A Bayesian Analysis of the Stock Price Decomposition**

Nathan S. Balke  
Department of Economics, Southern Methodist University  
Dallas, TX 75275  
and  
Research Department, Federal Reserve Bank of Dallas  
Office: (214) 768-2693  
Fax: (214) 768-1821  
E-mail: nbalke@mail.smu.edu

Jun Ma*  
Department of Economics, Finance and Legal Studies  
Culverhouse College of Commerce & Business Administration  
University of Alabama  
Tuscaloosa, AL 35487-0024  
Phone: 205-348-8985  
Email: jma@cba.ua.edu

Mark E. Wohar  
Department of Economics  
Mammel Hall 332S  
University of Nebraska at Omaha  
Omaha, NE 68182-0286  
Phone: 402-554-3712  
Fax: 402-554-2853  
Email: mwohar@mail.unomaha.edu

September 14, 2012

JEL classification: C11, C32, G12

Keywords: weak identification, Bayesian analysis, stock price decomposition, state-space model

* Corresponding Author.

** This paper is written in honor of Charles Nelson who wrote numerous influential articles in the fields of macroeconomics and finance. In particular, this paper follows in the wake of his well cited 2007 article on weak identification. The importance of weak identification in the empirical literature has been neglected in many important works. This paper attempts to reconcile this shortcoming.
A Bayesian Analysis of the Stock Price Decomposition

Abstract

This paper employs the state-space model to reexamine the fundamental issue in finance about whether it is the expected returns or the expected dividends growth that is primarily responsible for the stock price variations. We use Bayesian methods to conduct inference and show that there is a substantial uncertainty about the contributions of expected returns and expected dividends to the fluctuation in the price-dividend ratio when using the aggregate returns and dividends data. The substantial uncertainty of the contributions results from the model being weakly identified, in the sense of Nelson and Startz (2007). Our finding challenges the long held notion in the existing literature that it is the expected returns that contribute most to the price-dividend variations and calls for further investigations using more disaggregated data.
1. Introduction

Finance theory tells us that fluctuations in stock prices are attributable to movements in either expected future dividend growth (cash flows) or expected future returns (the discount rate). Campbell and Shiller (1988a, 1988b) derive an approximate accounting identity based on a linearization to decompose the stock price-dividend ratio into the sum of expectations of future dividend growth and returns. They show that if the price-dividend ratio can predict future stock returns then expected future returns should contribute to the price-dividend variation. Campbell (1991) further notes that the contribution of the expectations of future expected returns to stock price fluctuation will depend not only on the degree of return predictability, but also on the time series properties of expected returns. In particular the expected returns can have a large effect on the stock price provided that the expected returns are persistent, even if stock return predictability is low. Campbell (1991) and Campbell and Ammer (1993) conclude that news about future excess returns is the primary factor behind movements in US stock returns.¹

The relative contribution of expected returns and expected cash flows to movements in the price-dividend ratio has been an area of interest in the finance discipline for decades. A major challenge that researchers face when attempting to investigate the above question is that neither expected cash flows nor expected discount rates are observable. Two primary approaches have been proposed to capture these unobserved expectations of future variables: Vector Autoregression (VAR) and the state-space model. The Vector Autoregression (VAR) decomposition methodology has found that it is expected returns that contribute most to movement in the price-dividend ratio when traditional dividends are employed.² A second

¹ Cuthbertson et al. (1999) finds that similar results hold for the UK stock market.
approach employs a state-space methodology and finds that expected returns are the primary contributor to movements in the price-dividend ratio when dividends are used as the cash flow measure.\textsuperscript{3} Applications of the two approaches using dividends as the cash measure have reached the same conclusion: almost all aggregate stock price variations are driven by discount rate news and almost none by cash flow news.

Most of the above extant literature indicate that not only that expected returns are time varying but that it is expected returns, rather than expected dividend growth, that contributed most to movements in asset prices. A number of important theoretical papers set out to explain this result. Campbell and Cochrane (1999) focus on what they refer to as the surplus consumption ratio (which is current consumption relative to “habit”, essentially the history of aggregate consumption) and show that their model can successfully replicate many empirical features of asset prices. In their model if the surplus consumption ratio is low (typically in recessions), risk aversion will be high, current stock returns will be low, expected future equity returns will be high, and the price-dividend ratio will be low. The model suggests that most of the movement in the price-dividend ratio is due to expected returns as their endowment growth is \textit{i.i.d.}. On the other hand, Bansal and Yaron (2004) depart from the \textit{i.i.d} assumption and argue that there exist persistent components (i.e., the long-run risk) that are common in both consumption and dividend growth. They show that this small but persistent component helps resolve several asset pricing anomalies such as the risk premium puzzle.\textsuperscript{4} In particular, their model implies that consumption and dividend growth are predictable by price-dividend ratio and

\textsuperscript{3} For the State-space methodology see Balke and Wohar (2002), Binsbergen and Koijen (2010), and Ma and Wohar (2012).

