Netspar Network for Studies on Pensions, Aging and Retirement

Ralph Koijen

Jules van Binsbergen

Predictive Regressions: A Present – Value Approach

Discussion Paper 2008 - 029

January, 2009 (revised version from September, 2008)

Predictive Regressions: A Present-Value Approach*

Jules H. van Binsbergen
Stanford GSB[†]

Ralph S.J. Koijen
University of Chicago[‡]

This version: January 2009

Abstract

We propose a latent-variables approach within a present-value model to estimate the expected returns and expected dividend growth rates of the aggregate stock market. This approach aggregates information contained in the whole history of the price-dividend ratio and dividend growth rates to obtain predictors for future returns and dividend growth rates. We find that both returns and dividend growth rates are predictable with R-squared values ranging from 8.2-8.9 percent for returns and 13.9-31.6 percent for dividend growth rates. Both expected returns and expected dividend growth rates have a persistent component, but expected returns are more persistent than expected dividend growth rates.

*First version: November 2006. We thank Ravi Bansal, Lieven Baele, Geert Bekaert, Alan Bester, Michael Brandt, Alon Brav, Ben Broadbent, John Campbell, Hui Chen, John Cochrane, George Constantinides, Joost Driessen, Darrell Duffie, Gene Fama, Jesus Fernández-Villaverde, Xavier Gabaix, Eric Ghysels, Will Goetzmann, Cam Harvey, Chris Jones, Frank de Jong, Ron Kaniel, Martin Lettau, Hanno Lustig, Anthony Lynch, Toby Moskowitz, Justin Murfin, Theo Nijman, Ľuboš Pástor, Anamaría Pieschacón, Nick Polson, Matt Richardson, Juan Rubio-Ramírez, Yuliy Sannikov, Ken Singleton, Pilar Soriano, Allan Timmermann, Stijn Van Nieuwerburgh, Pietro Veronesi, Luis Viceira, Vish Viswanathan, Jessica Wachter, Bas Werker and seminar participants at the WFA meetings 2008, Chicago Booth, Duke University, the Stanford-Berkeley Seminar, Tilburg University, the Rotman School, UCSD-Rady School of Management, University of Southern California, University of Wisconsin-Madison, and UCLA for useful comments and suggestions.

†Stanford University, Graduate School of Business, Palo Alto, CA 94305. Phone: +1-650-721-1353. Email: jvb2@gsb.stanford.edu.

‡University of Chicago, Booth School of Business, Chicago, IL 60637. Phone: +1-773-834-4199. Email: ralph.koijen@chicagogsb.edu. Koijen is also associated with Netspar.

We propose a latent-variables approach within a present-value model to estimate the time series of expected returns and expected dividend growth rates of the aggregate stock market. We treat conditional expected returns and expected dividend growth rates as latent variables that follow an exogenously-specified time-series model. We combine this model with a Campbell and Shiller (1988) present-value model to derive the implied dynamics of the price-dividend ratio. We subsequently use a Kalman filter to construct the likelihood of our model and we estimate the parameters of the model by means of maximum likelihood. We find that both expected returns and expected dividend growth rates are time-varying and persistent, but expected returns are more persistent than expected dividend growth rates. The filtered series for expected returns and expected dividend growth rates are good predictors of realized returns and realized dividend growth rates, with R-squared values ranging from 8.2-8.9 percent for returns and 13.9-31.6 percent for dividend growth rates.

We consider an annual model to ensure that the dividend growth predictability we find is not simply driven by the seasonality in dividend payments.[1] Using an annual dividend growth series does however imply that we need to take a stance on how dividends received within a particular year are reinvested. Analogous to the way in which different investment strategies lead to different risk-return properties of portfolio returns, different reinvestment strategies for dividends within a year result in different dynamics of dividend growth rates.[2] We study two reinvestment strategies in detail. First, we reinvest dividends in a 30-day T-bill, which we call cash-invested dividends. Second, we reinvest dividends in the aggregate stock market. We will refer to these dividends as market-invested dividends. Market-invested dividends have been studied widely in the dividend-growth and return-forecasting literature.[3] We find the reinvestment strategy to matter for the time series properties of dividend growth. For instance, the volatility of market-invested dividend growth is twice as high as the volatility of cash-invested dividend growth. Within our model, we derive the link between the time-series models of dividend growth rates for different reinvestment strategies. This analysis demonstrates that if expected cash-invested dividend growth follows a first-order autoregressive process, then expected market-invested dividend growth has both a first-order autoregressive and a moving-average component.

[1]See also Cochrane (1994) and Lettau and Ludvigson (2005).

[2]For instance, the risk-return properties of the end-of-period capital will be different if an investor allocates its capital to stocks instead of Treasury bonds. By the same token, the properties of dividend growth rates depend on the reinvestment strategy chosen for dividends that are received within a particular year.

[3]See for instance Lettau and Ludvigson (2005), Cochrane (2007), and Lettau and Van Nieuwerburgh (2006).

The main assumptions we make in this paper concern the time-series properties for expected returns and expected dividend growth rates, which are the primitives of our model. We consider first-order autoregressive processes for expected *cash-invested* dividend growth and returns, and then derive the implied dynamics for expected *market-invested* dividend growth rates. Using this specification, we find that both returns and dividend growth rates are predictable, regardless of the reinvestment strategy. We can reject the null hypothesis that either expected returns or expected growth rates are constant at conventional significance levels. For both reinvestment strategies, we find, using a likelihood-ratio test, that expected returns are more persistent than expected growth rates. Also, innovations to both processes are positively correlated. Even though we find that future growth rates are predictable, we find that most of the unconditional variance in the price-dividend ratio stems from variation in discount rates, consistent with, for instance, Campbell (1991). If we decompose the conditional variance of stock returns, we find that cash-flow news can account for 34.6-49.4 percent of this variance, discount-rate news accounts for 118.4-215.3 percent, and the remainder is attributable to the covariance between cash-flow and discount-rate news.

Our model, in which we consider low-order autoregressive processes for expected returns and expected dividend growth rates, admits an infinite-order VAR representation in terms of dividend growth rates and price-dividend ratios.[4] Cochrane (2008) rigorously derives the link between our model in Section 1 and the VAR representation. This insightful analysis also demonstrates why our approach can improve upon predictive regressions that include only the current price-dividend ratio to predict future returns and dividend growth rates. Our latent-variables approach aggregates the whole history of price-dividend ratios and dividend growth rates to estimate expected returns and expected growth rates. This implies that we expand the information set that we use to predict returns and dividend growth rates. However, instead of adding lags to a VAR model, which increases the number of parameters to be estimated, we suggest a parsimonious way to incorporate the information contained in the history of price-dividend ratios and dividend growth rates. As Cochrane (2008) shows, our model introduces moving-average terms of price-dividend ratios and dividend growth rates, in addition to the current price-dividend ratios, and we find these moving-average terms to be relevant in predicting future returns and in particular dividend growth rates.

The insight that return predictability and dividend growth rate predictability are best studied jointly has already been pointed out by Cochrane (2007), Fama and French

[4] Pástor and Stambaugh (2006) show a similar result for return predictability. They abstract, however, from dividend growth predictability and do not impose the present-value relationship.

(1988), and Campbell and Shiller (1988). The main contribution of our paper is to model expected returns and expected dividend growth rates as latent processes and use filtering techniques to uncover them. Fama and French (1988) point out that the price-dividend ratio is only a noisy proxy for expected returns when the price-dividend ratio also moves due to expected dividend growth rate variation. This has also been argued by Menzly, Santos, and Veronesi (2004) and Goetzmann and Jorion (1995). The reverse argument also holds: the price-dividend ratio is only a noisy proxy for expected dividend growth when the price-dividend ratio also moves due to expected return variation. Our framework explicitly takes into account that the price-dividend ratio moves due to both expected return and expected dividend growth rate variation and the filtering procedure assigns price-dividend ratio shocks to either expected return and/or expected dividend growth rate shocks.

One may wonder why we choose an AR(1)-process to model expected *cash-invested* dividend growth as opposed to expected *market-invested* dividend growth. First, each reinvestment strategy for dividends corresponds to a different time-series model for expected returns and expected dividend growth rates. It could have been the case that, in fact, expected market-invested dividend growth is well described by an AR(1)-process. Given that most of the literature on return and dividend-growth-rate predictability has focused on market-invested dividend growth rates, this might seem like a more sensible first pass. In Section 5.4, we explore this specification and find that the persistence coefficient of expected market-invested dividend growth is negative in this case. By fixing the persistence parameter of expected dividend growth in the estimation of this specification, and maximizing over all other parameters, we show that the model's likelihood is bi-modal. This suggests that a simple first-order autoregressive process for expected market-invested dividend growth is too restrictive. We then perform a formal specification test and find that the model in Section 1, in which expected cash-invested dividend growth is an AR(1)-process and expected market-invested dividend growth an ARMA(1,1)-process, is preferred over a model in which expected market-invested dividend growth is an AR(1)-process.

Our paper is closely related to the recent literature on present-value models, see Cochrane (2007), Lettau and Van Nieuwerburgh (2006), Pástor and Veronesi (2003), Pástor and Veronesi (2006), Pástor, Sinha, and Swaminathan (2007), Bekaert, Engstrom, and Grenadier (2001), Bekaert, Engstrom, and Grenadier (2005), Burnside (1998), Ang and Liu (2004), and Brennan and Xia (2005). All of these papers provide expressions for the price-dividend or market-to-book ratio. However, in cases of Bekaert, Engstrom, and Grenadier (2001), Pástor and Veronesi (2003), Pástor and Veronesi (2006), Ang and Liu (2004), and Brennan and Xia (2005), the price-dividend ratio is an infinite sum, or indefinite integral, of exponentially-quadratic terms, which makes likelihood-based estimation and filtering computationally much more involved. Bekaert, Engstrom, and Grenadier (2001) and Ang and Liu (2004) estimate the model by means of GMM and model expected returns and expected growth rates as an affine function of a set of additional variables. Brennan and Xia (2005) use a two-step procedure to estimate their model and use long-term forecasts for expected returns to recover an estimate of the time series of (instantaneous) expected returns, in turn. Alternatively, Lettau and Van Nieuwerburgh (2006) set up a linearized present-value model and recover structural parameters from reduced-form estimators. They subsequently test whether the present-value constraints are violated. They impose, however, that the persistence of expected returns and expected growth rates are equal.[5] Our paper also relates to Brandt and Kang (2004), Pástor and Stambaugh (2006), and Rytchkov (2007) who focus on return predictability using filtering techniques. We contribute to this literature by focusing on the interaction between return and dividend growth predictability, and by showing that the reinvestment strategy of dividends has an impact on the specification of the present-value model.

The paper proceeds as follows. In Section 1, we present the linearized present-value model. In Section 2 we discuss the data, our estimation procedure, and the link between the two reinvestment strategies. In Section 3 we present our estimation results and compare our empirical results to predictive regressions. Section 4 discusses hypothesis testing including the tests for (the lack of) return and dividend growth rate predictability. Section 5 discusses several additional implications and some robustness checks, and Section 6 concludes.

1 Present-value model

In this section we present a log-linearized present-value model in the spirit of Campbell and Shiller (1988).[6] We assume that both expected returns and expected dividend growth rates are latent variables. We first consider the model in which both latent variables follow

[5]Under these assumptions, no filtering is required to uncover expected returns and expected dividend growth rates. We test, using a likelihood-ratio test, whether the persistence of expected returns and expected dividend growth rates is equal, and we reject this hypothesis.

[6]Cochrane (1991), Cochrane (2007), and Lettau and Van Nieuwerburgh (2006) present a version of this model in which it is assumed that expected growth rates are constant, or expected returns are equally persistent as expected growth rates.

an AR(1)-process.[7] However, we can allow for higher-order VARMA representations for these variables, some of which we further explore when we study different reinvestment strategies.

Let r_{t+1} denote the total log return on the aggregate stock market:

$$r_{t+1} \equiv \log\left(\frac{P_{t+1} + D_{t+1}}{P_t}\right), \tag{1}$$

let PD_t denote the price-dividend ratio of the aggregate stock market:

$$PD_t \equiv \frac{P_t}{D_t},$$

and let Δd_{t+1} denote the aggregate log dividend growth rate:

$$\Delta d_{t+1} \equiv \log\left(\frac{D_{t+1}}{D_t}\right).$$

We model both expected returns (μ_t) and expected dividend growth rates (g_t) as an AR(1)-process:

$$\mu_{t+1} = \delta_0 + \delta_1 (\mu_t - \delta_0) + \varepsilon^{\mu}_{t+1}, \tag{2}$$

$$g_{t+1} = \gamma_0 + \gamma_1 (g_t - \gamma_0) + \varepsilon^{g}_{t+1}, \tag{3}$$

where:

$$\begin{aligned} \mu_t &\equiv E_t [r_{t+1}], \\ g_t &\equiv E_t [\Delta d_{t+1}]. \end{aligned}$$

The distribution of the shocks ε^{μ}_{t+1} and ε^{g}_{t+1} will be specified shortly. The realized dividend growth rate is equal to the expected dividend growth rate plus an orthogonal shock:

$$\Delta d_{t+1} = g_t + \varepsilon^{D}_{t+1}.$$

[7] Many authors have argued that expected returns are likely to be persistent, including Fama and French (1988), Campbell and Cochrane (1999), Ferson, Sarkissian, and Simin (2003), and Pástor and Stambaugh (2006). Further, it has been argued that expected dividend growth rates have a persistent component, see for example Bansal and Yaron (2004), Menzly, Santos, and Veronesi (2004), and Lettau and Ludvigson (2005).

