.25 *	5.43 *	0.06	1.91	0.23
7 year	0.35 *	0.09	0.24	5.83 *	2.72	0.10	0.34 *	3.12	0.09	2.07	0.24
Panel C: Individual Industries (4 year horizon)											
Industry	$\frac{\bar{s}^i}{s^i_t}$	se($\frac{\bar{s}^i}{s^i_t}$)	R^2	$\frac{P^i_t}{D^i_t}$	se($\frac{P^i_t}{D^i_t}$)	R^2	$\frac{\bar{s}^i}{s^i_t}$	$\frac{P^i_t}{D^i_t}$	se($\frac{\bar{s}^i}{s^i_t}$)	se($\frac{P^i_t}{D^i_t}$)	R^2
Constr.	0.68 *	0.29	0.35	15.80	15.63	0.07	0.65 *	10.65	0.27	8.44	0.36
Railroads	0.58 *	0.14	0.37	33.70 *	8.00	0.29	0.43 *	13.46 *	0.13	4.23	0.38
Retail	0.49 *	0.23	0.19	8.32	5.99	0.10	0.41 *	3.93	0.19	6.46	0.19
Petroleum	0.60 *	0.17	0.31	7.36	12.47	0.02	0.60 *	0.97	0.15	6.22	0.29
Mining	0.30	0.23	0.08	16.07 *	8.20	0.22	0.03	15.49	0.24	10.09	0.20
Elect.Eq.	0.67 *	0.14	0.29	1.99	5.98	0.01	0.71 *	3.80	0.15	5.37	0.30
Apparel	0.44 *	0.08	0.39	12.49	8.49	0.11	0.44 *	-0.63	0.09	4.00	0.37
Machinery	0.64 *	0.09	0.43	12.22	7.36	0.19	0.59 *	9.55 *	0.06	3.34	0.54
Paper	0.72 *	0.18	0.38	1.24	3.91	0.00	0.73 *	2.55	0.14	2.08	0.38
Other Transp.	0.66 *	0.14	0.40	12.05	10.08	0.13	0.69 *	-1.68	0.13	5.32	0.39
Dept.Stores	0.87 *	0.31	0.43	6.85	6.18	0.08	0.86 *	0.08	0.31	3.51	0.42
Transp.Eq.	0.43 *	0.09	0.34	38.71 *	16.22	0.27	0.34 *	26.78 *	0.09	13.04	0.44
Manufact.	0.11	0.13	0.03	0.06	1.60	0.00	0.11	0.21	0.14	1.60	0.01
Other	0.37 *	0.15	0.27	3.53	6.06	0.01	0.37 *	3.04	0.15	4.77	0.27
Fab.Metals	0.74 *	0.22	0.34	13.81	9.39	0.18	0.63 *	6.78	0.19	7.92	0.37
Financial	0.54 *	0.20	0.19	-5.13	6.14	0.03	0.55 *	-5.61	0.22	5.21	0.21
Chemical	0.12	0.15	0.02	-5.85	3.42	0.08	0.17	-6.71	0.15	3.51	0.10
Prim.Metals	0.11 *	0.04	0.17	32.61	16.79	0.26	0.04	26.60	0.04	20.01	0.26
Utilities	-0.09	0.07	0.07	8.26 *	3.47	0.16	-0.02	7.66 *	0.07	3.23	0.14
Food	-0.16	0.15	0.05	-2.11	9.01	0.01	-0.16	-1.60	0.16	8.92	0.03

Notes to Table 3: Dividend growth predictability regressions for market portfolio (Panel A), pooled individual industries (Panel B) and individual industries (Panel C). For each of the three cases, the predictive regression is

$$\Delta d_{t,t+k} = \alpha + \boldsymbol{\beta}\mathbf{X}_t + \varepsilon_{t+k}$$

where $\Delta d^i_{t,t+\tau}$ denotes the cumulative dividend growth between time t and $t+\tau$, $\mathbf{X}_t$ is either the relative share $\frac{\bar{s}}{s_t}$, the price dividend ratio $\frac{P_t}{D_t}$, or both (heading multivariate regression). The pooled regression in Panel B uses fixed effects. For each regression we report the corresponding Newey-West adjusted standard errors, where the number of lags is double the number of years in the forecasting horizon. A star "*" on the regression coefficient indicates statistical significance at the 5% level. The last column under each heading reports the adjusted R^2. The sample period is 1947-2001. Data is annual.