\textsuperscript{4} Ma (2012) shows that the long-run risk model is weakly identified and the corrected inference suggests that the uncertainty of the persistent components for the consumption and dividend growth is large.
that it is the expected dividend growth that contributes most to movements in the price-dividend ratio.

Evidently the investigation of the empirical issue whether it is the expected future returns or the expected future dividends growth that is most responsible for the stock price variations has important implications for testing various theoretical asset pricing models. The conclusion of the large extant literature, that it is the expected future returns that are most responsible for movements in the price-dividend ratio is not without criticism. Inspired by Nelson and Kim (1989) who point out that the predictability of stock returns turns out to be insignificant once taking into account finite sample inference bias, Ma and Wohar (2012) apply a nonparametric bootstrap procedure to document the uncertainty of the VAR stock price decomposition and find that the seemingly large contribution of the expected future returns is not statistically significant. There are also a number of other limitations and pitfalls in such VAR decompositions. For example, the finding in the literature that expected returns dominate dividend growth in explaining movements in the price-dividend ratio, is sensitive to the time period (Chen, 2009), and to the choice of predictive variables (Goyal and Welch, 2008, and Chen and Zhao, 2009).

This paper investigates in a deeper way the time series dynamic properties of expected returns and expected cash flows and show that the results of earlier literature are questionable. In particular we employ state-space Bayesian methods to illustrate that the earlier empirical work in this area are subject to severe inference problems which make their findings unreliable. We find that aggregate returns and dividends data cannot provide sufficient statistical evidence to support the notion that it is expected returns that explains the majority of the fluctuation in the price-dividend ratio when dividends are used as the cash flow measure.
To understand early on what the issues are, note that the log price-dividend ratio is a fairly persistent series while (expected) returns and (expected) cash flows appear at a glance to be white noise series. In order to capture the persistence in the log price-dividend ratio, there needs to be a persistent component within (expected) returns and/or (expected) dividend growth. Within a state-space modeling framework, it turns out that one only needs a small permanent component to help explain movements in the log price-dividend ratio as the loading factor on this component in the log price-dividend equation is relatively large. Unfortunately, neither of these two series individually provides much information about their permanent component—they are both close to being white noise. As a result, most of information about movements in the permanent components in both dividends and return is actually contained in the log price-dividend ratio not in the individual series themselves, but this is not sufficient to separately identify a persistent component in dividend growth and returns.

The above is similar to the notion of a small signal-to-noise ratio of the state-space model implying weak identification (in the sense of Nelson and Startz (2007)). In particular, Ma and Nelson (2012) find that the state space models are typically subject to ZILC (Zero-Information-Limit-Condition) as formulated by Nelson and Startz (2007), and as a result, when the signal is small relative to noise for a particular process in the state space model, the process becomes weakly identified and the resulting uncertainty of the estimates are large. In our Bayesian analysis below, we find evidence consistent with weak identification of key variances and co-variances. The weak identification and the ZILC condition makes it very difficult to determine which factor explains movements in the price-dividend ratio.

Ma, Nelson and Startz (2007) also show that the weak identification in a GARCH model implies a great deal of uncertainty of the persistence estimate and the standard inference often fails to report the correct confidence interval.
2. Stock Price Decomposition

Let \( r_{t+1} \) denote the logarithm of equity return during the period \( t+1 \). By definition:

\[
r_{t+1} = \log(P_{t+1} + D_{t+1}) - \log(P_t)
\]

Here, \( P_{t+1} \) is the equity price at the end of time \( t+1 \), and \( D_{t+1} \) is the dividends distributed during the period \( t+1 \). Campbell and Shiller (1988a, 1988b) applied the log-linearization approximation to (2.1) and derived the well-known identity in terms of logarithm variables:

\[
r_{t+1} \approx \kappa + \Delta d_{t+1} + \rho \cdot pd_{t+1} - pd_t
\]

Where, \( \Delta d_{t+1} = \log\left(\frac{D_{t+1}}{D_t}\right) \), \( pd_{t+1} = \log\left(\frac{P_{t+1}}{D_{t+1}}\right) \), \( \rho = \frac{\exp(E[pd_t])}{1 + \exp(E[pd_t])} \), and \( \kappa = \log(1 + \exp(E[pd_t])) - \rho \cdot E[pd_t] \).