Defining $pd_t \equiv \log(PD_t)$, we can write the log-linearized return as:

$$r_{t+1} \simeq \kappa + \rho pd_{t+1} + \Delta d_{t+1} - pd_t,$$

with $\overline{pd} = E[pd_t]$, $\kappa = \log\left(1 + \exp\left(\overline{pd}\right)\right) - \rho\overline{pd}$ and $\rho = \frac{\exp\left(\overline{pd}\right)}{1+\exp\left(\overline{pd}\right)}$, as in Campbell and Shiller (1988). If we iterate this equation, and using the AR(1) assumptions (2)-(3), it follows that:

$$pd_t = A - B_1(\mu_t - \delta_0) + B_2(g_t - \gamma_0),$$

see Appendix A, with $A = \frac{\kappa}{1-\rho} + \frac{\gamma_0 - \delta_0}{1-\rho}$, $B_1 = \frac{1}{1-\rho\delta_1}$, and $B_2 = \frac{1}{1-\rho\gamma_1}$. The log price-dividend ratio is linear in the expected return μ_t and the expected dividend growth rate g_t. The loading of the price-dividend ratio on expected returns and expected dividend growth rates depends on the relative persistence of these variables (δ_1 versus γ_1). The three shocks in the model, which are shocks to expected dividend growth rates (ε^g_{t+1}), shocks to expected returns (ε^μ_{t+1}), and realized dividend growth shocks (ε^d_{t+1}), have mean zero, covariance matrix:

$$\Sigma \equiv var\left(\begin{bmatrix} \varepsilon^g_{t+1} \\ \varepsilon^\mu_{t+1} \\ \varepsilon^d_{t+1} \end{bmatrix}\right) = \begin{bmatrix} \sigma^2_g & \sigma_{g\mu} & \sigma_{gD} \\ \sigma_{g\mu} & \sigma^2_\mu & \sigma_{\mu D} \\ \sigma_{gD} & \sigma_{\mu D} & \sigma^2_D \end{bmatrix},$$

and are independent and identically distributed over time. Further, in the maximum-likelihood estimation procedure, we assume that the shocks are jointly normally distributed.

2 Data and estimation

2.1 Data

We obtain the with-dividend and without-dividend monthly returns on the value-weighted portfolio of all NYSE, Amex, and Nasdaq stocks for the period 1946-2007 from the Center for Research in Security Prices (CRSP). We then use these data to construct our annual data for aggregate dividends and prices. We consider two reinvestment strategies. First, we consider dividends reinvested in 30-day T-bills and compute the corresponding series for dividend growth, the price-dividend ratio, and returns. Data on the 30-day T-bill rate is also obtained from CRSP. Secondly, we consider dividends reinvested in the aggregate stock market and compute the corresponding series for dividend growth, the price-dividend ratio, and returns. The latter reinvestment strategy is commonly used in the return predictability literature. It causes annual dividend growth to be highly volatile with an annual unconditional volatility of 12.3% versus a volatility of 6.2% for cash-invested dividend growth, see also Table 1 and Figure 1 - 3. In the next three subsections we present our estimation procedure. In Section 2.2 we discuss our estimation procedure for the model where dividends are reinvested cash, and in Section 2.4 we discuss our estimation procedure when dividends are reinvested in the market. Section 2.3 discusses the link between both models.

2.2 State space representation: cash-invested dividends

Our model features two latent state variables, μ_t and g_t. We assume that each of these is an AR(1)-process. The de-meaned state variables are:

$$\begin{aligned}\hat{\mu}_t &= \mu_t - \delta_0, \\ \hat{g}_t &= g_t - \gamma_0.\end{aligned}$$

The model has two transition equations:

$$\begin{aligned}\hat{g}_{t+1} &= \gamma_1 \hat{g}_t + \varepsilon^g_{t+1}, \\ \hat{\mu}_{t+1} &= \delta_1 \hat{\mu}_t + \varepsilon^\mu_{t+1},\end{aligned}$$

and two measurement equations:

$$\begin{aligned}\Delta d_{t+1} &= \gamma_0 + \hat{g}_t + \varepsilon^D_{t+1}, \\ pd_t &= A - B_1 \hat{\mu}_t + B_2 \hat{g}_t.\end{aligned}$$

Because the second measurement equation contains no error term, we can substitute the equation for pd_t into the transition equation for de-meaned expected returns, to arrive at the final system that has just one transition and two measurement equations:

$$\begin{aligned}\hat{g}_{t+1} &= \gamma_1 \hat{g}_t + \varepsilon^g_{t+1}, \quad (4)\\ \Delta d_{t+1} &= \gamma_0 + \hat{g}_t + \varepsilon^D_{t+1}, \quad (5)\\ pd_{t+1} &= (1-\delta_1) A + B_2 (\gamma_1 - \delta_1) \hat{g}_t + \delta_1 pd_t - B_1 \varepsilon^\mu_{t+1} + B_2 \varepsilon^g_{t+1}. \quad (6)\end{aligned}$$

It may be surprising that there is no measurement equation for returns. However, the measurement equation for dividend growth rates and the price-dividend ratio implies the measurement equation for returns. As all equations are linear, we can compute the likelihood of the model using a Kalman filter (Hamilton (1994)). We then use conditional maximum-likelihood estimation (MLE) to estimate the vector of parameters:

$$\Theta \equiv \left(\gamma_0, \delta_0, \gamma_1, \delta_1, \sigma_g, \sigma_\mu, \sigma_D, \rho_{g\mu}, \rho_{gD}, \rho_{\mu D}\right).$$

The details of this estimation procedure are described in Appendix B. We use conditional maximum likelihood to enhance the comparison with the standard predictive regressions approach. When we use unconditional maximum likelihood estimation, this does not substantially change our qualitative or quantitative results, apart from slightly increasing both the R^2 value for returns and for dividend growth rates. We maximize the likelihood using simulated annealing. This maximization algorithm is designed to search for the global maximum (Goffe, Ferrier, and Rogers (1994)).

2.3 Reinvesting dividends and modeling growth rates

We assume that cash-invested expected growth rates are an AR(1)-process. In this section, we find the observable implications for market-invested dividends. To illustrate why the reinvestment strategy is potentially important for the time-series model of dividend growth rates, we present the following extreme example. Consider the case where year-t prices are recorded on December 31st and year-$(t+1)$ dividends are all paid out one day later, on January 1st. Denote by D_{t+1} the dividends paid out on January 1st. Assuming (for ease of exposition) that the one-year interest rate is 0, the end-of-year cash-invested dividends are simply given by D_{t+1}. However, the end-of-year market-invested dividends are given by:

$$D_{t+1}^M = D_{t+1}\exp\left(r_{t+1}\right), \tag{7}$$

where r_{t+1} denotes the aggregate stock market return defined in (1). Even though realized dividend growth rates are strongly dependent on the reinvestment strategy, the aggregate stock market return does not. The correlation between cum-dividend returns where dividends are reinvested in the market, denoted by r_{t+1}^M, and cum-dividend returns where dividends are reinvested in the risk free rate, denoted by r_{t+1}, is 0.9999, see also Figure 3. As such, from an empirical perspective, these two series can be used interchangeably. The observed market-invested dividend growth rates are then given by:

$$\Delta d_{t+1}^{M} = \log\left(\frac{D_{t+1}^{M}}{D_{t}^{M}}\right) = \log\left(\frac{D_{t+1}}{D_{t}}\right) + r_{t+1} - r_{t}.$$

First, this expression suggests that the lagged return on the market is a candidate predictor of market-invested dividend growth rates where a high past return predicts low future dividend growth. If the return on the market in period t is high, this increases the dividend growth rate at time t, but it implies a lower dividend growth rate at time $t+1$ relative to cash-invested dividends. Second, the expression suggests that reinvesting dividends in the market can add substantial volatility to dividend growth rates.

In reality, dividends are paid out throughout the year, for example at the end of each quarter. To capture the impact of reinvesting dividends in the aggregate stock market, we consider the following reduced-form representation:

$$D_{t+1}^{M} = D_{t+1}\exp(\varepsilon_{t+1}^{M}),$$

in which D_{t+1} denotes the cash-invested dividend. We assume that ε_{t+1}^{M} is i.i.d. over time with mean zero and standard deviation σ_M. Further, we allow for correlation between ε_{t+1}^{M} and aggregate market returns:

$$\rho_M \;\; = \;\; corr\left(\varepsilon_{t+1}^{M}, \varepsilon_{t+1}^{r}\right),$$

with $\varepsilon_{t+1}^{r} \equiv r_{t+1} - \mu_t \approx -B_1\rho\varepsilon_{t+1}^{\mu} + B_2\rho\varepsilon_{t+1}^{g} + \varepsilon_{t+1}^{d}$. In our previous example, where all dividends were paid out at the beginning of the year, this correlation is close to 1 and $r_{t+1} \approx \varepsilon_{t+1}^{M}$. If dividend payments are made throughout the year, we expect a positive value for ρ_M, but not necessarily close to one. Using this model, we can decompose ε_{t+1}^{M} into a part that is correlated with ε_{t+1}^{r} and a part that is orthogonal to ε_{t+1}^{r}:

$$\varepsilon_{t+1}^{M} = \beta_M\varepsilon_{t+1}^{r} + \varepsilon_{t+1}^{M\perp}, \tag{8}$$

in which $\beta_M = \rho_M\sigma_M/\sigma_r$, $\sigma_r = \sqrt{var(\varepsilon_{t+1}^{r})}$, and $\varepsilon_{t+1}^{M\perp}$ is orthogonal to ε_{t+1}^{r}. To keep the model parsimonious, we assume that all correlation between ε_{t+1}^{M} and the structural shocks in our model, that is, ε_{t+1}^{g}, ε_{t+1}^{μ}, and ε_{t+1}^{d}, comes via the aggregate market return. The latter assumption implies that $cov(\varepsilon_{t+1}^{M\perp}, \varepsilon_{t+1}^{g}) = cov(\varepsilon_{t+1}^{M\perp}, \varepsilon_{t+1}^{d}) = cov(\varepsilon_{t+1}^{M\perp}, \varepsilon_{t+1}^{\mu}) = 0$.

Using that expected growth rates for cash-invested dividends follow an AR(1)-process, we can derive the expected growth rate of market-invested dividends:

$$\begin{aligned}\Delta d^M_{t+1} &= \Delta d_{t+1} + \varepsilon^M_{t+1} - \varepsilon^M_t \\ &= g_t + \varepsilon^d_{t+1} + \varepsilon^M_{t+1} - \varepsilon^M_t,\end{aligned} \tag{9}$$

where $g_t \equiv E_t[\Delta d_{t+1}]$. This implies that $g^M_t \equiv E_t[\Delta d^M_{t+1}] = g_t - \varepsilon^M_t$. The dynamics of expected market-invested dividend growth therefore reads:

$$\begin{aligned}g^M_{t+1} &= g_{t+1} - \varepsilon^M_{t+1} \\ &= \gamma_0 + \gamma_1(g_t - \gamma_0) + \varepsilon^g_{t+1} - \varepsilon^M_{t+1} \\ &= \gamma_0 + \gamma_1(g^M_t - \gamma_0) + \gamma_1\varepsilon^M_t + \varepsilon^g_{t+1} - \varepsilon^M_{t+1}.\end{aligned} \tag{10}$$

This shows that expected market-invested dividend growth is not a first-order autoregressive process, but instead an ARMA(1,1)-process.

Summary of the model We can then summarize the model for market-invested dividend growth as follows:

$$\begin{aligned}\Delta d^M_{t+1} &= g^M_t + \varepsilon^{DM}_{t+1}, \\ g^M_{t+1} &= \gamma_0 + \gamma_1(g^M_t - \gamma_0) + \gamma_1\varepsilon^M_t + \varepsilon^{gM}_{t+1}, \\ \mu_{t+1} &= \delta_0 + \delta_1(\mu_t - \delta_0) + \varepsilon^\mu_{t+1}, \\ pd^M_t &= A - B_1(\mu_t - \delta_0) + B_2(g^M_t - \gamma_0) + (B_2 - 1)\,\varepsilon^M_t,\end{aligned}$$

in which we define:

$$\begin{aligned}\varepsilon^{DM}_{t+1} &\equiv \varepsilon^d_{t+1} + \varepsilon^M_{t+1}, \\ \varepsilon^{gM}_{t+1} &\equiv \varepsilon^g_{t+1} - \varepsilon^M_{t+1},\end{aligned}$$

and recall that $g^M_t = g_t - \varepsilon^M_t$. Market-invested dividend growth rates Δd^M_{t+1} are equal to expected market invested dividend growth rates g^M_t, plus an orthogonal shock ε^{DM}_{t+1}. This orthogonal shock consists of two parts. A part due to the unexpected change in the payout of firms, ε^d_{t+1}, and a part related to the performance of the stock market during that year, ε^M_{t+1}. The expected market-invested dividend growth rate g^M_t also consists of two parts. The first part g_t is driven by the expected change in the payout of firms and the second part ε^M_t relates to the stock market performance of the previous year. As g_t is an AR(1)-process, market-invested dividend growth is an ARMA(1,1)-process. Finally, the expected return still is an AR(1)-process as in the case of cash-invested dividends.

The parameters of the covariance matrix are given by:

$$\begin{aligned}
\sigma_D^{M2} &\equiv var(\varepsilon_{t+1}^{DM}) = \sigma_D^2 + \sigma_M^2 + 2\sigma_{DM}, \\
\sigma_g^{M2} &\equiv var(\varepsilon_{t+1}^{gM}) = \sigma_g^2 + \sigma_M^2 - 2\sigma_{gM}, \\
\sigma_{\mu D}^{M2} &\equiv cov(\varepsilon_{t+1}^{\mu}, \varepsilon_{t+1}^{DM}) = \sigma_{\mu D} + \sigma_{\mu M}, \\
\sigma_{\mu g}^{M2} &\equiv cov(\varepsilon_{t+1}^{\mu}, \varepsilon_{t+1}^{gM}) = \sigma_{\mu g} - \sigma_{\mu M}, \\
\sigma_{gD}^{M2} &\equiv cov(\varepsilon_{t+1}^{gM}, \varepsilon_{t+1}^{DM}) = -\sigma_M^2,
\end{aligned}$$

and σ_{DM}, σ_{gM}, and $\sigma_{\mu M}$ are derived in Appendix B. The correlations are subsequently defined as: $\rho_{gD}^M \equiv \sigma_{gD}^M/(\sigma_g^M \sigma_D^M)$, $\rho_{\mu D}^M \equiv \sigma_{\mu D}^M/(\sigma_\mu \sigma_D^M)$, and $\rho_{\mu g}^M \equiv \sigma_{\mu g}^M/(\sigma_\mu \sigma_g^M)$.