TABLE 4
The Predictability of dividend growth: Simulations

Panel A: Simulation Results

Industry	Horizon = 1 year $\bar{s}^i/s_t^i$	R^2	P_t^i/D_t^i	R^2	Horizon = 4 years $\bar{s}^i/s_t^i$	R^2	P_t^i/D_t^i	R^2
Mkt Ptfl	0.07	0.03	0.50	0.01	0.25	0.09	1.97	0.02
Constr.	0.38	0.21	9.66	0.15	0.86	0.30	22.11	0.22
Railroads	0.19	0.11	4.99	0.08	0.59	0.24	15.35	0.17
Retail	0.18	0.09	3.14	0.04	0.59	0.21	10.10	0.09
Petroleum	0.16	0.08	2.34	0.04	0.51	0.19	7.36	0.08
Mining	0.13	0.08	4.00	0.06	0.44	0.18	13.23	0.14
Elect.Eq.	0.14	0.07	1.85	0.03	0.46	0.17	6.28	0.07
Apparel	0.11	0.06	2.82	0.04	0.39	0.15	9.94	0.10
Machinery	0.11	0.06	2.37	0.03	0.38	0.14	8.54	0.08
Paper	0.10	0.05	1.58	0.02	0.36	0.13	5.57	0.06
Dept.Stores	0.09	0.05	2.81	0.04	0.32	0.13	10.03	0.09
Other Transp.	0.08	0.04	1.59	0.02	0.29	0.11	5.84	0.05
Transp.Eq.	0.07	0.04	2.24	0.02	0.26	0.10	8.04	0.06
Other	0.06	0.03	1.19	0.01	0.20	0.08	4.42	0.03
Manufact.	0.06	0.03	0.76	0.01	0.24	0.09	3.05	0.02
Fab.Metals	0.06	0.03	1.37	0.02	0.22	0.09	5.00	0.04
Financial	0.04	0.02	0.73	0.01	0.16	0.06	3.35	0.02
Chemical	0.03	0.02	0.70	0.01	0.11	0.05	3.15	0.02
Prim.Metals	0.01	0.01	0.88	0.01	0.03	0.02	3.35	0.02
Utilities	0.00	0.00	-0.06	0.00	0.00	0.00	-0.33	0.00
Food	0.00	0.00	0.20	0.00	0.00	0.01	0.92	0.00

Panel B: Data vs Model Predictions

	Horizon = 1 year $\bar{s}^i/s_t^i$	R^2	P_t^i/D_t^i	R^2	Horizon = 4 years $\bar{s}^i/s_t^i$	R^2	P_t^i/D_t^i	R^2
a_0	0.07	0.05	2.07	0.05	0.17	0.14	3.09	0.08
a_1	0.83	0.58	1.14	0.45	0.83	0.91	1.03	0.46
se(a_0)	0.03	0.02	1.23	0.02	0.10	0.04	4.20	0.03
se(a_1)	0.30	0.32	0.66	0.55	0.22	0.27	0.45	0.42
$R2$	0.39	0.20	0.24	0.07	0.40	0.25	0.19	0.07

Notes to Table 4: Panel A: Simulation results of the time series regressions $\Delta d_{t,t+k}^i = \alpha + \beta X_t^i + \varepsilon_{t+k}^i$ for $k = 1$, 4 years,for the market portfolio (line Mkt Ptfl) and each of the twenty industry portfolios. $\Delta d_{t,t+\tau}^i$ denotes the cumulative dividend growth between time t and $t + \tau$, and the regressor is either $X_t^i = \bar{s}^i/s_t^i$, the industry i's relative share, or $X_t^i = \frac{P_t^i}{D_t^i}$, the price dividend ratio. Simulations are based on 40,000 quarters of dividend and consumption data. Panel B: Results of the cross sectional regression of the coefficients of the dividend growth predictability regression on their population counterparts obtained in simulations, that is,

$$\hat{\beta}_X^{Data,i} = a_0 + a_1 \hat{\beta}_X^{Simulation,i} + \varepsilon,$$

where $\hat{\beta}_X^{Data,i}$ is the data coefficient associated with either the relative share or the price dividend ratio (from Table III for the four year regression) and $\hat{\beta}_X^{Simulation,i}$ is the population counterpart (from Panel A). The column denoted R^2 reports the corresponding cross sectional regression of the data R^2 on its population counterpart. For each cross sectional regression, we report the standard errors, $se(a_0)$ and $se(a_1)$, and the R^2 (R2 in Panel B).

TABLE 5
The predictability of returns: dividend yields and consumption to price ratio

Panel A: Market Portfolio

Horizon	$\frac{D_t^i}{P_t^i}$	se($\frac{D_t^i}{P_t^i}$)	R^2	$\frac{D_t^i}{P_t^i}$	$\frac{C_t}{P_t^i}$	se($\frac{D_t^i}{P_t^i}$)	se($\frac{C}{P_t^i}$)	R^2
1 year	4.78 *	1.46	0.12	4.78 *	-28.81	1.47	57.64	0.12
4 year	11.72 *	5.29	0.24	11.72 *	55.57	5.22	230.48	0.24
7 year	25.42 *	5.95	0.49	25.42 *	54.77	5.83	267.17	0.49

Panel B: Pooled Regression

Horizon	$\frac{D_t^i}{P_t^i}$	se($\frac{D_t^i}{P_t^i}$)	R^2	$\frac{D_t^i}{P_t^i}$	$\frac{C_t}{P_t^i}$	se($\frac{D_t^i}{P_t^i}$)	se($\frac{C}{P_t^i}$)	R^2
1 year	2.94 *	0.38	0.06	2.94 *	0.41 *	0.36	0.11	0.07
4 year	7.40 *	1.24	0.15	7.40 *	1.10 *	1.17	0.29	0.17
7 year	13.84 *	2.18	0.25	13.84 *	1.97 *	2.05	0.47	0.29