By iterating forward and excluding the explosive solution, we obtain:

\[
pd_t = \frac{\kappa}{1 - \rho} + E_t \sum_{j=0}^{\infty} \rho^j (\Delta d_{t+1+j} - r_{t+1+j})
\]

Equation (2.3) states that any variation in the price-dividend ratio must come from either the variation of the agent’s expectation of future dividend growth or future returns. It is important to study which of the two components on the right-hand side of the equation is more important in driving the price-dividend variations. The major challenge in this study is how to estimate the expectations of future dividend growth and future returns. Without a direct observation of the expectations, one has to make assumptions about the agent’s information set and explore possible dynamic patterns in these variables to estimate these expectations. As discussed in the introduction much research work in the literature has relied on the VAR model following Campbell and Shiller’s (1988a, 1988b) pioneering work.
3. State-Space Model Decomposition

Recent work by Balke and Wohar (2002), and Binsbergen and Koijen (2010) have applied state space models to directly model and estimate the expectation processes. The state-space framework offers a nice alternative to VAR model and has the advantages of modeling the expectations directly as latent factors and capturing the long-run serial correlations that a VAR with finite number of lags would have difficulty in doing. The latter point comes from the fact that the state-space model typically results in moving averaging terms in its reduced-form (see Morley, Nelson and Zivot (2003)).

The realized dividend growth and realized returns are the sum of a persistent component and a transitory component:

\[ \Delta d_{t+1} = d_{t+1}^p + d_{t+1}^r + d_{t+1}^e \]  
\[ r_{t+1} = r_{t+1}^p + r_{t+1}^r + r_{t+1}^e \]

For simplicity, we model \( d_{t+1}^p \) and \( r_{t+1}^p \) by random walks with innovations \( \epsilon_{t+1}^{dp} \) and \( \epsilon_{t+1}^{rp} \):

\[ d_{t+1}^p = d_t^p + \epsilon_{t+1}^{dp} \]  
\[ r_{t+1}^p = r_t^p + \epsilon_{t+1}^{rp} \]

Let transitory components of dividends and returns, \( d_{t+1}^r \) and \( r_{t+1}^r \), be modeled as stationary AR(\( p \)) processes:

\[ d_{t+1}^r = \phi_d(L) \cdot d_t^r + \epsilon_{t+1}^{dr} \]  
\[ r_{t+1}^r = \phi_r(L) \cdot r_t^r + \epsilon_{t+1}^{rr} \]
where, $\phi_i(L) = \sum_{j=1}^{n} \phi_{ij} L^{j-1}$, $i = d, r$. Note the expectation of future dividends is given by:

$$g_i = E \Delta d_{t+1} = d_i^p + \phi_d(L) \cdot d_i^r$$

(3.7)

and the expectation of future returns is:

$$\mu_i = E r_{t+1} = r_i^p + \phi_r(L) \cdot r_i^r$$

(3.8)

The terms $d_i^n$ and $r_i^n$ are white noise, $E d_i^n = 0$ and $E r_i^n = 0$. Below we set $d_i^n = 0$.

Write out the companion form of (3.5) and (3.6):

$$Z_{t+1}^i = A_i \cdot Z_t^i + V_{t+1}^i$$

(3.9)

where, $i = d, r$, $Z_{t+1}^d = (d_{t+1}, \ldots, d_{t-p+2})^\prime$, $V_{t+1}^d = (\varepsilon_{t+1, d}, \ldots, 0)^\prime$, $Z_{t+1}^r = (r_{t+1}, \ldots, r_{t-p+2})^\prime$, $V_{t+1}^r = (\varepsilon_{t+1, r}, \ldots, 0)^\prime$, and $A_i$ is the corresponding companion matrix. Plugging (3.9) into (2.3) we can derive the log price-dividend as follows:

$$pd_i = (1 - \rho)^{-1} \left(d_i^p - r_i^p \right) + \varepsilon_i^d \cdot \left[ A_d \cdot (I - \rho \cdot A_d)^{-1} \cdot Z_t^d - A_r \cdot (I - \rho \cdot A_r)^{-1} \cdot Z_t^r \right]$$

(3.10)

where $e_i$ is the selection vector that has 1 in its first element and zero elsewhere. Plugging this result into (2.2) we obtain the following result after some algebra.

$$r_{t+1}^n = (1 - \rho)^{-1} (\varepsilon_{t+1, d} \cdot A_d)^{-1} V_i = \left(1 - \phi_d(\rho)\right)^{-1} \varepsilon_{t+1, d}$$

(3.11)

Here notice that we use the result $e_i \cdot (I - \rho \cdot A_r)^{-1} \cdot V_i = \left(1 - \phi_r(\rho)\right)^{-1} \varepsilon_{t+1, r}$ for $i = d, r$.