2.4 State-space representation: market-invested dividends

We define the two de-meaned state variables as:

$$\begin{aligned}
\mu_t &= \delta_0 + \hat{\mu}_t, \\
g_t &= \gamma_0 + \hat{g}_t.
\end{aligned}$$

Again, the model has two transition equations:

$$\begin{aligned}
\hat{g}_{t+1} &= \gamma_1 \hat{g}_t + \varepsilon_{t+1}^g, \\
\hat{\mu}_{t+1} &= \delta_1 \hat{\mu}_t + \varepsilon_{t+1}^{\mu},
\end{aligned}$$

and two measurement equations:

$$\begin{aligned}
\Delta d_{t+1}^M &= \gamma_0 + \hat{g}_t + \varepsilon_{t+1}^D + \varepsilon_{t+1}^M - \varepsilon_t^M, \\
pd_t^M &= A - B_1 \hat{\mu}_t + B_2 \hat{g}_t - \varepsilon_t^M.
\end{aligned}$$

We are now using dividends reinvested in the market to compute the log dividend growth rate and the log price-dividend ratio. As before, we can substitute the price-dividend ratio for one latent variable to arrive at the final system consisting of two measurement equations and one transition equation:

$$\Delta d_{t+1}^{M} = \gamma_0 + \hat{g}_t + \varepsilon_{t+1}^{D} + \varepsilon_{t+1}^{M} - \varepsilon_t^{M}, \tag{11}$$

$$pd_{t+1}^{M} = (1-\delta_1)\,A + B_2\,(\gamma_1 - \delta_1)\,\hat{g}_t + \delta_1 pd_t^{M} - B_1\varepsilon_{t+1}^{\mu} + B_2\varepsilon_{t+1}^{g} - \varepsilon_{t+1}^{M} + \delta_1\varepsilon_t^{M}, \tag{12}$$

$$\hat{g}_{t+1} = \gamma_1\hat{g}_t + \varepsilon_{t+1}^{g}. \tag{13}$$

As all equations are still linear, we can compute the likelihood of the model using the Kalman filter, and use conditional MLE to estimate the parameters.

2.5 Identification

In our model, all but one of the parameters in the covariance matrix are identified.[8] We choose to normalize the correlation between realized dividend growth shocks (ε_{t+1}^{d}) and expected dividend growth shocks (ε_{t+1}^{g}) to zero.

3 Results

3.1 Estimation results: cash-invested dividends

Table 2 shows the maximum-likelihood estimates of the parameters of the present-value model described in equations (4)-(6) where we use dividend data reinvested in the risk-free rate. We estimate the unconditional expected log return equal to $\delta_0 = 9.0\%$ and the unconditional expected log growth rate of dividends to be $\gamma_0 = 6.2\%$. Further, we find expected returns to be highly persistent, consistent with Fama and French (1988), Campbell and Cochrane (1999), Ferson, Sarkissian, and Simin (2003), and Pástor and Stambaugh (2006) with an annual persistence coefficient (δ_1) of 0.932. The estimated persistence of expected dividend growth rates equals 0.354, which is less than the estimated persistence of expected returns. We test whether this difference is significant with a likelihood-ratio test in Section 4. Further, shocks to expected returns and expected dividend growth rates are positively correlated.[9] In our state-space model we compute the R^2 values for returns and dividend growth rates as (see also Harvey (1989)):

$$R_{Ret}^2 = 1 - \frac{\hat{var}\left(r_{t+1} - \mu_t^F\right)}{\hat{var}\left(r_t\right)}, \tag{14}$$

$$R_{Div}^2 = 1 - \frac{\hat{var}\left(\Delta d_{t+1} - g_t^F\right)}{\hat{var}\left(\Delta d_{t+1}\right)}, \tag{15}$$

[8] See also Cochrane (2008) and Rytchkov (2007).

[9] This is consistent with Menzly, Santos, and Veronesi (2004) and Lettau and Ludvigson (2005).

where $v\hat{a}r$ is the sample variance, μ_t^F is the filtered series for expected returns (μ_t), and g_t^F is the filtered series for expected dividend growth rates (g_t). Alternatively, we can compute the R-squared values within our model as if g_t and μ_t are observed and do not need to be filtered. However, to compare our results to OLS, we use the filtered series because g_t and μ_t are latent processes. The R^2 value for returns is equal to 8.2% and for dividend growth rates it equals 13.9%.

3.2 Estimation results: market-invested dividends

Table 3 shows maximum-likelihood estimates of the parameters of the present-value model described in equations (11)-(13). We now find $\delta_0 = 8.6\%$ and $\gamma_0 = 6.0\%$. Further, we find that the persistence coefficient of expected returns (δ_1) equals 0.957 and the persistence of expected dividend growth rates equals 0.638, indicating, as before, that expected returns are more persistent than expected dividend growth rates. Compared to the case of cash-invested dividends, we find a higher correlation $\rho_{\mu g}$ and a higher persistence coefficient γ_1. When we reinvest dividends in the market, a high expected return will lead to a high expected growth rate. This increases the correlation between expected returns and expected dividend growth rates. Further, the part of the expected dividend growth rate that relates to the higher expected return will be persistent, due the high persistence of expected returns. This increases the estimated persistence coefficient of expected dividend growth.

The R^2 values are in this case computed as:

$$R^2_{Ret} = 1 - \frac{v\hat{a}r\left(r^M_{t+1} - \mu^F_t\right)}{v\hat{a}r\left(r^M_t\right)},$$

$$R^2_{Div_M} = 1 - \frac{v\hat{a}r\left(\Delta d^M_{t+1} - g^{M,F}_t\right)}{v\hat{a}r\left(\Delta d^M_{t+1}\right)},$$

where $v\hat{a}r$ is the sample variance, μ_t^F is the filtered series for expected returns (μ_t) and $g_t^{M,F}$ is the filtered series for expected dividend growth rates (g_t^M). For returns, we find an R^2 value of 8.9% and for dividend growth rates, we find an R^2-value of 31.6%. Further, the standard deviation of the shock ε_t^M equals 5.4% and the correlation between ε_t^M and the unexpected return on the aggregate market is $\rho_M = 0.59$. If all dividends would have been paid out at the beginning of the year, ε_t^M would closely resemble the market return and we would expect a standard deviation σ_M equal to that of the aggregate market and the correlation to be close to 1. When all dividend payments are paid out at the end of the year, then we would expect a value of σ_M close to zero and a correlation close to

0. When dividends are paid out throughout the year, as they are in our data set, it is reassuring to find values of σ_M and ρ_M in between these two extreme cases. It suggests that this reduced-form representation indeed captures, at least partially, the reinvestment of dividends in the aggregate market.

3.3 Comparison with OLS regressions

As a benchmark for our latent-variables approach, we also report results from the following predictive OLS regressions:[10]

$$r_{t+1} = \alpha_r + \beta_r pd_t + \varepsilon_{t+1}^{r,OLS},$$
$$\Delta d_{t+1} = \alpha_d + \beta_d pd_t + \varepsilon_{t+1}^{d,OLS}.$$

The results are summarized in Table 4. For market-invested dividends, the return regression has a predictive coefficient of $\beta_r = -0.10$ with an R^2 value of 7.96% and a t-statistic of -2.19, where we use OLS standard errors to compute the t-statistic. The dividend growth rate regression results in a predictive coefficient of $\beta_d = -0.04$, with an R^2 value of 1.56% and a t-statistic of -0.91. The dividend growth rate regression has an insignificant coefficient, which seems to have the wrong sign, in the sense that a *high* price-dividend ratio predicts a *low* expected dividend growth rate as opposed to a high expected dividend growth rate.[11] Appendix C shows why our filtering approach can improve upon predictive regressions. In addition to the lagged price-dividend ratio, we use the entire history of dividend growth rates and price-dividend ratios to predict future growth rates and returns:

$$\begin{aligned}
\Delta d_t &= a_0^d + \sum_{i=0}^{\infty} a_{1i}^d pd_{t-i-1} + \sum_{i=0}^{\infty} a_{2i}^d \Delta d_{t-i-1} + \varepsilon_t^{d*}, \\
r_t &= a_0^r + \sum_{i=0}^{\infty} a_{1i}^r pd_{t-i-1} + \sum_{i=0}^{\infty} a_{2i}^r \Delta d_{t-i-1} + \varepsilon_t^{r*},
\end{aligned}$$

and Appendix C defines the coefficients and innovations (ε_t^{d*} and ε_t^{r*}). Our filtering approach therefore uses more information and aggregates this information in a

[10] These regressions have been studied widely in the literature, and an incomplete list of references includes Cochrane (2007), Lettau and Van Nieuwerburgh (2006), and Stambaugh (1999).

[11] Given that for values of γ_1 smaller than $1/\rho$, the coefficient B_2 is bigger than zero, we would expect a positive sign in this regression. The price-dividend ratio is a noisy proxy for expected dividend growth rates when the price-dividend ratio also moves due to expected return variation, which can lead to the wrong sign in the regression. Binsbergen and Koijen (2008) show that the price-dividend ratio relates negatively to expected growth rates if: $\frac{\sigma_g^2}{1-\gamma_1^2} < \frac{B_1}{B_2}\frac{\sigma_{\mu g}}{1-\gamma_1\delta_1}$.

parsimonious way (see also Cochrane (2008)). The present-value approach we propose, in combination with the Kalman filter, allows us therefore to expand the information set without increasing the number of parameters.

For cash-invested dividends, the return regression has a predictive coefficient of $\beta_r = -0.10$ with an R-squared value of 8.20% and a t-statistic of -2.32. The dividend growth rate regression results in a predictive coefficient of $\beta_d = -0.01$, with an R^2 value of 0.01% and a t-statistic of -0.91.

We have argued that the reinvestment strategy matters for realized dividend growth and that market-invested dividend growth is more volatile than cash-invested dividend growth due to the volatility of stock returns. Further, we have argued that apart from this added volatility as a result of the reinvestment return, realized market-invested dividend growth can be well described by an ARMA(1,1) process. To further explore this argument we present results for several OLS regressions for market-invested dividends. The results are summarized in Table 5. When we include in the regression a constant term and an AR(1)-term, we find a negative coefficient which is significant at the 10% level. The R^2 value is low and equal to 5.0%. Further, when we include a constant term, an AR(1)-term and an MA(1) term, none of the latter two coefficients shows up significantly in the regression and the R^2 value remains low at 6.3%.

However, when we control for the lagged return (r_{t-1}) in this regression the results substantially change, with an AR(1) coefficient value of 0.782, and an MA(1) coefficient of -0.979, both statistically significant at the 1% level. Further, the lagged return enters significantly, as expected, with a negative coefficient of -0.313. The R^2 value of the latter regression equals 27.5%. In fact, including the lagged return as the sole regressor already leads to an R^2 value of 22.3%.[12] The R^2 value of 27.5% is still lower than the 31.6% that we achieve by filtering, even though the OLS regressions allow for an additional degree of freedom compared to the specification for dividend growth in the filtering procedure. To increase the R^2 further we need to include the information contained in the price-dividend ratio, which in the OLS regressions above, we have not yet explored. However, including the lagged price-dividend ratio in the regression does not lead to a higher R^2 value, the coefficient is not significant and the coefficient has the wrong sign, consistent with the OLS regression above, where the price-dividend ratio is the only regressor. This is not very surprising. When the price-dividend ratio moves both due to expected returns

[12]Fama and French (1988) add up dividends throughout the year, which is close to our cash-invested dividend reinvestment strategy. They find that dividend growth rates are positively correlated with past returns. When dividends are reinvested in the market, this induces a negative correlation that more than offsets the positive correlation found by Fama and French (1988).

and expected dividend growth rates, the price-dividend ratio is only a noisy proxy for expected dividend growth rates. Our filtering approach explicitly takes into account that the price-dividend ratio also moves due to expected return variation, which allows us to filter out the relevant expected dividend growth rate information and achieve an R^2 value for dividend growth equal to 31.6%.

4 Hypothesis testing

Our estimates reveal several important properties of expected returns and expected dividend growth rates. In particular, both expected returns and expected growth rates seem to vary over time, expected returns seem to be more persistent than expected growth rates, and both seem to contain a persistent component. In this section, we perform a series of hypothesis tests to establish the statistical significance of these results.

Our likelihood-based estimation approach facilitates a straightforward way to address these questions using the likelihood-ratio (LR) test. Denote the log-likelihood that corresponds to the unconstrained model by $\mathcal{L}^1$. The log-likelihood that follows from estimating the model under the null hypothesis is denoted by $\mathcal{L}^0$. The likelihood-ratio test statistic is given by:

$$LR = 2(\mathcal{L}^1 - \mathcal{L}^0),$$

which is asymptotically chi-squared distributed with the degrees of freedom equal to the number of constrained parameters. We perform our test for both market-reinvested dividend growth rates and cash-invested dividend growth rates.

First, we test for a lack of return predictability. The associated null hypothesis is:

$$H_0 : \delta_1 = \sigma_\mu = \rho_{\mu g} = \rho_{\mu D} = 0,$$

whose LR statistic has a χ^2_4-distribution. Under this null hypothesis, all variation in the log price-dividend ratio comes from variation in expected growth rates. In this case, we can uncover expected dividend growth rates through an OLS regression of dividend growth rates on the lagged price-dividend ratio.

Secondly, we test for the lack of dividend growth rate predictability. The null hypothesis that corresponds to this test for cash-invested dividends reads:

$$H_0 : \gamma_1 = \sigma_g = \rho_{\mu g} = 0,$$

whose LR statistic follows a χ^2_3-distribution. If dividend growth is unpredictable, we can uncover expected returns through an OLS regression of returns on the lagged price-dividend ratio. For market-reinvested dividends, the null hypothesis is:

$$H_0 : \gamma_1 = \sigma_g = \rho_{\mu g} = \sigma_M = \rho_M = 0,$$

and the LR statistic has a χ^2_5-distribution. The absence of dividend growth predictability also requires that σ_M and ρ_M are zero. If not, ε^M_{t+1} correlates with returns and forecasts subsequent dividend growth rates. Under this null hypothesis, all variation in the log price-dividend ratio comes from variation in expected returns.