Panel C: Individual industries (4 year horizon)

Industry	$\frac{D_t^i}{P_t^i}$	se($\frac{D_t^i}{P_t^i}$)	R^2	$\frac{D_t^i}{P_t^i}$	$\frac{C_t}{P_t^i}$	se($\frac{D_t^i}{P_t^i}$)	se($\frac{C}{P_t^i}$)	R^2
Constr.	4.03	3.06	0.08	4.03	2.01	2.15	1.16	0.16
Railroads	4.06 *	1.61	0.11	4.06 *	3.50 *	1.10	1.39	0.18
Retail	11.37 *	4.85	0.29	11.37 *	0.35	4.86	1.48	0.29
Petroleum	3.89	2.20	0.08	3.89 *	58.56 *	0.92	7.23	0.48
Mining	2.89	3.39	0.03	2.89 *	9.67 *	1.19	1.99	0.37
Elect.Eq.	14.00	7.76	0.17	14.00	5.33	7.60	7.60	0.18
Apparel	9.72 *	3.47	0.25	9.72 *	0.53 *	2.44	0.25	0.31
Machinery	4.44	6.85	0.03	4.44	26.96 *	3.64	4.35	0.53
Paper	15.53 *	3.54	0.39	15.53 *	6.98 *	2.00	1.49	0.53
Other Transp.	10.31	6.36	0.16	10.31 *	2.13 *	4.10	1.00	0.27
Dept.Stores	9.86	5.51	0.17	9.86	2.53	5.28	3.23	0.17
Transp.Eq.	5.97	3.77	0.10	5.97	10.29	3.95	6.22	0.18
Manufact.	13.56 *	6.90	0.19	13.56 *	3.79 *	6.50	1.66	0.26
Other	7.18	5.62	0.08	7.18	3.18	5.57	5.67	0.08
Fab.Metals	12.46 *	3.50	0.33	12.46 *	0.87	3.43	1.14	0.33
Financial	14.63 *	5.51	0.25	14.63 *	0.61	5.48	2.37	0.25
Chemical	18.99 *	4.01	0.38	18.99 *	-17.63	4.34	21.13	0.41
Prim.Metals	9.35	5.69	0.19	9.35	0.36	5.68	1.11	0.19
Utilities	2.89	2.67	0.05	2.89	0.01	2.67	2.17	0.05
Food	13.78 *	4.83	0.31	13.78 *	-7.65	4.81	4.56	0.36

Notes to Table 5: Return predictability regressions for the market portfolio (Panel A), pooled individual industries (Panel B) and individual industries (Panel C). For each of the three cases, the predictive regression is

$$r_{t,t+k} = \alpha + \boldsymbol{\beta}\mathbf{X}_t + \varepsilon_{t+k}$$

where $r^i_{t,t+\tau}$ denotes the cumulative log excess return between time t and $t+\tau$, $\mathbf{X}_t = \frac{D_t}{P_t}$, or $\mathbf{X}_t = \left[\frac{D_t}{P_t}, \frac{C_t}{P_t}\right]$, where the consumption price ratio is orthogonalized with respect to the dividend yield. The pooled regression in Panel B uses fixed effects. For each regression we report the corresponding Newey-West adjusted standard errors, where the number of lags is double the number of years in the forecasting horizon. A star "*" on the regression coefficient indicates statistical significance at the 5% level. The last column under each heading reports the adjusted R^2. The sample period is 1947-2001. Data is quarterly.