Due to this implicit nonlinear restriction derived from the accounting identity that ties together the three variables $\Delta d, r, pd$, one only needs to use two of them to estimate the model and the last can be backed out. We follow Binsbergen and Koijen (2010) to choose to model dividends growth explicitly together with the log price-dividend ratio.

The above model can be put in a state-space representation as follows:
\[ S_t = F(\theta)S_{t-1} + V_t, \quad V_t \sim MVN(0, Q(\theta)), \quad (3.12) \]

Where, \( \theta \) is the vector of structural parameters, the vector of unobserved state variables,

\[ S_t = \left( d_{t}^p, r_{t}^p, d_{t}^r, r_{t}^r, d_{t-1}^r, r_{t-1}^r, \ldots, d_{t-(p-1)}^r, r_{t-(p-1)}^r \right), \]

the vector of innovations to the state variables,

\[ V_t = \left( e_{t}^{dp}, e_{t}^{rp}, e_{t}^{dr}, e_{t}^{rr}, 0_{2x(p-1)} \right), \]

and the transition matrix

\[
F(\theta) = \begin{bmatrix}
I_{2x2} & 0_{2x2} & \cdots & 0_{2x2} \\
0_{2x2} & \phi_1 & \phi_2 & \cdots & \phi_{p-1} & \phi_p \\
I_{2x2} & 0_{2x2} & \cdots & 0_{2x2} & 0_{2x2} \\
\vdots & 0_{2x2} & I_{2x2} & \ddots & \vdots & \vdots \\
0_{2x2} & 0_{2x2} & \cdots & 0_{2x2} & I_{2x2} & 0_{2x2}
\end{bmatrix}, \quad \text{where } \phi_i = \begin{bmatrix} \phi_{d,i} & 0 \\ 0 & \phi_{r,i} \end{bmatrix} \text{ for } i = 1, 2, \ldots, p.
\]

The observation equation is

\[ Y_{t}^{obs} = H(\theta)S_t, \quad (3.13) \]

Where, \( Y_{t}^{obs} = (\Delta d_t, pd_t)' \), and \( H(\theta) = \begin{bmatrix} H_d \\ (H_d - H_r)(I - \rho F)^{-1} F \end{bmatrix}, \) with \( H_d = (1,0,1,0,0_{2x(p-1)}) \) and \( H_r = (0,1,0,1,0_{2x(p-1)}). \)

The log-likelihood function of the above state-space model can be written out via Kalman Filter by following the standard procedure as developed in Kim and Nelson (1999). We retrieve the CRSP market indices of NYSE/AMEX/NASDAQ stocks as the market portfolio and follow Hansen, Heaton and Li’s (2008) aggregation procedure to construct the quarterly real equity return and real dividend growth series for the period 1952-2011. The maximization of the log-likelihood function gives rise to the point estimates which are reported in Table 1.

Balke and Wohar (2002) and Ma and Wohar (2012) show that the decomposition of the stock price-dividend is subject to weak identification issue, in the sense of Nelson and Startz

\[ 6 \text{ Hansen, Heaton, and Li’s data sources and procedures to compute return and dividend series are available on Nan Li’s website: } \text{http://www.bschool.nus.edu.sg/staff/biznl/}. \]
(2007). In particular as we observe from in Figure 1, the stock price-dividend is very persistent but neither the dividends growth nor the returns are persistent. The persistent price-dividend series implies that at least one of the expectations of future dividend growth and returns must be persistent but the observed dividends growth and returns do not contain enough information for us to infer which one is more responsible for most of the stock price variations. To better document the large uncertainty of the stock price decomposition due to weak identification, we resort to a Bayesian approach.

The predictive log likelihood of the above state-space model is given by:

$$
\log(L(Y_T, \theta)) = \sum_{t=1}^{T} \{-0.5 \log(\det( H(\theta) P_{t|t-1} H(\theta))) \\
-0.5(Y_t - H(\theta) S_{t|t-1}')(H(\theta) P_{t|t-1} H(\theta))^{-1}(Y_t - H(\theta) S_{t|t-1})\}
$$

(3.14)

where $S_{t|t-1}$ and $P_{t|t-1}$ are the conditional mean and variance of $S_t$ from the Kalman filter. Given a prior distribution over parameters, $h(\theta)$, the posterior distribution, $P(\theta \mid Y_T)$, is

$$
P(\theta \mid Y_T) \propto L(Y_T, \theta) h(\theta) .
$$

(3.15)