Third, we test whether the persistence coefficient of expected dividend growth rates equals zero. The null hypothesis that corresponds to this test is:

$$H_0 : \gamma_1 = 0,$$

where the LR statistic has a χ^2_1-distribution. The question of whether expected dividend growth rates are time-varying, and what their persistence is, plays an important role in general equilibrium models with long-run risk (Bansal and Yaron (2004) and Hansen, Heaton, and Li (2008)).

Fourth, we test whether the persistence coefficients of expected dividend growth rates and expected returns are equal, which has been assumed by Cochrane (2007) and Lettau and Van Nieuwerburgh (2006) for analytical convenience. The null hypothesis for this test is:

$$H_0 : \gamma_1 = \delta_1,$$

and the LR statistic has a χ^2_1-distribution. Under the null hypothesis of equal persistence coefficients, the price-dividend ratio is an AR(1)-process, which has been used as a reduced-form model by many authors.[13] Under the alternative hypothesis, the price-dividend ratio is not an AR(1)-process, as the sum of two AR(1)-processes is an ARMA(2,1)-process.

Finally, we test whether the inclusion of ε^M_t adds significantly to the fit of the model. The null hypothesis for this test is:

$$H_0 : \sigma_M = \rho_M = 0,$$

[13] See for example Stambaugh (1999), Lewellen (2004), and Lettau and Van Nieuwerburgh (2006).

and the LR statistic has a χ^2_1-distribution.

We summarize the LR statistics of all these tests in Table 6. The critical values at the 5% and 1% significance levels for the χ^2 with N degrees of freedom are summarized in Table 7. The tables show that all the null hypotheses stated above can be rejected at the 5% level. This suggests, in the context of our model, that both returns and dividend growth rates are predictable. Furthermore, it seems that expected returns are more persistent than expected dividend growth rates, given that (i) we find a lower value of γ_1 than for δ_1 in our unconstrained estimates and (ii) the hypothesis that these two coefficients are equal can be rejected at the 1% level. Finally, the inclusion of the term ε_t^M in our specification for market-invested dividend growth seems to add significantly to the fit of the model. The correlation between returns and ε_t^M is significantly different from 0 and positive, lending further support to our interpretation of ε_M as a reduced-form representation for reinvesting dividends in the market throughout the year.

5 Additional results

5.1 Comparing the filtered series

In this section, we compare the filtered series for both expected returns and expected dividend growth rates for both reinvestment strategies. In Figure 4, we plot the filtered series for μ_t as well as the realized log return when dividends are reinvested in the risk-free rate. We compare it to the fitted return series from an OLS regression of realized log returns on the lagged price-dividend ratio. The figure shows that the two expected return series are almost identical, consistent with the comparable R^2 values we find for both approaches. In Figure 5, we then plot the same series when dividends are reinvested in the market. In this case, the expected return series of our filtering procedure is different from the OLS series. The filtered series is lower in the eighties and higher by the end of the nineties. Consequently, the OLS regression predicts a negative return in the nineties, whereas the filtered series remains positive.[14]

In Figure 6 we plot the filtered series for g_t when dividends are reinvested in the risk-free rate as well as the fitted value from an OLS regression of realized log dividend growth rates (again reinvested in the risk-free rate) on the lagged price-dividend ratio. The difference between the two series is large. The filtered series picks up much more

[14] Campbell and Thompson (2007) suggest to impose the restriction that the equity risk premium is always positive in predictive regressions, which, as they show, enhances the out-of-sample predictability of stock returns.

variation of realized dividend growth than the fitted values from the OLS regression does. Further, it seems that expected dividend growth has a positive autocorrelation, but its persistence is not as high as that of the price-dividend ratio. The price-dividend ratio is mainly driven by expected returns, which are more persistent than expected dividend growth rates, as we formally tested in Section 4.

In Figure 7, we plot the same series, but now for the reinvestment strategy that reinvests dividends in the market. The filtered series pick up a large fraction of the variation of market-reinvested dividend growth rates. This implies that a substantial fraction of market-invested dividend growth is predictable.

5.2 Variance decompositions

We now derive variance decompositions of both the price-dividend ratio and of unexpected returns in both models. The variance decomposition of the price-dividend ratio for cash-invested dividends is given by:

$$\begin{aligned} var(pd_t) &= B_1^2 var(\mu_t) + B_2^2 var(g_t) - 2B_1 B_2 cov(\mu_t, g_t) \\ &= \frac{(B_1 \sigma_\mu)^2}{1-\delta_1^2} + \frac{(B_2 \sigma_g)^2}{1-\gamma_1^2} - \frac{2B_1 B_2 \sigma_{g\mu}}{1-\delta_1 \gamma_1}. \end{aligned} \tag{16}$$

The first term, $B_1^2 var(\mu_t)$, represents the variation of the price-dividend ratio due to discount rate variation. The second term, $B_2^2 var(g_t)$, measures the variation of the price-dividend ratio due to expected dividend growth rate variation. The last term measures the covariation between these two components. For market-invested dividends, the variance decomposition is given by:

$$var\left(pd_t^M\right) = B_1^2 var\left(\mu_t\right) + var\left(B_2 g_t^M + (B_2-1)\varepsilon_t^M\right) + 2cov\left(B_1\mu_t, B_2 g_t^M + (B_2-1)\,\varepsilon_t^M\right).$$

We include the variance due to ε_t^M as part of expected dividend growth variation. This enhances the comparison with cash-invested dividends because we can now also summarize the decomposition using three terms: variation due to discount rates given by $B_1^2 var(\mu_t)$, variation due to expected dividend growth variation given by $var\left(B_2 g_t^M + (B_2-1)\varepsilon_t^M\right)$, and the covariance between these two components. Table 8 summarizes the results, where we use sample covariances and we standardize all terms on the right-hand side of (16) and (17) by the left-hand side, so that the sum of the terms is 100%. We find that for both reinvestment strategies, most variation of the price-dividend ratio is related to expected return variation.

We decompose the variance of unexpected stock returns as in Campbell (1991). In case of cash-invested dividend growth rates, the unexpected return can be written as:

$$r_{t+1} - \mu_t = -\rho B_1 \varepsilon^{\mu}_{t+1} + \rho B_2 \varepsilon^{g}_{t+1} + \varepsilon^{d}_{t+1}. \quad (17)$$

We group the last two terms together to decompose the unexpected return into the influence of discount rates, dividend growth variation, and the covariance between the two. In case of market-invested dividend growth rates, the unexpected return can be written as:[15]

$$\begin{aligned} r_{t+1} - \mu_t &= -\rho B_1 \varepsilon^{\mu}_{t+1} + \rho B_2 \varepsilon^{g}_{t+1} + \varepsilon^{d}_{t+1} + (1-\rho)\varepsilon^{M}_{t+1} \\ &= -\rho B_1 \varepsilon^{\mu}_{t+1} + \rho B_2 \varepsilon^{gM}_{t+1} + \varepsilon^{DM}_{t+1} + (B_2 - 1)\,\varepsilon^{M}_{t}. \end{aligned} \quad (18)$$

As before, we group all the terms after $-\rho B_1 \varepsilon^{\mu}_{t+1}$ together and compute the influence of discount rates, dividend growth rates, and the covariance between these two components. In the results we report below, we use sample covariances and standardize all terms on the right-hand side of equation (17) and (18) by the left-hand side, so that the sum of the terms is 100%.

The variance decomposition of unexpected returns is quite different across reinvestment strategies as summarized in Table 9. This difference is caused by the difference in the correlation between ε^{μ} and ε^{g}, which is higher in the case of market-invested dividends and the difference in the persistence of expected dividend growth rates, γ_1, which is higher in the case of market-invested dividends. Finally, the decomposition of unexpected returns suggests that there is a substantial role for dividend growth variation when explaining unexpected returns.

5.3 Robustness to log-linearizations

In deriving the expression for the log price-dividend ratio in Section 1, we use the approximation to the log total stock return in equation (4). In Binsbergen and Koijen (2008) we study a non-linear present-value model within the class of linearity-inducing models developed by Menzly, Santos, and Veronesi (2004) and generalized by Gabaix (2007).[16] Because the transition equation is non-linear in this model, we use non-linear

[15] The first of these two equations illustrates why the quantitative influence of ε^{M}_{t+1} on unexpected returns is negligible, as it is premultiplied by $1-\rho$ which equals 0.032. In the variance decomposition, the contribution of the term is less than 1%.

[16] We refer to Fernández-Villaverde and Rubio-Ramírez (2004), Fernández-Villaverde, Rubio-Ramírez, and Santos (2006), and Fernández-Villaverde and Rubio-Ramírez (2006) for theoretical results related

filtering techniques to estimate the time series of expected returns. More specifically, we use an unscented Kalman filter (Julier and Uhlmann (1997)) and a particle filter. We find that the main results that we report in this paper are not sensitive to the linearization of log total stock returns. Both expected returns and expected growth rates are persistent processes, but expected returns are more persistent than expected growth rates. Innovations to expected returns and expected growth rates are positively correlated, and we find that the filtered series are good predictors of future returns and dividend growth rates.

5.4 Reinvestment strategy and model specification

We have assumed that the conditional expected dividend growth rate is an AR(1)-process if dividends are reinvested in the risk-free rate. We have subsequently derived the implied dynamics for market-invested dividends. We stress again that there is a present-value model for each reinvestment strategy of dividends, reflected in the time-series properties of expected returns and expected dividend growth rates. We now consider the present-value model in equations (4)-(6) for market-invested dividends instead of cash-invested dividends. That is, we estimate an alternative specification in which expected growth rates of *market-invested* dividends are modeled as an AR(1)-process. The parameter estimates of this model are presented in Table 10. The table shows that the estimated value of γ_1 is not only lower than in the model in which cash-invested expected dividend growth is an AR(1)-process, it is in fact estimated to be *negative*. Despite this negative value for γ_1, we still find relatively high R^2-values for both returns and dividend growth rates.

To further explore this evidence of a negative estimated value for γ_1 in this model, we construct a grid of possible levels of γ_1. For each point in the grid, we optimize over the other parameters and record the associated likelihoods and parameter estimates, shown in Table 11. The main results are summarized in Panel A of Figure 8, where we plot the likelihood as a function of γ_1. The picture shows that the likelihood has two peaks, of which one is positive; the other is negative. Panels B and C show plots of the R^2 values for returns and dividend growth rates as a function of γ_1. The R^2 value for dividend growth rates also exhibits a bi-model shape, and, perhaps surprisingly, the R^2 value is higher for the positive root than for the negative root of γ_1. Furthermore, the R^2 value for returns is also higher for the positive root of γ_1. The pictures therefore illustrate that maximizing R^2 values is not necessarily equivalent to maximizing the likelihood.[17]

to maximum-likelihood estimation in linearized models as well as applications to DSGE models in macroeconomics.

[17] See also Harvey (1989).

5.5 Finite-sample properties

We analyze the finite-sample properties of our maximum-likelihood estimators. We focus on the model for cash-invested dividends in Section 1. We simulate 1,000 samples with the same number of observations as in the data, starting with a draw from the unconditional distribution of the state variables. We use the point estimates of Table 2 in the simulation. We subsequently estimate the model for each of the simulated samples. Table 12 reports the true parameters along with the average, standard deviation, and quantiles of distribution of 1,000 parameter estimates in Panel A. Panel B reports the correlation between the parameter estimates.

Panel A shows that δ_1 is somewhat downward biased, while γ_1 is upward biased. This corresponds to an upward bias in σ_μ and a downward bias in σ_g. Further, it seems that the correlation between expected returns and unexpected growth rates, $\rho_{\mu D}$, is not estimated precisely. Panel B shows that the estimates for the persistence of expected returns (δ_1) and expected growth rates (γ_1) are negatively correlated. Also, we find the persistence parameters and the conditional volatility parameters to be negatively correlated (e.g., δ_1 and σ_μ).

6 Conclusion

We propose a new approach to predictive regressions by assuming that conditional expected returns and conditional expected dividend growth rates are latent, following an exogenously specified ARMA-model. We combine this with a Campbell and Shiller (1988) present-value model to derive the implied dynamics of the price-dividend ratio, and use filtering techniques to uncover estimated series of expected returns and expected dividend growth rates. The filtered series turn out to be good predictors for future returns and for future dividend growth rates.

We find that the reinvestment strategy of dividends that are received within a particular year can have a non-negligible effect on dividend growth rates. For instance, if dividends are reinvested in the aggregate stock market instead of the T-bill rate, the annual volatility of dividend growth is twice as high. We provide a parsimonious model to relate the two reinvestment strategies. It shows, for instance, that if cash-invested expected growth rates are an AR(1)-process, market-invested expected growth rates are an ARMA(1,1)-process.

Our likelihood setup allows for straightforward hypothesis testing using the likelihood-ratio test. We can statistically reject the hypotheses that returns and dividend growth rates are unpredictable or that they are not persistent. Further, we can reject the hypothesis that expected returns and expected dividend growth rates are equally persistent and we find that expected dividend growth rates are less persistent than expected returns.

References

Ang, A., and J. Liu (2004): "How to Discount Cashflows with Time-Varying Expected Returns," *Journal of Finance*, 59, 2745–2783.

Bansal, R., and A. Yaron (2004): "Risks for the Long-Run: A Potential Resolution of Asset Pricing Puzzles," *Journal of Finance*, 59(4), 1481–1509.

Bekaert, G., E. Engstrom, and S. R. Grenadier (2001): "Stock and Bond Pricing in an Affine Economy," NBER Working Paper No. 7346.

——— (2005): "Stock and Bond Returns with Moody Investors," Working Paper, Columbia University.

Binsbergen, J. H. v., and R. S. Koijen (2008): "The Impact of Non-Linearities in Present-Value Models," Unpublished paper, Stanford GSB and Chicago GSB.

Brandt, M. W., and Q. Kang (2004): "On the Relation Between the Conditional Mean and Volatility of Stock Returns: A Latent VAR Approach," *Journal of Financial Economics*, 72, 217–257.

Brennan, M. J., and Y. Xia (2005): "Persistence, Predictability, and Portfolio Planning," Working Paper, UCLA.

Burnside, C. (1998): "Solving Asset Pricing Models with Gaussian Shocks," *Journal of Economic Dynamics and Control*, 22, 329–340.

Campbell, J. Y. (1991): "A Variance Decomposition for Stock Returns," *Economic Journal*, 101, 157–179.