TABLE 6
The Predictability of returns: Simulations

Panel A: Simulation Results

Industry	Horizon = 1 year: Univariate $\frac{D_t^i}{P_t^i}$	Univariate R^2	Multivariate $\frac{D_t^i}{P_t^i}$	Multivariate $\frac{C_t}{P_t^i}$ ($x1000$)	Multivariate R^2	Horizon = 4 year: Univariate $\frac{D_t^i}{P_t^i}$	Univariate R^2	Multivariate $\frac{D_t^i}{P_t^i}$	Multivariate $\frac{C_t}{P_t^i}$ ($x1000$)	Multivariate R^2
Mkt. Ptfl	4.00	0.06	4.00	39.99	0.06	10.37	0.15	10.37	98.91	0.15
Constr.	2.15	0.03	2.15	1.34	0.06	5.53	0.08	5.53	3.46	0.15
Railroads	2.17	0.03	2.17	3.12	0.06	5.40	0.07	5.40	8.49	0.15
Retail	3.36	0.05	3.36	1.44	0.06	8.61	0.12	8.61	3.87	0.15
Petroleum	3.59	0.06	3.59	14.57	0.06	9.32	0.14	9.32	38.75	0.15
Mining	2.29	0.03	2.29	1.48	0.06	5.76	0.08	5.76	3.93	0.14
Elect.Eq.	3.44	0.05	3.44	3.56	0.06	8.96	0.13	8.96	8.85	0.15
Apparel	2.96	0.04	2.96	0.32	0.06	7.62	0.10	7.62	0.83	0.14
Machinery	3.27	0.05	3.27	3.75	0.06	8.38	0.12	8.38	9.58	0.15
Paper	3.71	0.06	3.71	1.25	0.06	9.67	0.14	9.67	2.99	0.15
Other Transp.	2.55	0.03	2.55	0.41	0.06	6.50	0.08	6.50	1.08	0.14
Dept.Stores	3.42	0.05	3.42	2.67	0.06	8.87	0.12	8.87	6.76	0.15
Transp.Eq.	2.94	0.04	2.94	6.50	0.06	7.74	0.09	7.74	16.23	0.14
Manufact.	3.54	0.05	3.54	1.27	0.06	9.20	0.13	9.20	3.22	0.15
Other	3.76	0.06	3.76	4.65	0.06	9.62	0.13	9.62	12.80	0.15
Fab.Metals	3.58	0.05	3.58	0.69	0.06	9.19	0.12	9.19	1.84	0.15
Financial	3.74	0.05	3.74	0.99	0.06	9.60	0.13	9.60	2.49	0.15
Chemical	3.62	0.05	3.62	5.62	0.06	9.36	0.13	9.36	13.27	0.15
Prim.Metals	3.78	0.04	3.78	0.46	0.05	9.77	0.09	9.77	1.27	0.11
Utilities	4.17	0.06	4.17	0.00	0.06	10.88	0.15	10.88	0.00	0.15
Food	4.16	0.06	4.16	0.00	0.06	10.77	0.13	10.77	0.00	0.13

Panel B: Data versus Model Predictions

	Horizon = 1 year: Univariate $\frac{D_t^i}{P_t^i}$	Univariate R^2	Multivariate $\frac{D_t^i}{P_t^i}$	Multivariate $\frac{C_t}{P_t^i}$	Multivariate R^2	Horizon = 4 year: Univariate $\frac{D_t^i}{P_t^i}$	Univariate R^2	Multivariate $\frac{D_t^i}{P_t^i}$	Multivariate $\frac{C_t}{P_t^i}$	Multivariate R^2
a_0	-0.26	0.01	-0.26	-1.19	0.24	-2.00	0.00	-2.00	-3.26	0.17
a_1	1.15	1.48	1.15	0.89	-2.22	1.34	1.60	1.34	1.27	0.73
se(a_0)	1.45	0.05	1.45	0.65	0.09	4.38	0.10	4.38	1.85	0.32
se(a_1)	0.50	1.02	0.50	0.30	1.55	0.57	0.95	0.57	0.31	2.30
$R2$	0.20	0.11	0.20	0.48	0.04	0.20	0.12	0.20	0.57	0.00

Notes to Table 6: Panel A: Return predictability univariate regressions on D_t^i/P_t^i and multivariate regressions on D_t^i/P_t^i and C_t/P_t^i in simulated data for one and four year horizons. Simulations are based on 40,000 quarters of dividend and consumption data. Results are for both the market portfolio (line Mkt Ptfl) and the 20 industries. Panel B: Results of the cross sectional regression of the coefficients of the return predictability regression, reported in Table 5 for the 4 year regression, on their population counterparts obtained in simulations, that is,

$$\hat{\beta}_X^{Data,i} = a_0 + a_1 \hat{\beta}_X^{Simulation,i} + \varepsilon,$$

where $\hat{\beta}_X^{Data,i}$ is the data coefficient associated with either the dividend yield or the consumption to price ratio and $\hat{\beta}_X^{Simulation,i}$ is the population counterpart. The column denoted R^2 reports the corresponding cross sectional regression of the data R^2 on its population counterpart. We report the standard errors, $se(a_0)$ and $se(a_1)$ for each coefficient, and the R^2 (denoted by R2 in last line of Panel B).

TABLE 7
The source of return predictability

Panel A: Data

$\beta_{D/P}$ versus ϕ^i and θ^i

Horizon	α_0	α_1	α_2 (/100)	se(α_0)	se(α_1)	se(α_2) (/100)	R^2
1 year	3.87	-4.81	-15.36	0.53	1.74	17.55	0.13
4 years	11.38	-16.14	10.13	1.62	5.66	63.84	0.15
7 years	23.01	-35.05	55.40	3.02	10.58	132.86	0.21

$\beta_{C/P}$ versus ϕ^i and θ^i

Horizon	α_0	α_1	α_2 (/100)	se(α_0)	se(α_1)	se(α_2) (/100)	R^2
1 year	0.36	7.78	-0.22	0.95	7.12	36.28	0.04
4 years	2.76	24.04	-13.33	3.16	26.40	111.26	0.04
7 years	7.26	28.04	-65.60	3.84	36.56	159.60	0.03