In this section, we consider the case of very diffuse priors so that the likelihood function is the principal determinant of the posterior distribution.\(^7\)

Because the log-likelihood is a nonlinear function of the structural parameter vector, it is not possible to write an analytical expression for the posterior distribution. As a result, we use Bayesian Markov Chain Monte Carlo methods to estimate the posterior distribution of the

----

\(^7\) Formally, for each of the models the individual autoregressive parameters have a prior distribution of joint truncated normal $N(0,1000)$, the variances of innovations in (3.12) are distributed $U(0,10000)$ while the correlations of innovations are distributed $U(-1,1)$. Draws in autoregressive parameters that imply nonstationarity are rejected. These prior distributions ensure that for this model and data, the acceptance in the Metropolis-Hastings sampler depends only on the likelihoods.
parameter vector, $\theta$. In particular, we employ a Metropolis-Hasting sampler to generate draws from the posterior distributions. The algorithm is as follows:

(i) Given a previous draw of the parameter vector, $\theta^{(i-1)}$, draw a candidate vector $\theta^c$ from the distribution $g(\theta | \theta^{(i-1)})$.

(ii) Determine the acceptance probability for the candidate draw,

$$\alpha(\theta^c, \theta^{(i-1)}) = \min \left[ \frac{L(Y_T, \theta^c) h(\theta^c)}{L(Y_T, \theta^{(i-1)}) h(\theta^{(i-1)})}, \frac{g(\theta^{(i-1)} | \theta^c)}{g(\theta^c | \theta^{(i-1)})} \right].$$

(iii) Determine a new draw from the posterior distribution, $\theta^{(i)}$.

$$\theta^{(i)} = \theta^c \quad \text{with probability } \alpha(\theta^c, \theta^{(i-1)})$$

$$\theta^{(i)} = \theta^{(i-1)} \quad \text{with probability } 1 - \alpha(\theta^c, \theta^{(i-1)}).$$

(iv) Return to (i).

Starting from an initial parameter vector and repeating enough times, the distribution parameters draws, $\theta^{(i)}$, will converge to the true posterior distribution.

In our application, $\theta^c = \theta^{(i-1)} + \nu$, where $\nu$ is drawn from a multivariate $t$-distribution with 50 degrees of freedom and a covariance matrix $\Sigma$. We set $\Sigma$ to be a scaled value of the Hessian matrix of $- \log(L(Y_T, \theta))$ evaluated at the maximum likelihood estimates. We choose the scaling so that between 25-40 percent of the candidate draws are accepted. We can also obtain the posterior distributions for the unobserved states. Given a parameter draw, we draw from the conditional posterior distribution for the unobserved states, $P(S_T | \theta^{(i)} , Y_T)$. Here we use the “filter forward, sample backwards” approach proposed by Carter and Kohn (1994) and discussed in Kim and Nelson (1999). The posterior distribution is based on a 750,000 sample of every fifth draw from the MH-MCMC after a burn-in period of 250,000 draws.
4. Estimation Results and Discussions

Table 1 presents the mean, standard deviation, 5th and 95th percentiles of the posterior distributions of the structural state-space model parameters as well as the maximum likelihood estimate of the parameters. Figures 2, 3, and 4 plot the posterior distributions of the parameters. From Table 1 and Figure 2, we see that the mean of the posterior distribution for both of the autoregressive coefficients for transitory returns is centered around zero. The sum of the AR(1) and AR(2) coefficient from the ML estimation is not significantly different from zero. The mean of the posterior distribution for the AR(1) and AR(2) coefficients for the transitory dividend process are 0.32 and 0.23, respectively. One interesting result is the MCMC wants the AR coefficients of transitory returns to be very small and the variance of the returns to be large. The maximum posterior probability point estimates suggest AR(1) coefficient for returns to be about 0.3 and the variance to be small relative to MCMC results. This would appear to be a symptom of identification problems as the MCMC appears to explore other significant modes in the posterior distribution.8

Figure 3 displays posterior distribution of the variances and correlation coefficients of innovations in the state vector. From Table 1 and Figure 3, one observes that the posterior distributions of the correlations are very disperse—the posterior distributions of many of the correlation coefficients span the nearly the entire range from -1 to 1. Interestingly, the posterior distributions of the variances of the two permanent components look quite similar—they have similar means, medians, and standard deviations even though the variances of actual dividend growth and returns are very different (see Figure 1).