Campbell, J. Y., and J. H. Cochrane (1999): "By Force of Habit: A Consumption-Based Explanation of Aggregate Stock Market Behavior," *Journal of Political Economy*, 107, 205–251.

Campbell, J. Y., and R. J. Shiller (1988): "The Dividend-Price Ratio and Expectations of Future Dividends and Discount Factors," *Review of Financial Studies*, 1, 195–227.

Campbell, J. Y., and S. Thompson (2007): "Predicting Excess Stock Returns Out of Sample: Can Anything Beat the Historical Average?," *Review of Financial Studies,* forthcoming.

Cochrane, J. H. (1991): "Explaining the Variance of Price-Dividend Ratios," *Review of Financial Studies*, 5(2), 243–280.

——— (1994): "Permanent and Transitory Components of GDP and Stock Prices," *Quarterly Journal of Economics*, 109, 241–265.

——— (2007): "The Dog That Did Not Bark: A Defense of Return Predictability," *Review of Financial Studies*, 21(4), 1533–1575.

——— (2008): "State-Space vs. VAR models for Stock Returns," Unpublished paper, Chicago GSB.

Fama, E. F., and K. R. French (1988): "Dividend Yields and Expected Stock Returns," *Journal of Financial Economics*, 22, 3–27.

Fernández-Villaverde, J., and J. Rubio-Ramírez (2004): "Estimating Dynamic Equilibrium Economies: Linear versus Nonlinear Likelihood," *Journal of Applied Econometrics*, 20, 891–910.

——— (2006): "Estimating Macroeconomic Models: A Likelihood Approach," *Review of Economic Studies,* forthcoming.

Fernández-Villaverde, J., J. Rubio-Ramírez, and M. Santos (2006): "Convergence Properties of the Likelihood of Computed Dynamic Models," *Econometrica*, 74, 93–119.

Ferson, W. E., S. Sarkissian, and T. T. Simin (2003): "Spurious regressions in financial economics?," *Journal of Finance*, 58(4), 1393–1413.

Gabaix, X. (2007): "Linearity-Generating Processes: A Modelling Tool Yielding Closed Forms for Asset Prices," Working Paper MIT.

Goetzmann, W. N., and P. Jorion (1995): "A Longer Look at Dividend Yields," *Journal of Business*, 68, 483–508.

Goffe, W. L., G. D. Ferrier, and J. Rogers (1994): "Global Optimization of Statistical Functions with Simulated Annealing," *Journal of Econometrics*, 60, 65–99.

Hamilton, J. D. (1994): *Time Series Analysis.* Princeton University Press, Princeton, NY.

Hansen, L. P., J. C. Heaton, and N. Li (2008): "Consumption Strikes Back? Measuring Long-Run Risk," *Journal of Political Economy*, 116(2), 260–302.

Harvey, A. (1989): *Forecasting, Structural Time Series Models and the Kalman Filter.* Cambridge University Press, Cambridge, UK.

Julier, S., and J. Uhlmann (1997): "A New Extension of the Kalman Filter to Nonlinear Systems," *Proceedings SPIE Signal Processing, Sensor Fusion, and Target Recognition VI*, 3068, 182–193.

Lettau, M., and S. C. Ludvigson (2005): "Expected Returns and Expected Dividend Growth," *Journal of Financial Economics*, 76, 583–626.

Lettau, M., and S. Van Nieuwerburgh (2006): "Reconciling the Return Predictability Evidence," forthcoming Review of Financial Studies.

Lewellen, J. W. (2004): "Predicting Returns with Financial Ratios," *Journal of Financial Economics*, 74, 209–235.

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

Pástor, L., M. Sinha, and B. Swaminathan (2007): "Estimating the Intertemporal Risk-Return Tradeoff Using the Implied Cost of Capital," *Journal of Finance*, forthcoming.

Pástor, L., and R. F. Stambaugh (2006): "Predictive Systems: Living with Imperfect Predictors," Working Paper University of Chicago and CRSP.

Pástor, L., and P. Veronesi (2003): "Stock Valuation and Learning About Profitability," *Journal of Finance*, 58, 1749–1789.

——— (2006): “Was There a Nasdaq Bubble in the Late 1990s?,” *Journal of Financial Economics*, 81, 61–100.

Rytchkov, O. (2007): “Filtering Out Expected Dividends and Expected Returns,” Unpublished paper, Texas A&M.

Stambaugh, R. F. (1999): “Predictive Regressions,” *Journal of Financial Economics*, 54, 375–421.

A Derivations of the present-value model

We consider the model:

$$\begin{aligned}
\Delta d_{t+1} &= g_t + \varepsilon^D_{t+1}, \\
g_{t+1} &= \gamma_0 + \gamma_1 \left(g_t - \gamma_0 \right) + \varepsilon^g_{t+1}, \\
\mu_{t+1} &= \delta_0 + \delta_1 \left(\mu_t - \delta_0 \right) + \varepsilon^\mu_{t+1},
\end{aligned}$$

with:

$$\begin{aligned}
\Delta d_{t+1} &\equiv \log \left(\frac{D_{t+1}}{D_t} \right), \\
\mu_t &\equiv E_t[r_{t+1}], \\
r_{t+1} &\equiv \log \left(\frac{P_{t+1} + D_{t+1}}{P_t} \right).
\end{aligned}$$

We also define: $pd_t = \log(PD_t)$. Now consider the log-linearized return, with $\overline{pd} = E\left[pd_t\right]$:

$$\begin{aligned}
r_{t+1} &= \log\left(1 + \exp\left(pd_{t+1}\right)\right) + \Delta d_{t+1} - pd_t \\
&\simeq \log\left(1 + \exp\left(\overline{pd}\right)\right) + \frac{\exp\left(\overline{pd}\right)}{1 + \exp\left(\overline{pd}\right)} pd_{t+1} + \Delta d_{t+1} - pd_t \\
&= \kappa + \rho pd_{t+1} + \Delta d_{t+1} - pd_t,
\end{aligned}$$

or, equivalently, we have:

$$pd_t = \kappa + \rho pd_{t+1} + \Delta d_{t+1} - r_{t+1},$$

in which:

$$\begin{aligned}
\kappa &= \log\left(1 + \exp\left(\overline{pd}\right)\right) - \rho \overline{pd}, \\
\rho &= \frac{\exp\left(\overline{pd}\right)}{1 + \exp\left(\overline{pd}\right)}.
\end{aligned}$$

By iterating this equation we find:

$$\begin{aligned}
pd_t &= \kappa + \rho pd_{t+1} + \Delta d_{t+1} - r_{t+1} \\
&= \kappa + \rho \left(\kappa + \rho pd_{t+2} + \Delta d_{t+2} - r_{t+2} \right) + \Delta d_{t+1} - r_{t+1} \\
&= \kappa + \rho \kappa + \rho^2 pd_{t+2} + \Delta d_{t+1} - r_{t+1} + \rho \left(\Delta d_{t+2} - r_{t+2} \right) \\
&= \sum_{j=0}^{\infty} \rho^j \kappa + \rho^\infty pd_\infty + \sum_{j=1}^{\infty} \rho^{j-1} \left(\Delta d_{t+j} - r_{t+j} \right) \\
&= \frac{\kappa}{1 - \rho} + \sum_{j=1}^{\infty} \rho^{j-1} \left(\Delta d_{t+j} - r_{t+j} \right),
\end{aligned}$$

assuming that $\rho^{\infty} pd_{\infty} = \lim_{j\to\infty} \rho^j pd_{t+j} = 0$ (in expectation would suffice for our purpose). Next, we take expectations conditional upon time-t:

$$\begin{aligned}
pd_t &= \frac{\kappa}{1-\rho} + \sum_{j=1}^{\infty} \rho^{j-1} E_t \left[\Delta d_{t+j} - r_{t+j}\right] \\
&= \frac{\kappa}{1-\rho} + \sum_{j=1}^{\infty} \rho^{j-1} E_t \left[g_{t+j-1} - \mu_{t+j-1}\right] \\
&= \frac{\kappa}{1-\rho} + \sum_{j=0}^{\infty} \rho^{j} E_t \left[g_{t+j} - \mu_{t+j}\right], \\
&= \frac{\kappa}{1-\rho} + \sum_{j=0}^{\infty} \rho^{j} \left(\gamma_0 + \gamma_1^j (g_t - \gamma_0) - \delta_0 - \delta_1^j (\mu_t - \delta_0)\right) \\
&= \frac{\kappa}{1-\rho} + \frac{\gamma_0 - \delta_0}{1-\rho} + \sum_{j=0}^{\infty} \rho^{j} \left(\gamma_1^j (g_t - \gamma_0) - \delta_1^j (\mu_t - \delta_0)\right) \\
&= \frac{\kappa}{1-\rho} + \frac{\gamma_0 - \delta_0}{1-\rho} + \frac{g_t - \gamma_0}{1-\rho\gamma_1} - \frac{\mu_t - \delta_0}{1-\rho\delta_1},
\end{aligned}$$

which uses:

$$E_t \left[x_{t+j}\right] = \alpha_0 + \alpha_1^j (x_t - \alpha_0),$$

provided that:

$$x_{t+1} = \alpha_0 + \alpha_1 (x_t - \alpha_0) + \varepsilon_{t+1}.$$

B Kalman filter

In this section we discuss the Kalman filtering procedure of our model. We discuss the most general case in which dividends are reinvested in the market. The other models that we discuss are special cases of this general setup.

We first reformulate the model in standard state-space form. Define an expanded state vector:

$$X_t = \begin{bmatrix} \hat{g}_{t-1} \\ \varepsilon_t^D \\ \varepsilon_t^g \\ \varepsilon_t^\mu \\ \varepsilon_t^M \\ \varepsilon_{t-1}^M \end{bmatrix},$$

which satisfies:

$$X_{t+1} = F X_t + \Gamma \varepsilon_{t+1}^X,$$

with

$$
\begin{aligned}
F &= \begin{bmatrix} \gamma_1 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 \end{bmatrix}, \\
\Gamma &= \begin{bmatrix} 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \end{bmatrix},
\end{aligned}
$$

and where

$$
\varepsilon_{t+1}^X = \begin{bmatrix} \varepsilon_t^D \\ \varepsilon_t^g \\ \varepsilon_t^\mu \\ \varepsilon_t^M \end{bmatrix}.
$$

which we assume to be jointly normally distributed.

The measurement equation, which has the observables $Y_t = (\Delta d_t, pd_t)$, is:

$$
Y_t = M_0 + M_1 Y_{t-1} + M_2 X_t,
$$

with

$$
\begin{aligned}
M_0 &= \begin{bmatrix} \gamma_0 \\ (1 - \delta_1)\, C \end{bmatrix}, \\
M_1 &= \begin{bmatrix} 0 & 0 \\ 0 & \delta_1 \end{bmatrix}, \\
M_2 &= \begin{bmatrix} 1 & 1 & 0 & 0 & 1 & -1 \\ B_2(\gamma_1 - \delta_1) & 0 & B_2 & -B_1 & -1 & \delta_1 \end{bmatrix}.
\end{aligned}
$$

The Kalman procedure is given by:

$$
\begin{aligned}
X_{0|0} &= E\left[X_0\right] = 0_{6\times 1}, \\
P_{0|0} &= E\left[X_t X_t'\right], \\
X_{t|t-1} &= F X_{t-1|t-1}, \\
P_{t|t-1} &= F P_{t-1|t-1} F' + \Gamma \Sigma \Gamma', \\
\eta_t &= Y_t - M_0 - M_1 Y_{t-1} - M_2 X_{t|t-1}, \\
S_t &= M_2 P_{t|t-1} M_2', \\
K_t &= P_{t|t-1} M_2' S_t^{-1}, \\
X_{t|t} &= X_{t|t-1} + K_t \eta_t, \\
P_{t|t} &= \left(I - K_t M_2\right) P_{t|t-1}.
\end{aligned}
$$

The likelihood is based on prediction errors (η_t) and their covariance matrix ($P_{t|t-1}$), which is subject to change in every iteration:

$$L = -\sum_{t=1}^{T} \log(\det(S_t)) - \sum_{t=1}^{T} \eta_t' S_t^{-1} \eta_t.$$

Finally, the covariance matrix of the shocks is:

$$\Sigma \equiv var\left(\begin{bmatrix} \varepsilon_{t+1}^{g} \\ \varepsilon_{t+1}^{\mu} \\ \varepsilon_{t+1}^{d} \\ \varepsilon_{t+1}^{M} \end{bmatrix}\right) = \begin{bmatrix} \sigma_g^2 & \sigma_{g\mu} & \sigma_{gD} & \sigma_{gM} \\ \sigma_{g\mu} & \sigma_\mu^2 & \sigma_{\mu D} & \sigma_{\mu M} \\ \sigma_{gD} & \sigma_{\mu D} & \sigma_D^2 & \sigma_{DM} \\ \sigma_{gM} & \sigma_{\mu M} & \sigma_{DM} & \sigma_M^2 \end{bmatrix}.$$

Recall that we have assumed that:

$$\varepsilon_{t+1}^{M} = \beta_M \varepsilon_{t+1}^{r} + \varepsilon_{t+1}^{M\perp},$$

in which $\beta_M = \rho_M \sigma_M / \sigma_r$ and $\sigma_r = \sqrt{var(\varepsilon_{t+1}^{r})}$ and:

$$\varepsilon_{t+1}^{r} \equiv r_{t+1} - \mu_t \approx -B_1 \rho \varepsilon_{t+1}^{\mu} + B_2 \rho \varepsilon_{t+1}^{g} + \varepsilon_{t+1}^{d}.$$

Therefore,

$$
\begin{aligned}
\sigma_r^2 &= \sigma_D^2 + \rho^2 B_1^2 \sigma_\mu^2 + \rho^2 B_2^2 \sigma_g^2 - 2\rho B_1 \sigma_{\mu D} - 2\rho^2 B_1 B_2 \sigma_{\mu g}, \\
\sigma_{gM} &= -\beta_M \rho B_1 \sigma_{g\mu} + \beta_M \rho B_2 \sigma_g^2, \\
\sigma_{\mu M} &= \beta_M \sigma_{\mu D} - \beta_M \rho B_1 \sigma_\mu^2 + \beta_M B_2 \rho \sigma_{\mu g}, \\
\sigma_{DM} &= \beta_M \sigma_D^2 - \beta_M \rho B_1 \sigma_{\mu D}.
\end{aligned}
$$