Panel B: Simulations

$\beta_{D/P}$ versus ϕ^i and θ^i

Horizon	α_0	α_1	α_2 (/100)	se(α_0)	se(α_1)	se(α_2) (/100)	R^2
1 year	3.78	-3.62	5.19	0.17	0.66	9.67	0.51
4 years	9.82	-9.50	16.38	0.44	1.75	25.35	0.50
7 years	11.26	-10.92	18.85	0.51	2.09	28.89	0.49

$\beta_{C/P}$ versus ϕ^i and θ^i

Horizon	α_0	α_1	α_2 (/100)	se(α_0)	se(α_1)	se(α_2) (/100)	R^2
1 year	2.68	3.15	23.91	0.86	5.53	37.01	0.02
4 years	6.73	8.67	52.53	2.18	14.55	95.15	0.02
7 years	7.87	9.96	52.68	2.51	17.34	107.54	0.01

Notes to Table 7: Panel A: The top panel reports the result of the cross-sectional regression of the predictive regression coefficient $\beta^i_{D/P}$ (from Table 5) on the speed of mean reversion for the share, ϕ^i, and cash flow risk parameter, θ^i_{CF} (from Table 1). The bottom panel in Panel A reports the same quantities but for the predictive regression coefficient $\beta^i_{C/P}$. Here $\beta_{D/P}$ and $\beta_{C/P}$ denote the regression coefficients in the return predictability regression

$$r^i_{t,t+k} = \beta^i_0 + \beta^i_{D/P}\left(\frac{D^i_t}{P^i_t}\right) + \beta^i_{C/P}\left(\frac{C_t}{P^i_t}\right) + \varepsilon^i_{t,t+\tau} \qquad \text{for} \qquad k = 1, 4, 7 \text{ years},$$

where the data is quarterly and the sample period is 1947-2001. Results are for the one, four, and seven year horizon. Panel B: Same as Panel A, but for simulated data, where these consist of 40,000 quarters of simulated dividend and consumption data.