---

8 This does not appear to be the result of the Markov chain being stuck at an inferior mode, as the chain traversed over a large part of the parameter space, especially for the correlations between innovations.
Figure 4 and the last row of Table 1 display features of the posterior distribution of \(\text{var}(\varepsilon_t^{dp} - \varepsilon_t^{rp})\). For comparison, we also include the posterior distributions components of this variance: \(\text{var}(\varepsilon_t^{dp})\), \(\text{var}(\varepsilon_t^{rp})\), and \(\text{corr}(\varepsilon_t^{dp}, \varepsilon_t^{rp})\). We truncate the plots of the histograms of \(\text{var}(\varepsilon_t^{dp})\) and \(\text{var}(\varepsilon_t^{rp})\) so that they can be more readily compared with that of \(\text{var}(\varepsilon_t^{dp} - \varepsilon_t^{rp})\). Note the posterior distribution of \(\text{var}(\varepsilon_t^{dp} - \varepsilon_t^{rp})\) is much more concentrated than the posterior distributions of \(\text{var}(\varepsilon_t^{dp})\) and \(\text{var}(\varepsilon_t^{rp})\). This implies that the data is informative about \(\text{var}(\varepsilon_t^{dp} - \varepsilon_t^{rp})\) (i.e. it is precisely estimated) but not so informative about \(\text{var}(\varepsilon_t^{dp})\), \(\text{var}(\varepsilon_t^{rp})\), or \(\text{corr}(\varepsilon_t^{dp}, \varepsilon_t^{rp})\) individually (i.e. these are estimated imprecisely). This suggests that the variances of the innovations in the permanent factors are not well identified; however, the variance of \((\varepsilon_t^{dp} - \varepsilon_t^{rp})\) is.

Figure 5 displays the posterior distribution of the historical decomposition of the log price-dividend ratio. The top panel displays the contribution of expected future dividends while the bottom panel displays the contribution of expected future returns. While the median contribution of expected future returns seems to track the log price-dividend ratio, the 5th and 95th percentile band is quite large. The posterior distribution of the contribution of expected future dividends is also quite disperse. Together these suggest that the model is not too informative about whether it is expectations of future dividends or future returns that are driving stock prices.

To determine whether it is the permanent or transitory components that are important, Figure 6 displays the contribution of just the permanent components to the log price-dividend ratio. The top panel displays the contribution of the permanent component of dividend growth, the middle panel displays the contribution of the permanent component of returns, and the
bottom panel displays the joint contribution of the two permanent components. Two important features stand out from Figure 6. First, the individual contribution of expected future dividends and expected future returns is largely driven by the permanent components and not the transitory components. The posterior distribution for the total contribution of expected dividends and returns (displayed in Figure 5) is similar to the top two panels in Figure 6. This is not too surprising as the transitory components are estimated to have relative small autoregressive parameters and, hence, do not provide much information about future dividends or returns. Second, while there is substantial uncertainty about the contributions of the individual permanent components, the joint contribution of the permanent components has little dispersion and tracks the actual log price-dividend ratio closely. This is consistent with fact that there was substantial uncertainty about the individual variance of (and correlation between) innovations in the permanent components but much less uncertainty about the value of $\var(\epsilon_{t}^{dp} - \epsilon_{t}^{rp})$. Overall, Figures 6 suggests that the permanent components of dividends and returns are driving the log price-dividend ratio, but the data cannot tell which of these two permanent components is the most important.

Given that jointly the permanent components of dividend growth and returns were estimated to be important for fluctuations in the log price-dividend ratio, one might wonder whether such important permanent components are plausible given the lack of persistence in actual dividend growth and excess returns. Figure 7 plots actual dividend growth and actual returns against the median, 5\textsuperscript{th}, and 95\textsuperscript{th} percentiles of the permanent component for dividend growth and returns. What stands out from Figure 7 is that the permanent component for both dividend growth and returns is very small compared with fluctuations in those variables. Note that this was also the case for the variance of innovations in the permanent and temporary
components displayed in Table 1 and Figure 3. Our finding is reminiscent of Nelson and Schwert (1977) who point out that it is possible for the expectation dynamics to be drastically different from the observed series provided that the expectation shock is small relative to the realized shock. Figure 7 also points to the source of the lack of identification in the relative contribution of dividends and returns to fluctuations in the price-dividend ratio. Because the factors loading on the permanent components are so large in the log price-dividend equation, small fluctuations in a permanent component have large effects on the log price-dividend ratio. As both dividend growth and returns have large transitory components yet have small permanent components, these variables individually have little direct information about the permanent components. As a result, nearly all of the information about the two persistent components comes from a single variable: the log price-dividend ratio. Unfortunately, this is not enough information to separately identify a permanent component in dividend growth and a permanent component in returns.

