A Bayesian Examination of Weak Identification in Stock Price Decompositions **

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

March 11, 2013

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 Examination of Weak Identification in Stock Price Decompositions

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.
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 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
ccontributor 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

\begin{footnotesize}
\begin{itemize}
\item broader definition of dividends is employed (called net payout), expected net payout is found to contribute most to
movement in the price-net payout ratio.
\item For the State-space methodology see Balke and Wohar (2002), Binsbergen and Koijen (2010), and Ma and Wohar (2012).
\item Ma (2013) 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.
\end{itemize}
\end{footnotesize}
that it is the expected dividend growth that contributes most to movements in the price-dividend ratio.

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 (1993) who point out that the predictability of stock returns turn 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 specification of the VAR (Balke and Wohar, 2006), to the sample 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 shows 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. In particular, we argue that stock price decompositions are plagued by weak identification. While this has been hinted at in Balke and Wohar (2002) and Ma and Wohar (2012), what separates the current analysis from our earlier work is that the Bayesian methods employed in this paper allow us to pinpoint the source of the weak identification as due to uncertainty about the relative variances of the persistent components of dividend growth and returns. The Bayesian methods also allow us to better document just how large this uncertainty is. Once the uncertainty about
the variances of the persistent components of dividend growth and returns is taken into account, 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 returns and 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.\(^5\) In our Bayesian

---

\(^5\) 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.
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.

2. Stock Price Decomposition

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

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

(2.1)

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:

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

(2.2)

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

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

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

(2.3)

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:

\begin{align}
    d_{t+1} &= d_{t+1}^p + d_{t+1}^r \quad (3.1) \\
    r_{t+1} &= r_{t+1}^p + r_{t+1}^r + r_{t+1}^n \quad (3.2)
\end{align}

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

\begin{align}
    d_{t+1}^p &= d_t + \epsilon_{t+1}^{dp} \quad (3.3) \\
    r_{t+1}^p &= r_t + \epsilon_{t+1}^{rp} \quad (3.4)
\end{align}
Let transitory components of dividends and returns, \( d_{t+1}^r \) and \( r_{t+1}^r \), be modeled as stationary AR(\( p \)) processes:

\[
d_{t+1}^r = \phi_d (L) \cdot d_t^r + \varepsilon_{t+1}^{dr} \quad (3.5)
\]

\[
r_{t+1}^r = \phi_r (L) \cdot r_t^r + \varepsilon_{t+1}^{rr} \quad (3.6)
\]

where, \( \phi_i (L) = \sum_{j=1}^{p} \phi_{i,j} L^{j-1} \), \( i = d, r \). Note the expectation of future dividends is given by:

\[
E d_{t+1} = d_t^p + \phi_d (L) \cdot d_t^r \quad (3.7)
\]

and the expectation of future returns is:

\[
E r_{t+1} = r_t^p + \phi_r (L) \cdot r_t^r \quad (3.8)
\]

The additional term in the returns equation (3.2), \( r_{t+1}^n \), is white noise, \( E r_{t+1}^n = 0 \). To see this, write out the companion form of (3.5) and (3.6):

\[
Z_{t+1}^i = A_i \cdot Z_t^i + V_{t+1}^i \quad (3.9)
\]

where, \( i = d, r \), \( Z_{t+1}^d = (d_{t+1}^r, \ldots, d_{t-p+r}^r)' \), \( V_{t+1}^d = (\varepsilon_{t+1}^{dr}, \ldots, 0)' \), \( Z_{t+1}^r = (r_{t+1}^r, \ldots, r_{t-p+r}^r)' \), \( V_{t+1}^r = (\varepsilon_{t+1}^{rr}, \ldots, 0)' \), and \( A_i \) is the corresponding companion matrix. Plugging (3.9) into (2.3) we can derive the log price-dividend as follows:

\[
pd_t = (1 - \rho)^{-1} (d_t^p - r_t^p) + e_t^i \cdot \left[ A_d \cdot (I - \rho \cdot A_d)^{-1} \cdot Z_t^d - A_r \cdot (I - \rho \cdot A_r)^{-1} \cdot Z_t^r \right] \quad (3.10)
\]

where \( e_1 \) is the selection vector that has 1 in its first element and zero elsewhere. Plugging this result into (2.2) we obtain after some algebra:

\[
r_{t+1}^n = (1 - \rho)^{-3} (\varepsilon_{t+1}^{dp} - \varepsilon_{t+1}^{rp}) + (1 - \phi_d (\rho))^{-1} \varepsilon_{t+1}^{dr} - (1 - \phi_r (\rho))^{-1} \varepsilon_{t+1}^{rr} \quad (3.11)
\]

Here notice that we use the result \( e_t^i \cdot (I - \rho \cdot A_i)^{-1} \cdot V_{t+1}^i = (1 - \phi_d (\rho))^{-1} \varepsilon_{t+1}^{dr} \) for \( i = d, r \).
Note that $r_{jt}^n$ is a linear combination of shocks to the other underlying components and that this particular linear combination has no effect on the log price-dividend ratio. Due to this restriction derived from the accounting identity that ties together the three variables $d, r, pd$, one only needs to use two of them to estimate the model while the third variable 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,

$$ S_t = \left( d_t^p, r_t^p, d_t^c, r_t^c, \ldots, d_{t-(p-1)}^p, r_{t-(p-1)}^c \right), $$