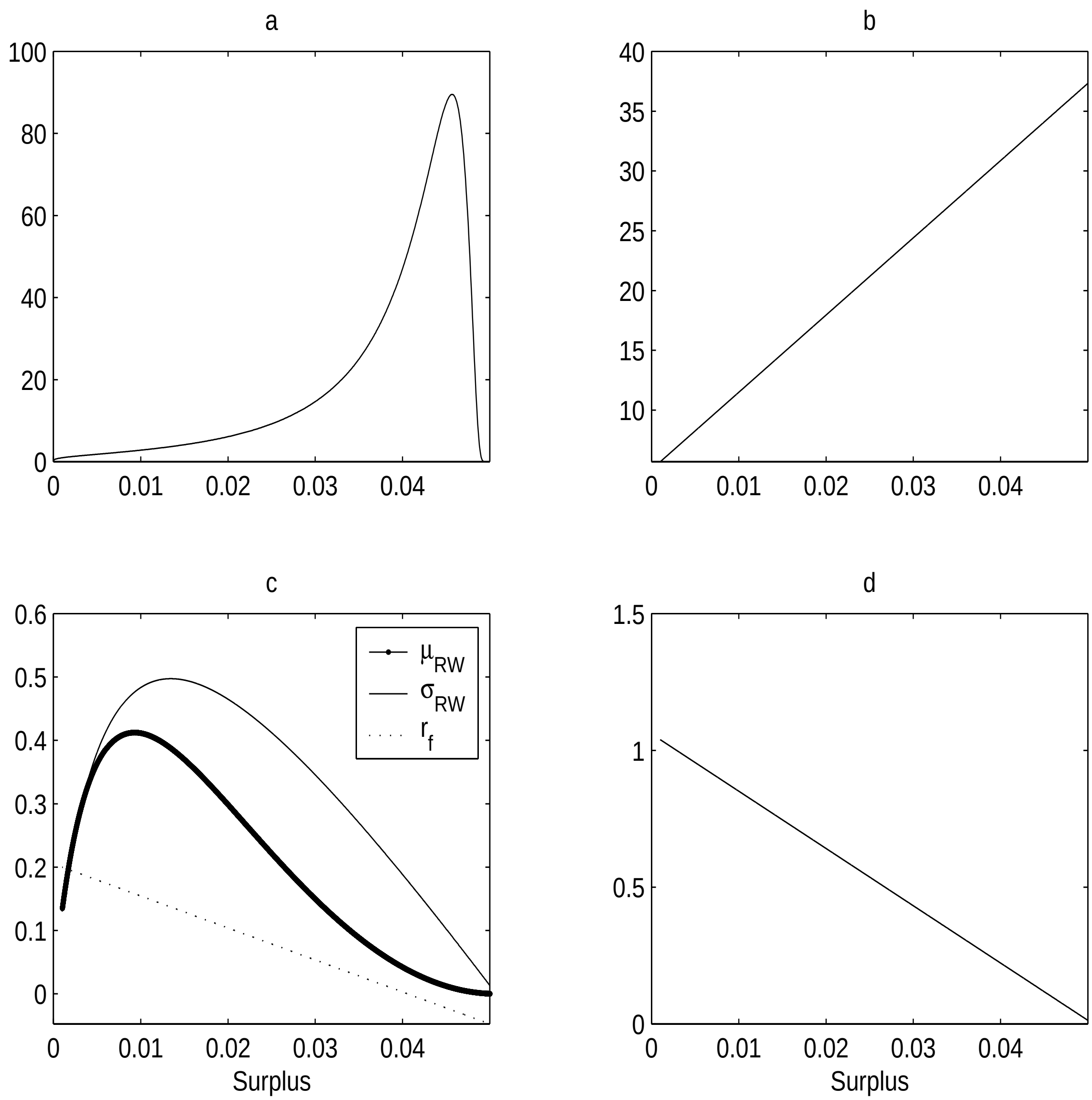
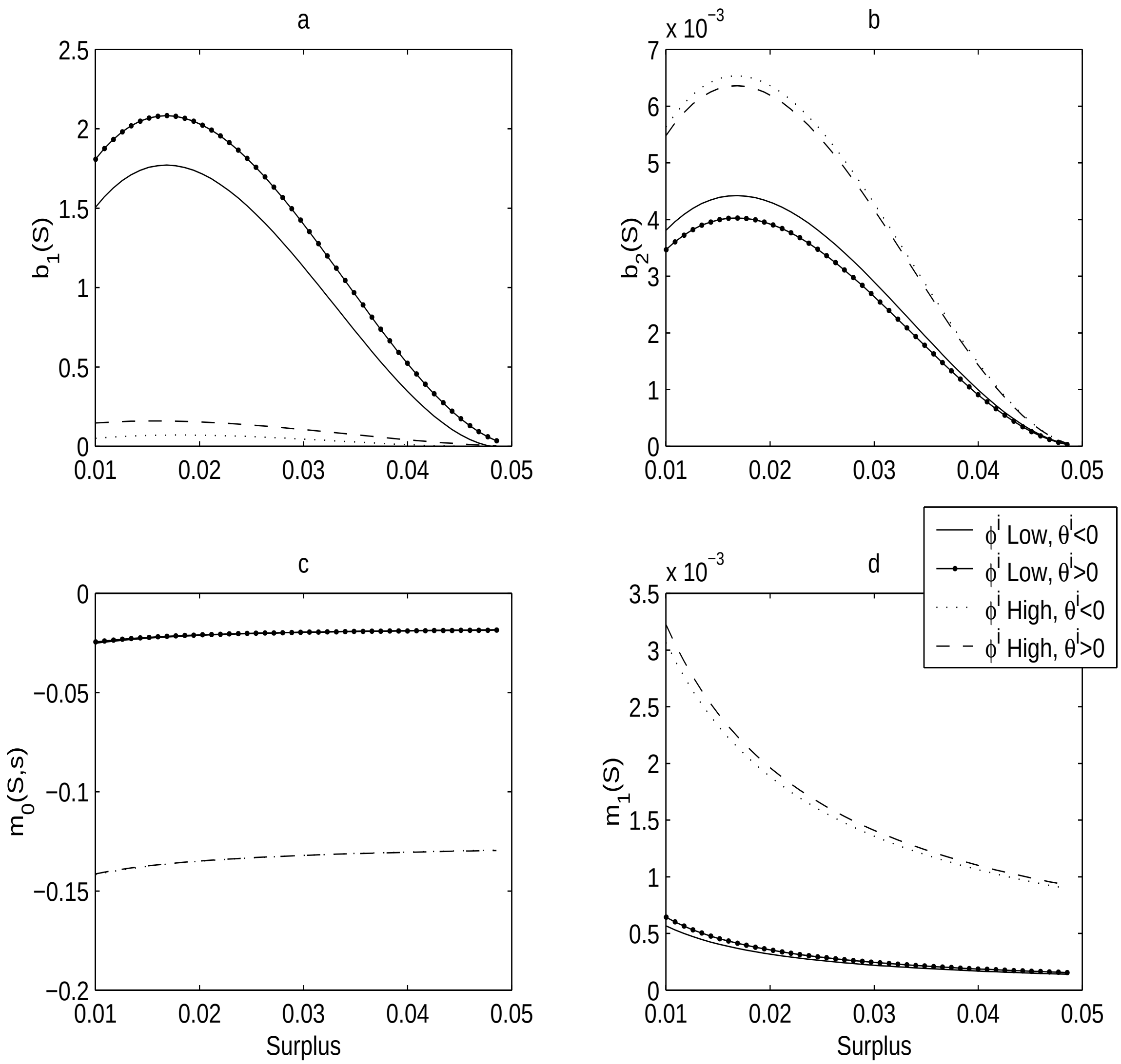

a
100
80
60
40
20
0
0
0.01
0.02
0.03
0.04
b
40
35
30
25
20
15
10
0
0.01
0.02
0.03
0.04
c
0.6
0.5
0.4
0.3
0.2
0.1
0
0
0.01
0.02
0.03
0.04
Surplus
μRW
σRW
rf
d
1.5
1
0.5
0
0
0.01
0.02
0.03
0.04
Surplus