We subsequently maximize the likelihood over the parameters:

$$\Theta \equiv \left(\gamma_0, \delta_0, \gamma_1, \delta_1, \sigma_g, \sigma_\mu, \sigma_D, \rho_{g\mu}, \rho_{gD}, \rho_{\mu D}, \sigma_M, \rho_M\right).$$

C Wold decomposition

Using the Kalman filter in Appendix B in stationary state ($K_t = K$), we can express the filtered state in terms of historical growth rates and price-dividend ratios:

$$\begin{aligned}
X_{t|t} &= X_{t|t-1} + K\left(Y_t - M_0 - M_1 Y_{t-1} - M_2 X_{t|t-1}\right) \\
&= \left(I - KM_2\right) X_{t|t-1} + K\left(Y_t - M_0 - M_1 Y_{t-1}\right) \\
&= \left(I - KM_2\right) F X_{t-1|t-1} + K\left(Y_t - M_0 - M_1 Y_{t-1}\right) \\
&= \ldots \\
&= \sum_{i=0}^{\infty} \left[\left(I - KM_2\right) F\right]^i K\left(Y_{t-i} - M_0 - M_1 Y_{t-1-i}\right).
\end{aligned}$$

Using the return definition:

$$r_{t+1} = \kappa + \rho pd_{t+1} + \Delta d_{t+1} - pd_t,$$

this also implies a representation in terms of price-dividend ratios and returns. The first element of X_t is $\hat{g}_{t-1} = g_{t-1} - \gamma_0$. Hence, it is more natural to think about $X_{t|t-1}$:

$$\begin{aligned}
X_{t|t-1} &= F X_{t-1|t-1} \\
&= F \sum_{i=0}^{\infty} \left[\left(I - KM_2\right) F\right]^i K\left(Y_{t-1-i} - M_0 - M_1 Y_{t-2-i}\right),
\end{aligned}$$

implying that the first element of $X_{t|t-1}$ equals $\hat{g}_{t-1|t-1}$, the filtered value of expected growth rates up to time $t-1$. Using the expression for the log price-dividend ratio, we obtain a similar representation for the filtered value of expected returns, $\hat{\mu}_{t-1|t-1}$:

$$\hat{\mu}_{t-1|t-1} = B_1^{-1}\left(pd_{t-1} - A - B_2 \hat{g}_{t-1|t-1}\right).$$

This represents expected returns and expected growth rates as a function of lagged growth rates and price-dividend ratios. We define $\varepsilon_t^{d*} \equiv \Delta d_t - \gamma_0 - g_{t-1|t-1}$, and obtain:

$$\begin{aligned}
\Delta d_t &= \gamma_0 + e_1' X_{t|t-1} + \varepsilon_t^{d*} \\
&= \gamma_0 + e_1' F \sum_{i=0}^{\infty} \left[\left(I - KM_2\right) F\right]^i K\left(Y_{t-1-i} - M_0 - M_1 Y_{t-2-i}\right) + \varepsilon_t^{d*} \\
&= a_0^d + \sum_{i=0}^{\infty} a_{1i}^d pd_{t-i-1} + \sum_{i=0}^{\infty} a_{2i}^d \Delta d_{t-i-1} + \varepsilon_t^{d*},
\end{aligned}$$

with:

$$
\begin{aligned}
a_0^d &= \gamma_0 - e_1' F \sum_{i=0}^{\infty} \left[(I - KM_2) F \right]^i KM_0, \\
a_{1i}^d &= e_1' FKe_1, \text{ if } i = 0 \\
&= e_1' F \left[(I - KM_2) F \right]^{i-1} \left((I - KM_2) FK - KM_1 \right) e_1, \text{ if } i \neq 0, \\
a_{2i}^d &= e_1' FKe_2, \text{ if } i = 0 \\
&= e_1' F \left[(I - KM_2) F \right]^{i-1} \left((I - KM_2) FK - KM_1 \right) e_2, \text{ if } i \neq 0.
\end{aligned}
$$

For returns, we have:

$$
r_t = a_0^r + \sum_{i=0}^{\infty} a_{1i}^r pd_{t-i-1} + \sum_{i=0}^{\infty} a_{2i}^r \Delta d_{t-i-1} + \varepsilon_t^{r*},
$$

with:

$$
\begin{aligned}
a_0^r &= a_0^d - B_1^{-1} A, \\
a_{1i}^r &= -\frac{B_2}{B_1} a_{1i}^d + \frac{1}{B_1}, \text{ if } i = 0, \\
&= -\frac{B_2}{B_1} a_{1i}^d, \text{ if } i \neq 0, \\
a_{2i}^r &= -\frac{B_2}{B_1} a_{2i}^d.
\end{aligned}
$$

	Δd_t^M	Δd_t
Mean	0.0586	0.0611
Median	0.0558	0.0540
Standard Deviation	0.1232	0.0622
Maximum	0.3699	0.2616
Minimum	-0.2912	-0.0579

Table 1: Summary statistics of dividend growth rates.
The table shows summary statistics for both market-invested and cash-invested dividend growth rates using data from 1946-2007.

Panel A: Maximum-likelihood estimates					
	Estimate	S.e.		Estimate	S.e.
δ_0	0.090	(0.020)	γ_0	0.062	(0.011)
δ_1	0.932	(0.128)	γ_1	0.354	(0.271)
σ_μ	0.016	(0.013)	σ_g	0.058	(0.017)
$\rho_{D\mu}$	-0.147	(0.579)	σ_D	0.002	(0.022)
$\rho_{\mu g}$	0.417	(0.375)			
Panel B: Implied present-value model parameters					
A	3.571	(0.421)	ρ	0.969	
B_1	10.334	(4.088)	B_2	1.523	(2.001)
Panel C: R-squared values					
$R^2_{Returns}$	8.2%		R^2_{Div}	13.9%	

Table 2: Estimation results: cash-invested dividends.
We present the estimation results of the present-value model in equations (4)-(6). The model is estimated by conditional maximum likelihood using data from 1946 to 2007 on cash-invested dividend growth rates and the corresponding price-dividend ratio. Panel A presents the estimates of the coefficients of the underlying processes (bootstrapped standard errors between parentheses). Panel B reports the resulting coefficients of the present-value model ($pd_t = A - B_1 (\mu_t - \delta_0) + B_2 (g_t - \gamma_0)$). The constants A, B_1 and B_2 are non-linear transformations of the underlying present-value parameters. Therefore, when interpreting the standard errors, it should be taken into account that the distribution of these constants is not symmetric. In Panel C we report the R-squared values for returns and dividend growth rates.

	Panel A: Maximum-likelihood estimates				
	Estimate	S.e.		Estimate	S.e.
δ_0	0.086	(0.039)	γ_0	0.060	(0.014)
δ_1	0.957	(0.055)	γ_1	0.638	(0.170)
σ_μ	0.016	(0.012)	σ_g^M	0.077	(0.015)
$\rho_{D\mu}^M$	-0.344	(0.171)	σ_D^M	0.089	(0.011)
$\rho_{\mu g}^M$	0.805	(0.078)	σ_M	0.054	(0.016)
ρ_M	0.586	(0.191)			
	Panel B: Implied present-value model parameters				
A	3.612	(0.953)	ρ	0.968	
B_1	13.484	(5.626)	B_2	2.616	(2.723)
	Panel C: R-squared values				
$R^2_{Returns}$	8.9%		R^2_{Div}	31.6%	

Table 3: Estimation results: market-invested dividends.
We present the estimation results of the present-value model in equations (11)-(13). The model is estimated by conditional maximum likelihood using data from 1946 to 2007 on market-invested dividend growth rates and the corresponding price-dividend ratio. Panel A presents the estimates of the coefficients of the underlying processes (bootstrapped standard errors in parentheses). Panel B reports the resulting coefficients of the present-value model ($pd_t = A - B_1(\mu_t - \delta_0) + B_2(g_t - \gamma_0)$). The constants A, B_1 and B_2 are non-linear transformations of the underlying present-value parameters. Therefore, when interpreting the standard errors, it should be taken into account that the distribution of these constants is not symmetric. In Panel C we report the R-squared values for returns and dividend growth rates.

Dependent Variable	r_t^M		r_t		Δd_t^M	Δd_t
constant	0.4539	***	0.4555	***	0.1814	0.1085
	(0.1537)		(0.1524)		(0.1266)	(0.0645)
pd_{t-1}^M	-0.1023	**	-		-0.0361	-
	(0.0449)		-		(0.0370)	-
pd_{t-1}	-		-0.1020	**	-	-0.0138
	-		(0.0441)		-	(0.0186)
R^2	7.96%		8.20%		1.56%	0.90%
Adj. R^2	6.43%		6.67%		-0.07%	-0.75%

Table 4: OLS predictive regressions.
The table reports the OLS regression results of log returns and log dividend growth rates on the lagged log price-dividend ratio using data between 1946 and 2007. The second and fourth column report the results using dividends that are reinvested in the aggregate stock market, whereas the third and the fifth column report the results using cash-invested dividends. Two asterisks (**) indicates significance at the 5% level, and three asterisks indicates significance at the 1% level.

Dependent Variable: Market-invested Dividend Growth (Δd_t^M)								
constant	**0.0604**	***	**0.0605**	***	**0.1007**	***	**0.0872**	***
	(0.0138)		(0.0127)		(0.0170)		(0.0130)	
AR(1)	**-0.5261**		**-0.2214**	*	-		**0.7817**	***
	(0.3506)		(0.1255)		-		(0.1148)	
MA(1)	**0.3652**		**-**		**-**		**-0.9793**	***
	(0.3937)		-		-		(0.0320)	
r_{t-1}	**-**		**-**		**-0.3717**	***	**-0.3125**	***
	-		-		(0.0904)		(0.1162)	
R^2	6.34%		5.01%		22.27%		27.46%	
Adj. R^2	3.11%		3.40%		20.95%		23.58%	

Table 5: OLS predictive regressions.
The table reports the results for several reduced-form specifications of realized log dividend growth estimated with OLS using data between 1946 and 2007. Dividends are reinvested in the aggregate stock market. One asterisk (*) denotes significance at the 10% level, two asterisks indicates significance at the 5% level, and three asterisks indicates significance at the 1% level.

	LR	Sign	Log Lik. H_0	Log Lik. H_a	δ_0	δ_1	γ_0	γ_1	σ_μ	σ_g	σ_D	$\rho_{g\mu}$	$\rho_{\mu D}$	σ_M	ρ_M
					Parameters under H_0										
Test for Lack of Return Predictability															
Cash reinv. dividends	**28.67**	***	7.0593	7.5218	0.0936	0	0.0637	0.9900	0	0.0065	0.0659	0	0	-	-
Market reinv. dividends	**22.37**	***	6.4773	6.8381	0.0926	0	0.0666	0.9936	0	0.0057	0.0780	0	0	0.0607	0.8521
Test for Lack of Div. Growth Predictability															
Cash reinv. dividends	**9.23**	**	7.3730	7.5218	0.0882	0.9261	0.0607	0	0.0164	0	0.0617	0	0.3494	-	-
Market reinv. dividends	**29.59**	***	6.3609	6.8381	0.0833	0.9514	0.0587	0	0.0104	0	0.1222	0	0.2973	0	0
Test for Lack of Persistence in Expected Div. Growth															
Cash reinv. dividends	**8.26**	***	7.3886	7.5218	0.0882	0.9288	0.0610	0	0.0156	0.0605	0.0121	0.2550	0.2636	-	-
Market reinv. dividends	**5.89**	**	6.7431	6.8381	0.0852	0.9262	0.0584	0	0.0174	0.0619	0.0470	0.7449	-0.2207	0.0501	0.6792
Test whether g_t and μ_t are Equally Persistent															
Cash reinv. dividends	**8.60**	***	7.3831	7.5218	0.0867	0.9437	0.0595	0.9437	0.0157	0.0022	0.0617	0.9493	0.3090	-	-
Market reinv. dividends	**5.10**	**	6.7558	6.8381	0.0782	0.9478	0.0548	0.9478	0.0166	0.0033	0.0764	0.9351	0.3541	0.0631	0.9254
Test for exclusion of ε_M															
Market reinv. dividends	**11.00**	***	6.6607	6.8381	0.0854	0.9324	0.0591	-0.3253	0.0149	0.0939	0.0635	0.9064	-0.4212	0	0
Test $\rho_M = 0$															
Market reinv. dividends	**6.93**	***	6.7264	6.8381	0.0853	0.9321	0.0584	0.4419	0.0209	0.0595	0.0633	0.9945	-0.1048	0.0479	0

Table 6: Likelihood-ratio tests.
We report the LR statistics for the tests described in Section 4. We do the first four tests for the two specifications that we explore in this paper. "Cash" refers to the system in equations (4)-(6) using the data where dividends are reinvested in the risk free rate. "Market" refers to the system in equations (11)-(13) using the data where dividends are reinvested in the aggregate stock market. Two asterisks (**) denotes that we reject the hypothesis at the 5% level and 3 asterisks (***) indicates that we reject the hypothesis at the 1% level.

Degrees of freedom (N)	1	2	3	4	5
$\chi^2_{N,0.05}$	3.841	5.991	7.815	9.488	11.070
$\chi^2_{N,0.01}$	6.635	9.210	11.345	13.277	15.086

Table 7: Critical values of the likelihood-ratio tests.

Reinvestment strategy	Discount rates	Div. Growth	Covariance
Cash	104.6%	4.6%	-9.2%
Market	117.9%	4.9%	-22.8%

Table 8: Variance decomposition of the price-dividend ratio.

Reinvestment strategy	Discount Rates	Div. Growth	Covariance
Cash	118.4%	34.6%	-53.0%
Market	215.3%	49.4%	-164.7%

Table 9: Variance decomposition of unexpected returns.