5. Solving weak identification by imposing restrictions on the number of permanent components

One can in principle solve the weak identification problem by imposing restrictions on the more general state-space model. In this section, we explore the consequences of dropping one of the permanent components from the state-space model (i.e. restricting the variance of innovations in that component to be zero).

Figure 8 displays the posterior distributions of the variances and correlations of innovations to the states when there is no permanent component in returns. Comparing the posterior distribution in Figure 8 with those in Figure 3, it is clear that the dispersion in the
posterior distributions of the correlation parameters is substantially smaller in the model with no permanent component in returns than in the model with permanent components in both dividends and returns. Note that the posterior distribution of $\text{var}(\varepsilon_t^{dp})$ for the model with no permanent component in returns looks more like the posterior distribution of $\text{var}(\varepsilon_t^{dp} - \varepsilon_t^p)$ for the model with two permanent components. Figure 9 displays the posterior distributions of variances and correlations for the model where there is no permanent component of dividends. Again, compared to the model with permanent components in both dividends and returns the correlations are much more precisely estimated. Similarly, the posterior distribution of innovations in the single permanent component (in Figure 9, $\text{var}(\varepsilon_t^p)$) looks much like the posterior distribution of $\text{var}(\varepsilon_t^{dp} - \varepsilon_t^p)$ in the model with two permanent components.

While both models’ parameters appear to be estimated with precision, which of the two models does the data prefer? Figure 10 plots the cumulative density function of the likelihood values based on draws from the MCMCs used to estimate the posterior distribution of the parameters. The CDFs for the empirical likelihood values lie nearly on top of one another—neither model appears to stochastically dominate the other—in other words, the data views these models as almost equally likely.

While the data does not appear to prefer one of the models over the other, the two models imply very different implications for stock price decomposition. In fact, for the model with a permanent dividend component and no permanent returns component it is expectations of future dividends that explain nearly all of the fluctuations in the log price-dividend. For the model with a permanent returns component and no permanent component for dividends it is the expectations of future returns explains stock price fluctuations. The top panel of Figure 11 plots the posterior distribution of the historical contribution of the permanent component of dividends to the log
price-dividend ratio for the model without permanent returns; the middle panel plots the historical contribution of the permanent component of returns to the log price-dividend ratio for the model without permanent dividends; for comparison, the bottom panel plots the posterior distribution of the combined historical contribution of the permanent components for the model with both permanent dividends and returns. In all three panels, the permanent component(s) explains the vast majority of price-dividend movements over the sample. In all three panels, the posterior distribution of the contribution of the permanent component(s) is very tight—given a model, there is little uncertainty that the permanent components explain nearly all of the log price-dividend ratio regardless of whether it is in dividend growth alone (top panel), returns alone (middle panel), or in both dividends and returns (bottom panel). Yet, when one takes account that one is uncertain across models, the data is unable to ascertain whether the permanent component(s) is primarily in dividends or in returns.

6. Conclusion

The relative contribution of expected returns and expected cash flows to movements in the price-dividend ratio has been an area of interest in the finance discipline for decades. A major challenge in finance research is to ascertain whether it is expected returns or expected dividends growth that contribute to movements in the price-dividend ratio.