Figure 1



a
b
c
d
b_1(S)
b_2(S)
m_0(S,s)
m_1(S)
x 10^-3
x 10^-3
Surplus
Surplus
φ^i Low, θ^i<0
φ^i Low, θ^i>0
φ^i High, θ^i<0
φ^i High, θ^i>0

Figure 2

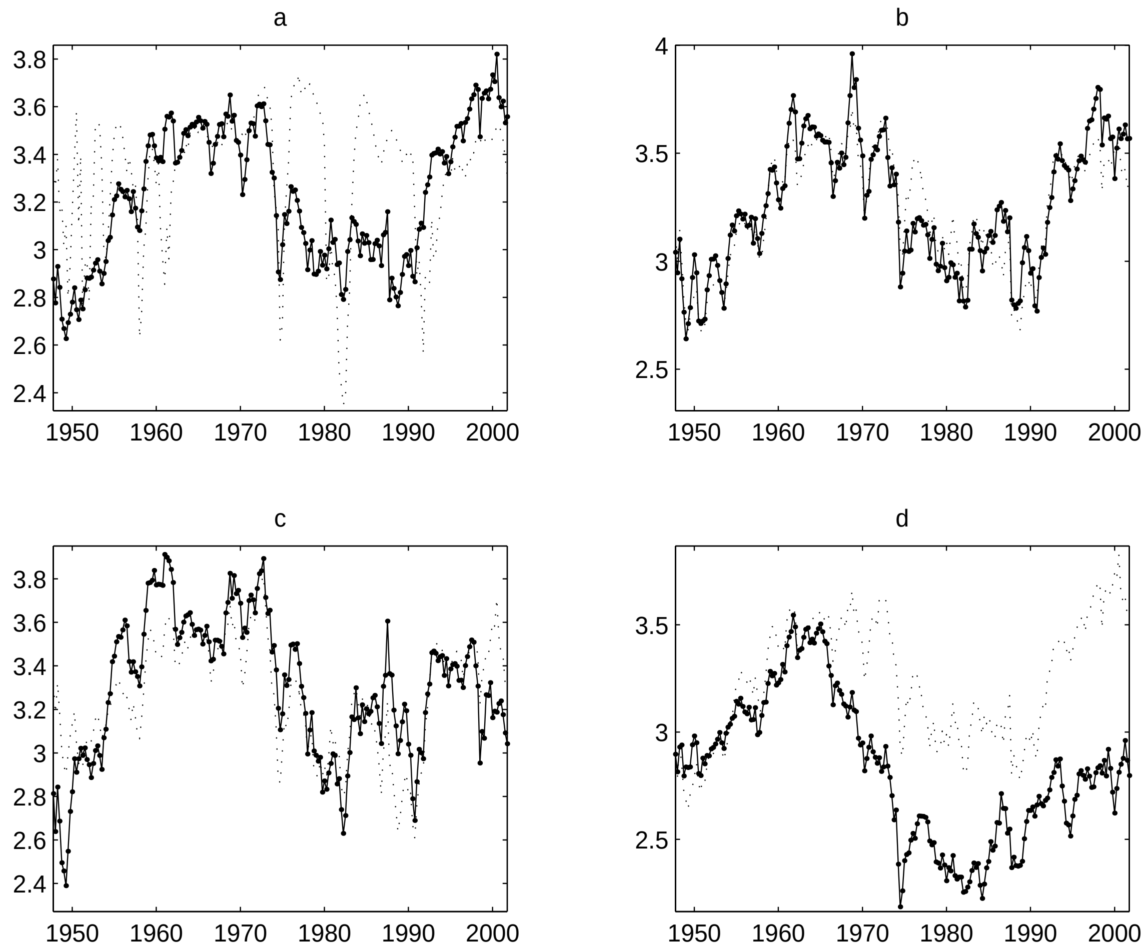


Figure 3

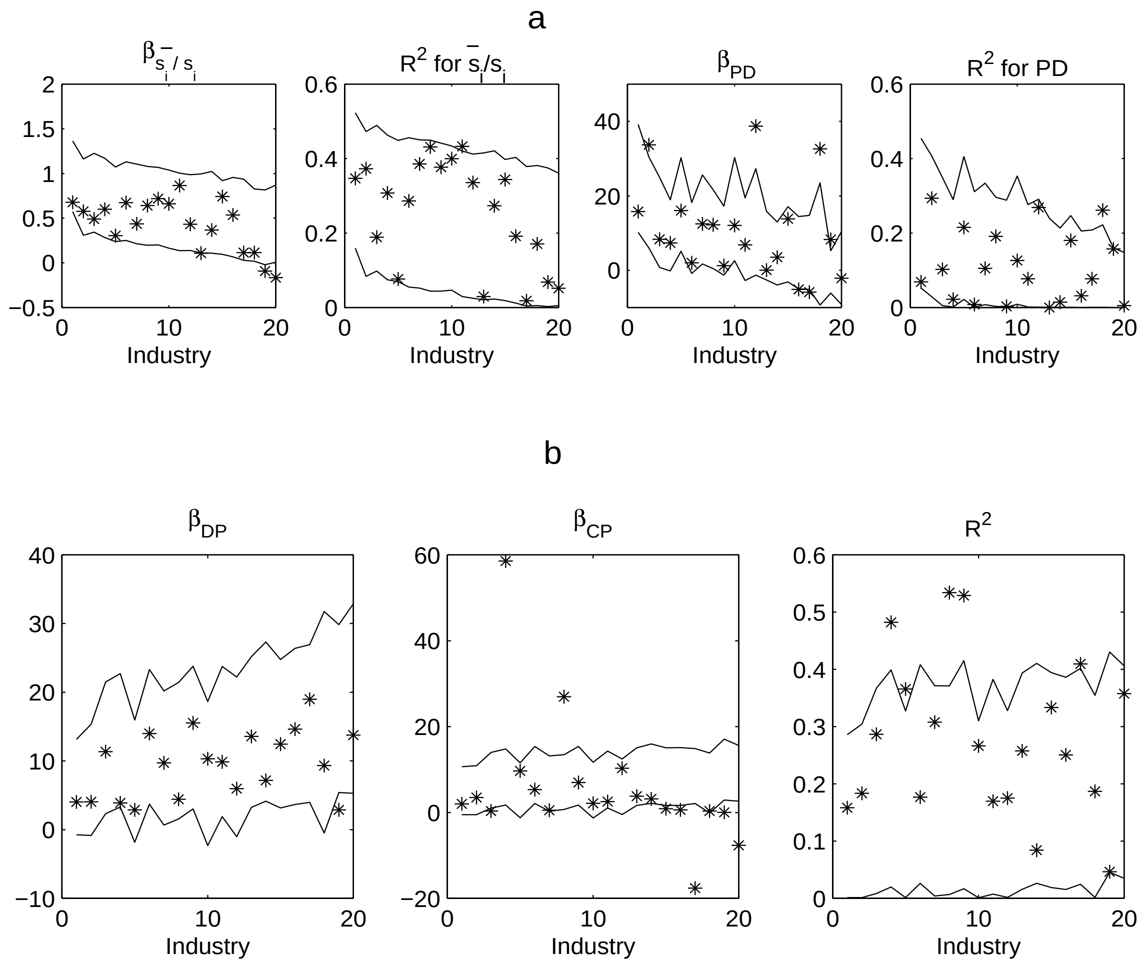

a
$\beta_{\bar{s}_i / s_i}$
R^2 for $\bar{s}_i/s_i$
β_{PD}
R^2 for PD
Industry
b
β_{DP}
β_{CP}
R^2
Industry

Figure 4

Figure 1: Aggregate Quantities. *a*: Stationary density function of S_t. *b*: Price consumption ratio of the total wealth portfolio. *c*: Expected excess returns and volatility of returns of the total wealth portfolio, and the risk free rate. *d*: Sharpe ratio of the total wealth portfolio. The parameters used are those of Table 1.

Figure 2: Predictive Regression Coefficients. Panels *a* and *b*: Theoretical slope coefficients $b_1(S)$ and $b_2(S)$ on the dividend price ratio and consumption to price ratio, respectively, in the return predictability regression. Panels *c* and *d*: Theoretical intercept $m_0(S, \mathbf{s})$ and slope coefficient $m_1(S)$ on the price dividend ratio in the dividend growth predictability regression. In addition to the parameters in Table 1, all plots use $\phi = 0.5$ or 0.05, and $\theta = 2/10000$ or $-2/10000$.

Figure 3: Log Price Dividend Ratios. *a*: Market Portfolio; *b* - *d*: Industries "Financials", "Paper" and "Utilities," respectively. All panels plot the log price dividend ratio in the data (solid dot) and implied by the model (dotted). The model-implied plots use the relative share $\overline{s}^i/s_t^i$ from the data, and the surplus consumption ratio S_t, as arguments in the price function. For the market portfolio, the surplus S_t is computed from consumption shocks, while for the three industry portfolios it is extracted from the price dividend ratio of the market portfolio itself. The three industries are those with the highest, median and lowest R^2 from a time series regression of price dividend ratios onto the model implied counterparts (Table 2).

Figure 4: Small Sample Simulation Results. Panel *a*: For each quantity reported on the title of each panel and for each of the 20 industries on the horizontal axis, we report the 90% confidence bands obtained through the simulation of 1000 samples of 55 years each (solid lines) and the data estimates (stars). The first two panels refer to the four year ahead predictive regression of future dividend growth using the relative share $\overline{s}^i/s_t^i$ as predictor, while the second two panels refer to the four year ahead predictive regression of future dividend growth using the price dividend ratio P_t^i/D_t^i as predictor. Industries are ordered on the horizontal axis inversely to the size of their respective speed of mean reversion ϕ^i, estimated in Table 1 (Industry 1 has the largest speed of mean reversion ϕ). Panel *b*: same quantities as in *a*, but for the multivariate four year ahead predictive regression of future returns, where both the dividend yield D_t^i/P_t^i and the consumption to price ratio C_t/P_t^i are used as predictors.