Panel A: Maximum-likelihood estimates

	Estimate	S.e.		Estimate	S.e.
δ_0	0.085	(0.019)	γ_0	0.059	(0.012)
δ_1	0.933	(0.148)	γ_1	-0.324	(0.282)
σ_μ	0.015	(0.014)	σ_g	0.094	(0.026)
$\rho_{D\mu}$	-0.422	(0.276)	σ_D	0.065	(0.022)
$\rho_{\mu g}$	0.905	(0.076)			

Panel B: Implied present-value model parameters

A	3.596	(0.349)	ρ	0.968	
B_1	10.263	(3.439)	B_2	0.761	(2.883)

Panel C: R-squared values

$R^2_{Returns}$	8.6%	R^2_{Div}	18.7%

Table 10: Estimation results of the model in (4)-(6) using market-invested dividends. We present the estimation results of the present-value model in equations (4)-(6) using market-reinvested dividend data. The model is estimated by conditional maximum likelihood using data from 1946 to 2007 on the dividend growth rate and the price-dividend ratio. Panel A presents the estimates of the coefficients of the underlying processes (bootstrapped standard errors in parentheses). Panel B reports the resulting coefficients of the present-value model ($pd_t = A - B_1(\mu_t - \delta_0) + B_2(g_t - \gamma_0)$). These parameters are non-linear transformations of the original present-value parameters. When interpreting the standard errors, it should be taken into account that the distribution of the coefficients is not symmetric. In Panel C we report the R-squared values for returns and dividend growth rates.

	$\gamma_1 > 0$	γ_1																	
		-0.9	-0.8	-0.7	-0.6	-0.5	-0.4	-0.3	-0.2	-0.1	0	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8
δ_0	0.084	0.083	0.083	0.085	0.083	0.085	0.085	0.085	0.086	0.087	0.086	0.087	0.085	0.086	0.085	0.083	0.084	0.083	0.077
δ_1	0.933	0.957	0.954	0.952	0.949	0.9441	0.936	0.931	0.926	0.921	0.918	0.920	0.922	0.924	0.926	0.933	0.944	0.957	0.976
γ_0	0.058	0.059	0.059	0.060	0.059	0.060	0.059	0.059	0.059	0.059	0.058	0.059	0.058	0.058	0.058	0.057	0.058	0.059	0.059
γ_1	0.472	-0.900	-0.800	-0.700	-0.600	-0.500	-0.400	-0.300	-0.200	-0.100	0.000	0.100	0.200	0.300	0.400	0.500	0.600	0.700	0.800
σ_μ	0.022	0.010	0.010	0.011	0.011	0.012	0.014	0.015	0.017	0.019	0.020	0.021	0.021	0.022	0.023	0.022	0.022	0.020	0.017
σ_g	0.053	0.029	0.047	0.062	0.076	0.084	0.089	0.095	0.101	0.105	0.051	0.051	0.053	0.053	0.053	0.052	0.050	0.045	0.040
σ_D	0.107	0.109	0.101	0.092	0.082	0.075	0.069	0.062	0.057	0.053	0.109	0.109	0.108	0.106	0.106	0.106	0.107	0.108	0.109
$\rho_{g\mu}$	0.978	0.857	0.887	0.915	0.944	0.928	0.915	0.902	0.897	0.897	0.953	0.968	0.971	0.974	0.977	0.980	0.982	0.986	0.990
$\rho_{\mu D}$	0.208	0.093	-0.030	-0.158	-0.314	-0.369	-0.403	-0.431	-0.442	-0.442	0.255	0.249	0.237	0.225	0.213	0.200	0.187	0.169	0.142
R^2_{Ret}	0.105	0.075	0.077	0.078	0.079	0.081	0.084	0.086	0.088	0.090	0.092	0.094	0.097	0.100	0.103	0.106	0.106	0.104	0.095
R^2_{Div}	0.242	0.036	0.075	0.105	0.130	0.155	0.176	0.191	0.200	0.200	0.191	0.204	0.217	0.229	0.237	0.240	0.236	0.221	0.193
Log L	6.617	6.466	6.521	6.568	6.609	6.640	6.657	6.660	6.648	6.621	6.578	6.590	6.600	6.609	6.616	6.617	6.611	6.590	6.547

Table 11: Estimating a model with an AR(1)-process for expected growth rates in case of market-invested dividends. In the column "$\gamma_1 > 0$" we report the maximum-likelihood estimates of equations (4)-(6), but using dividends that are reinvested in the market. In the first column, we impose that the persistence coefficient of expected dividend growth rates is positive. We then define a grid for γ_1 between -0.9 and 0.8 with increments of 0.1, and compute for each of these values of γ_1 the likelihood while optimizing over all the other parameters.

Panel A: Mean, standard deviation, and quantiles

	True	Average	St.dev.	Q(0.10)	Q(0.25)	Q(0.50)	Q(0.75)	Q(0.90)
δ_0	0.090	0.090	0.020	0.067	0.077	0.089	0.101	0.113
δ_1	0.932	0.864	0.128	0.765	0.837	0.887	0.926	0.952
γ_0	0.062	0.061	0.011	0.047	0.054	0.061	0.069	0.076
γ_1	0.354	0.429	0.271	0.218	0.304	0.417	0.565	0.764
σ_μ	0.016	0.025	0.013	0.012	0.016	0.022	0.030	0.041
σ_g	0.058	0.045	0.017	0.017	0.036	0.052	0.057	0.061
σ_D	0.002	0.022	0.019	0.003	0.006	0.014	0.040	0.051
$\rho_{g\mu}$	0.417	0.318	0.375	-0.009	0.254	0.403	0.516	0.605
$\rho_{\mu D}$	-0.147	0.176	0.579	-0.808	-0.180	0.298	0.640	0.860
A	3.612	3.546	0.421	3.135	3.345	3.551	3.771	3.979
B_1	13.484	8.009	4.088	3.870	5.288	7.116	9.709	12.891
B_2	2.616	2.281	2.001	1.268	1.418	1.678	2.212	3.855

Panel B: Correlation matrix

	δ_0	δ_1	γ_0	γ_1	σ_μ	σ_g	σ_D	$\rho_{g\mu}$	$\rho_{\mu D}$
δ_0	1.000	0.008	0.783	0.063	-0.021	0.011	-0.022	-0.021	-0.015
δ_1	0.008	1.000	0.007	-0.175	-0.686	0.235	-0.189	0.061	-0.149
γ_0	0.783	0.007	1.000	0.067	-0.024	0.012	-0.029	-0.034	-0.034
γ_1	0.063	-0.175	0.067	1.000	0.107	-0.280	0.337	0.152	0.072
σ_μ	-0.021	-0.686	-0.024	0.107	1.000	-0.105	0.104	0.105	0.233
σ_g	0.011	0.235	0.012	-0.280	-0.105	1.000	-0.885	0.496	-0.167
σ_D	-0.022	-0.189	-0.029	0.337	0.104	-0.885	1.000	-0.402	0.194
$\rho_{g\mu}$	-0.021	0.061	-0.034	0.152	0.105	0.496	-0.402	1.000	-0.072
$\rho_{\mu D}$	-0.015	-0.149	-0.034	0.072	0.233	-0.167	0.194	-0.072	1.000

Table 12: Finite-sample properties of the maximum-likelihood estimators.
The table contains results about the finite-sample properties of our maximum-likelihood estimators. We focus on the model for cash-invested dividends in Section 1. We simulate 1,000 samples with the same number of observations as in the data, starting with a draw from the unconditional distribution of the state variables. We use the point estimates of Table 2 in the simulation. We subsequently estimate the model for each of the simulated samples. Panel A reports the true parameters along with the average, standard deviation, and quantiles of distribution of 1,000 parameter estimates. Panel B depicts the correlation between the parameter estimates.

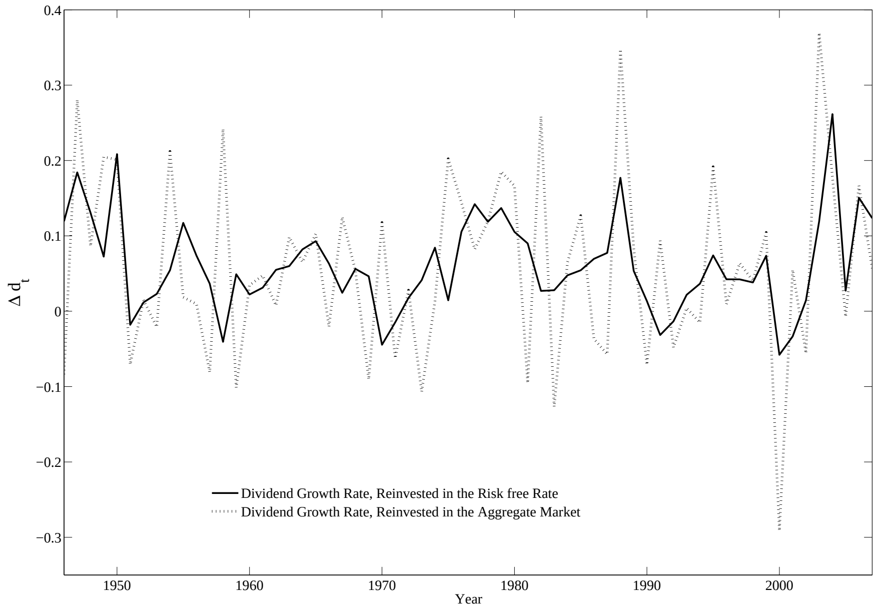

Figure 1: Dividend-growth rates: Reinvesting in either the risk-free rate or in the market. The graph plots the log dividend growth rate for two dividend reinvestment strategies: reinvesting in the risk-free rate and reinvesting in the market.

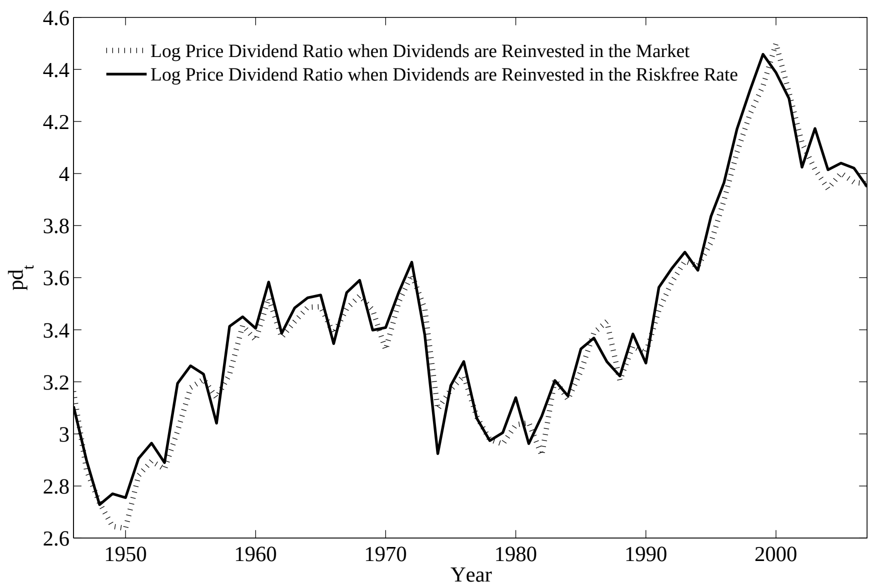

Figure 2: Price-dividend ratio: Reinvesting in either the risk-free rate or in the market. The graph plots the log price dividend ratio for two dividend reinvestment strategies: reinvesting in the risk free rate and reinvesting in the market.

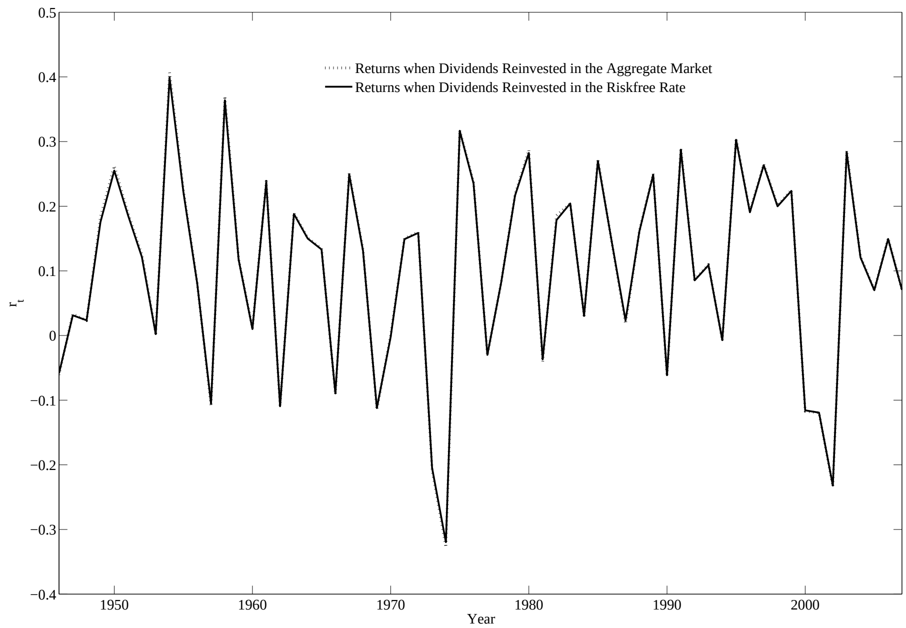

Figure 3: Cum-dividend returns: Reinvesting in either the risk-free rate or in the market. The graph plots the log cum dividend return for two dividend reinvestment strategies: reinvesting in the risk free rate (r_t) and reinvesting in the market (r_t^M).