This paper contributes to the existing literature by illustrating that the earlier empirical work in this area is subject to severe inference problems which make their findings unreliable. We employ quarterly real equity return and real dividend growth series for the period 1952-2011 and use Bayesian methods to conduct inference that, in contrast to much of the previous literature, there is substantial uncertainty about whether expected returns or expected dividends
drive stock price fluctuations. Our results show that using aggregate returns and dividends data one cannot provide sufficient statistical evidence to support the long held notion that it is expected returns that explain the majority of the fluctuation in the price-dividend ratio when dividends are used as the cash flow measure. What the data are sure of is that there is a small, persistent component affecting the log price-dividend ratio but the data are essentially silent on whether it is in dividends or returns.
<table>
<thead>
<tr>
<th>parameter</th>
<th>Maximum of posterior probability function</th>
<th>Mean of posterior distribution</th>
<th>St. dev. of posterior distribution</th>
<th>5\textsuperscript{th} percentile of posterior distribution</th>
<th>95\textsuperscript{th} percentile of posterior distribution</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\phi_{d,1}$</td>
<td>0.33168</td>
<td>0.32766</td>
<td>0.066572</td>
<td>0.21769</td>
<td>0.43601</td>
</tr>
<tr>
<td>$\phi_{r,1}$</td>
<td>0.33531</td>
<td>0.022255</td>
<td>0.11652</td>
<td>-0.11624</td>
<td>0.23836</td>
</tr>
<tr>
<td>$\phi_{d,2}$</td>
<td>0.2229</td>
<td>0.22914</td>
<td>0.064297</td>
<td>0.12356</td>
<td>0.33448</td>
</tr>
<tr>
<td>$\phi_{r,2}$</td>
<td>0.30176</td>
<td>0.023044</td>
<td>0.10237</td>
<td>-0.0996</td>
<td>0.22675</td>
</tr>
<tr>
<td>$\text{var}(\varepsilon_{i}^{dp})$</td>
<td>8.53E-05</td>
<td>0.017412</td>
<td>0.040751</td>
<td>2.77E-05</td>
<td>0.081385</td>
</tr>
<tr>
<td>$\text{var}(\varepsilon_{i}^{dr})$</td>
<td>3.6304</td>
<td>3.778</td>
<td>0.43494</td>
<td>3.126</td>
<td>4.5319</td>
</tr>
<tr>
<td>$\text{var}(\varepsilon_{i}^{rp})$</td>
<td>0.002087</td>
<td>0.019619</td>
<td>0.037781</td>
<td>0.001915</td>
<td>0.079824</td>
</tr>
<tr>
<td>$\text{var}(\varepsilon_{i}^{rt})$</td>
<td>62.907</td>
<td>2575.3</td>
<td>3985.1</td>
<td>129.15</td>
<td>9337.9</td>
</tr>
<tr>
<td>$\text{corr}(\varepsilon_{i}^{dp}, \varepsilon_{i}^{dr})$</td>
<td>-0.26609</td>
<td>-0.07817</td>
<td>0.64965</td>
<td>-0.99392</td>
<td>0.98655</td>
</tr>
<tr>
<td>$\text{corr}(\varepsilon_{i}^{dp}, \varepsilon_{i}^{rp})$</td>
<td>-0.67979</td>
<td>0.55852</td>
<td>0.47759</td>
<td>-0.51583</td>
<td>0.98435</td>
</tr>
<tr>
<td>$\text{corr}(\varepsilon_{i}^{dp}, \varepsilon_{i}^{rt})$</td>
<td>0.66205</td>
<td>-0.00138</td>
<td>0.50914</td>
<td>-0.82478</td>
<td>0.82</td>
</tr>
<tr>
<td>$\text{corr}(\varepsilon_{i}^{dr}, \varepsilon_{i}^{rp})$</td>
<td>0.30062</td>
<td>0.14527</td>
<td>0.506</td>
<td>-0.86245</td>
<td>0.87093</td>
</tr>
<tr>
<td>$\text{corr}(\varepsilon_{i}^{dr}, \varepsilon_{i}^{rt})$</td>
<td>0.029226</td>
<td>0.076433</td>
<td>0.27935</td>
<td>-0.33432</td>
<td>0.61116</td>
</tr>
<tr>
<td>$\text{corr}(\varepsilon_{i}^{rp}, \varepsilon_{i}^{rt})$</td>
<td>-0.94262</td>
<td>0.007345</td>
<td>0.49364</td>
<td>-0.77055</td>
<td>0.77131</td>
</tr>
<tr>
<td>$\text{var}(\varepsilon_{i}^{dp} - \varepsilon_{i}^{rp})$</td>
<td>0.002746</td>
<td>0.004402</td>
<td>0.000969</td>
<td>0.002977</td>
<td>0.006078</td>
</tr>
</tbody>
</table>
Figure 1.
Figure 2.
Histogram of posterior distribution of autoregressive parameters for model with permanent components in both dividends and returns
Figure 3.
Posterior distributions of variances and correlations of innovations for state-space model with permanent components in dividend growth and returns.
Figure 4.
Posterior distribution of the components of $\text{var}(\varepsilon_i^{dp} - \varepsilon_i^{tr})$
Figure 5.
Historical decomposition of log price-dividend ratio: contribution of expected future dividends and expected future returns. Median, 5th, and 95th percentiles of the posterior distribution.
Figure 6.
Contribution of permanent components to log price-dividend ratio. Median, 5\textsuperscript{th}, and 95\textsuperscript{th} percentiles of the posterior distribution.
Figure 7.
Actual dividend growth and returns along with permanent components of dividend growth and returns. Median, 5th, and 95th percentiles.
Figure 8.
Posterior distribution of variances and correlations of innovations in state-space model with a permanent component in dividend growth but no permanent component in returns.
Figure 9.
Posterior distribution of variances and correlations of innovations in state-space model with a permanent component in returns but no permanent component in dividend growth.
Figure 10.
Cumulative density function for likelihood values from MCMC draws for models with a single permanent component
Figure 11.
Contribution of permanent components to log price-dividend. Median, 5th, and 95th percentiles of the posterior distribution.
References