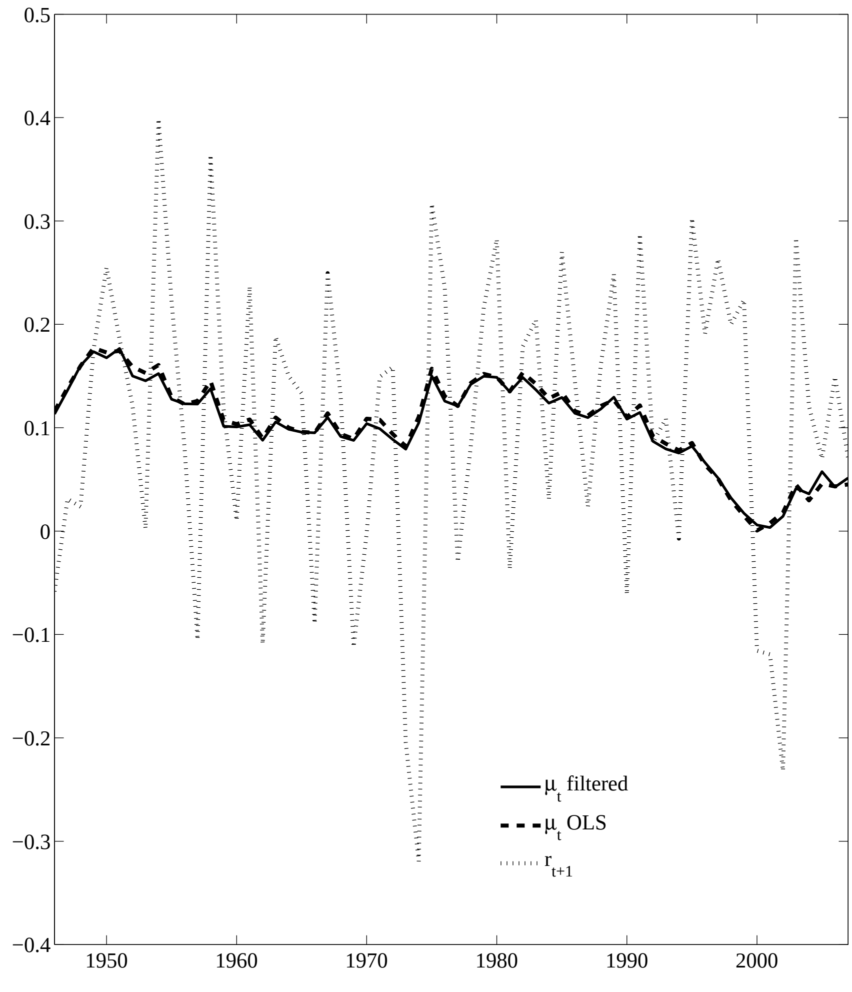

Figure 4: Filtered series for expected returns for reinvesting in the risk-free rate. The graph plots the filtered series of expected returns (μ_t) when dividends are reinvested in the risk-free rate. The graph also plots the realized return r_{t+1} as well as the expected return obtained from an OLS regression of r_{t+1} on the lagged price-dividend ratio.

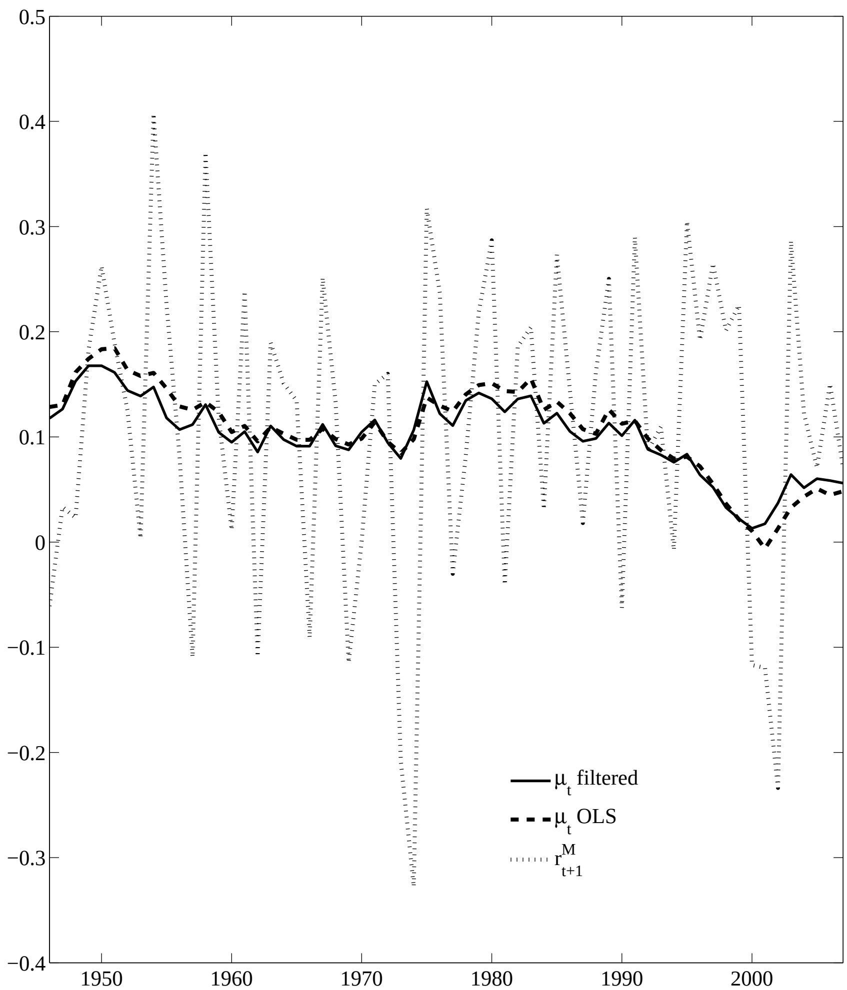

Figure 5: Filtered series for expected returns for reinvesting in the market.
The graph plots the filtered series of expected returns (μ_t) when dividends are reinvested in the market. The graph also plots the realized return r_{t+1}^{M} (again when dividends are reinvested in the market) as well as the expected return obtained from an OLS regression of r_{t+1}^{M} on the lagged price-dividend ratio.

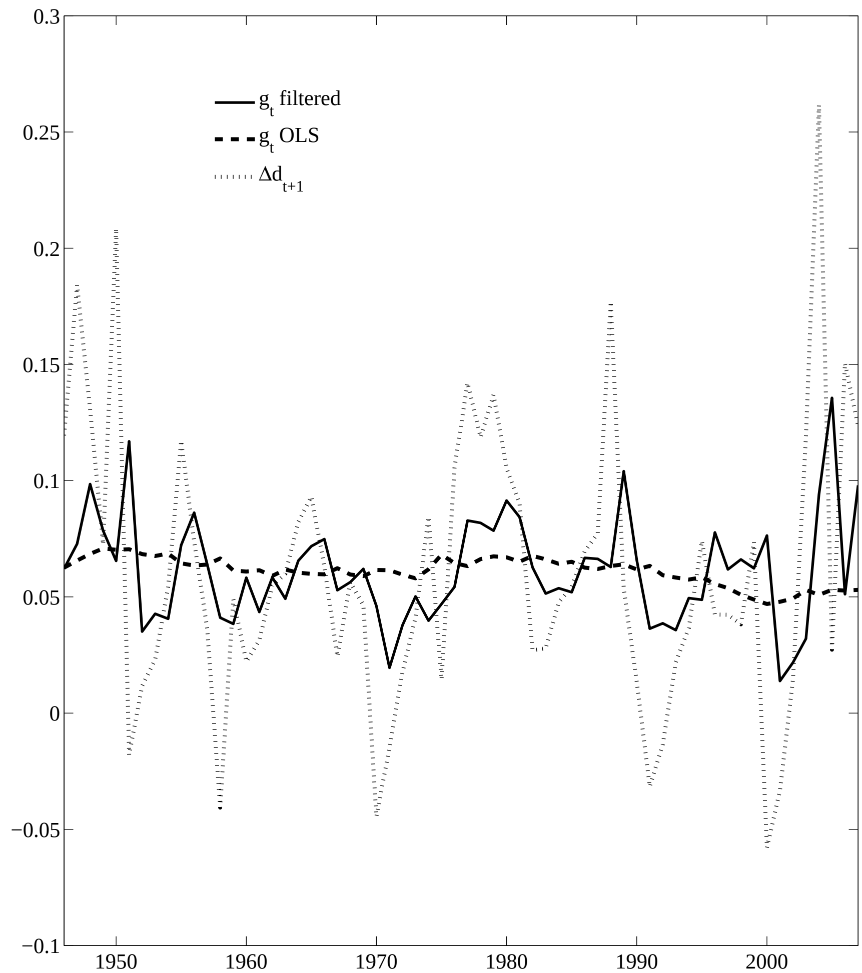

Figure 6: Filtered series for expected dividend growth for reinvesting in the risk-free rate. The graph plots the filtered series of expected dividend growth (g_t) when dividends are reinvested in the risk-free rate. The graph also plots the realized dividend growth Δd_{t+1} (again when dividends are reinvested in the risk-free rate) as well as the expected dividend growth rate obtained from an OLS regression of realized dividend growth Δd_{t+1} on the lagged price-dividend ratio.

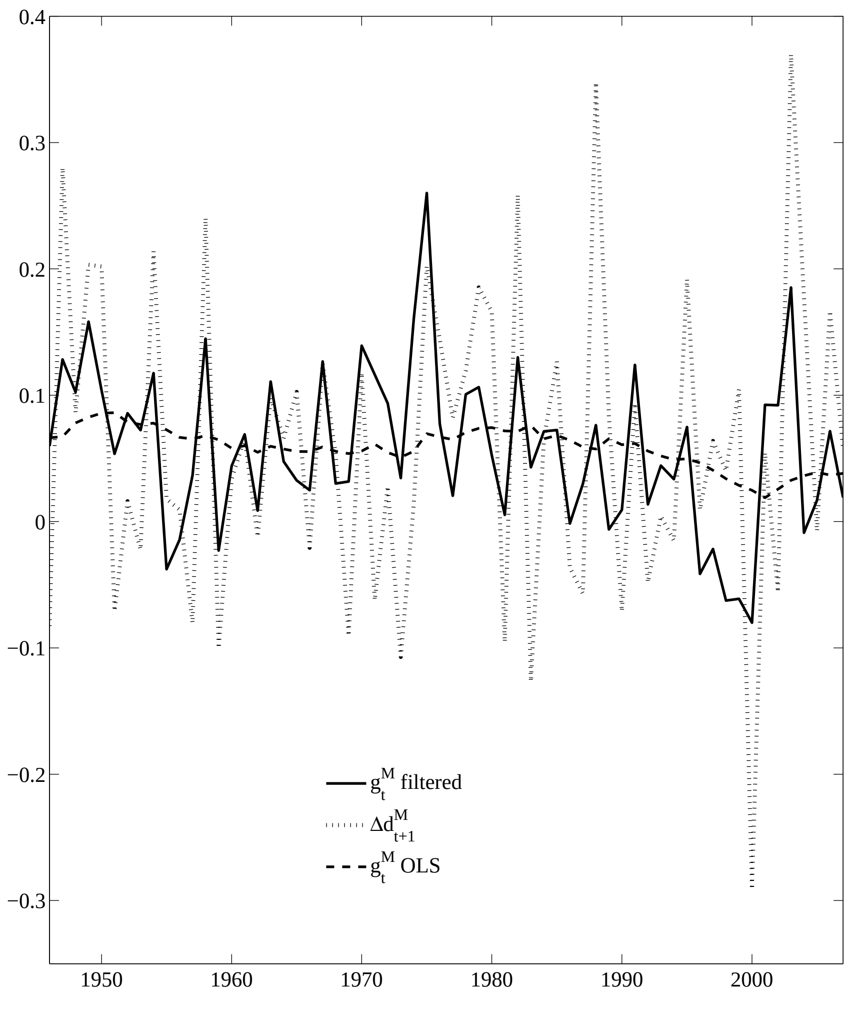

Figure 7: Filtered series for expected dividend growth for reinvestment in the market. The graph plots the filtered series of expected dividend growth, g_t^M, for market-invested dividends, the fitted OLS value, where log dividend growth rates are regressed on the lagged price-dividend ratio, and also the realized dividend growth rate Δd_{t+1}^M.

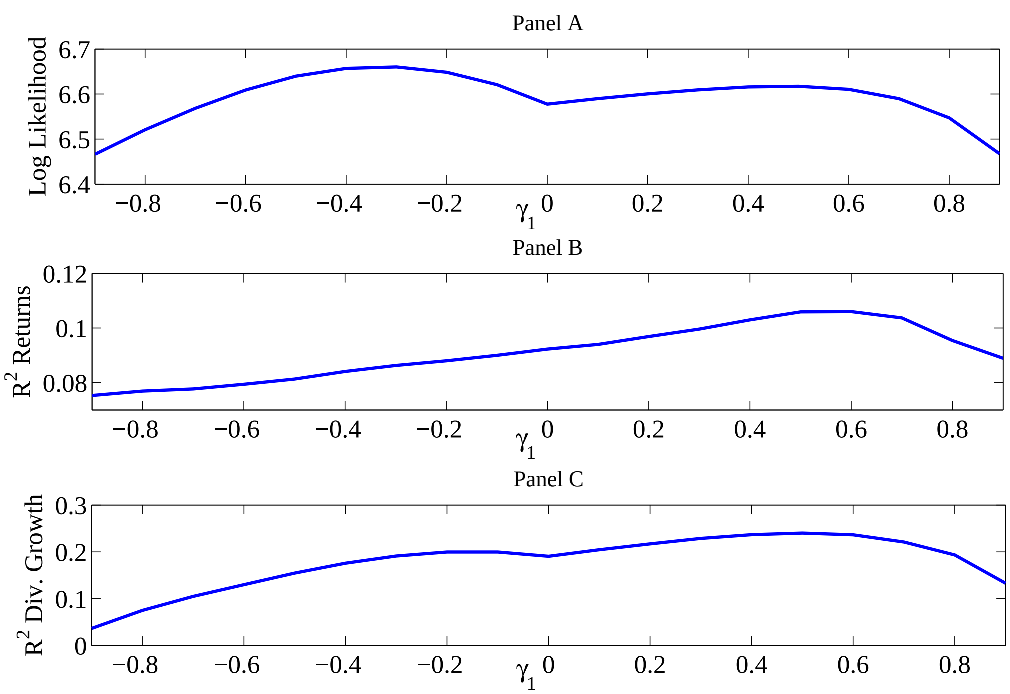

Figure 8: Log likelihood and R^2 values as a function of γ_1.
The graph plots the log likelihood and the R^2 values as a function of the persistence of expected dividend growth, γ_1, using the system described in equations (4)-(6) and data where dividends are reinvested in the aggregate market. We define a grid for γ_1 between -0.9 and 0.9 with step size 0.1, and compute for each of these grid points the likelihood and the R^2 values of the model while optimizing over all the other parameters.