the vector of innovations to the state variables,

$$ V_t = \left( \varepsilon_t^{dp}, \varepsilon_t^{cp}, \varepsilon_t^{dc}, \varepsilon_t^{cc}, 0_{2x(p-1)} \right), $$

and the transition matrix

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

The observation equation is

$$ Y_t^{obs} = H(\theta)S_t, \quad (3.13) $$

where, $Y_t^{obs} = (d_t, pd_t)'$ and $H(\theta) = H_d F$, with $H_d = (I, 0, 1, 0, 0, 0_{2x(p-1)})$ and $H_r = (0, 1, 0, 1, 0, 0_{2x(p-1)})$. The observation variables, dividend growth $(d_t)$ and the log price-dividend ratio $(pd_t)$, are from CRSP market indices of NYSE/AMEX/NASDAQ. We 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. In the estimated model, we set the number of lags in the transitory components of dividend growth and returns to be two (p=2).

Before moving onto the empirical results, a short discussion of how to interpret the unobserved components, $d_t^p, r_t^p, d_t^r, r_t^r$, might be useful. The market fundamentals interpretation of these components are that they reflect the predictable components in dividend growth ($d_t^p, d_t^r$) and in returns ($r_t^p, r_t^r$). These, through the present value relationship, determine the log price-dividend ratio. Given that dividend growth is observed and is assumed not to depend on the return components, an alternative interpretation is to think of the return components as idiosyncratic components of the log price-dividend ratio. These idiosyncratic components could reflect market fundamentals such as expectations about future returns but they could also reflect non-market fundamental effects on the log price-dividend ratio. For the discussion of weak identification, it does not matter which interpretation one uses.

### 4. Estimation Results

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

---

6 Hansen, Heaton, and Li’s data sources and procedures to compute return and dividend series are available on Nan Li’s website: [http://www.bschool.nus.edu.sg/staff/biznl/](http://www.bschool.nus.edu.sg/staff/biznl/).

7 Balke and Wohar (2009) consider the case of an explicitly non-fundamental component in the log price-dividend ratio in the form of bursting and reoccurring, rational bubble. This differed from the expected future returns components in that a rational bubble imposes fairly strong restrictions on the stochastic process governing the bubble component.
estimate of the parameters.\(^8\) 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 (given the relatively uniformed priors essentially the maximum likelihood) 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.\(^9\)

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).

---

\(^8\) The posterior distribution is based on a 750,000 sample of every fifth draw from the Metropolis Hasting (MH) MCMC after a burn-in period of 250,000 draws. 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.

\(^9\) 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 \( \text{var}(\varepsilon_i^{dp} - \varepsilon_i^{rp}) \). For comparison, we also include the posterior distributions components of this variance: \( \text{var}(\varepsilon_i^{dp}) \), \( \text{var}(\varepsilon_i^{rp}) \), and \( \text{corr}(\varepsilon_i^{dp}, \varepsilon_i^{rp}) \). We truncate the plots of the histograms of \( \text{var}(\varepsilon_i^{dp}) \) and \( \text{var}(\varepsilon_i^{rp}) \) so that they can be more readily compared with that of \( \text{var}(\varepsilon_i^{dp} - \varepsilon_i^{rp}) \).

Note the posterior distribution of \( \text{var}(\varepsilon_i^{dp} - \varepsilon_i^{rp}) \) is much more concentrated than the posterior distributions of \( \text{var}(\varepsilon_i^{dp}) \) and \( \text{var}(\varepsilon_i^{rp}) \). This implies that the data is informative about \( \text{var}(\varepsilon_i^{dp} - \varepsilon_i^{rp}) \) (i.e. it is precisely estimated) but not so informative about \( \text{var}(\varepsilon_i^{dp}) \), \( \text{var}(\varepsilon_i^{rp}) \), or \( \text{corr}(\varepsilon_i^{dp}, \varepsilon_i^{rp}) \) 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 \( \text{var}(\varepsilon_i^{dp} - \varepsilon_i^{rp}) \) 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 5\(^{th}\) and 95\(^{th}\) 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 $\text{var}(\epsilon_{t}^{dp} - \epsilon_{t}^{rp})$. 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, 5th, and 95th 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 $\text{var}(\varepsilon_{it}^{dp})$ for the model with no permanent component in returns looks more like the posterior distribution of $\text{var}(\varepsilon_{it}^{dp} - \varepsilon_{it}^{np})$ 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, $\text{var}(\varepsilon_{it}^{np})$) looks much like the posterior distribution of $\text{var}(\varepsilon_{it}^{dp} - \varepsilon_{it}^{np})$ 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. Further Analysis Using Disaggregated Portfolios

The results in the previous sections applied to essentially the market portfolio. As expected returns largely reflects systematic risk, one might expect movements in the market log price-dividend ratio to reflect changes in expected returns (as has been the dominant finding in the literature). More disaggregated portfolios are more likely to be influenced by idiosyncratic cash flows than in aggregate portfolios where this idiosyncratic cash flow gets diversified away. This suggests that at a more disaggregated level, dividend growth should matter more for stock return decompositions. In terms of decompositions of the price-dividend ratio, however, it is

---

10 Cochrane (2008) argues that much of the cash flow variance is idiosyncratic in nature while the expected return variance is common (i.e. related to systematic risk). Studies by Vuolteenaho (2002), Callen and Segal (2004), and Chen, Da, and Zhao (2012) have found that cash flow news is the primary factor explaining the variance in stock
not clear whether examining disaggregated portfolios helps to solve the weak identification problem pointed out in the previous sections. While idiosyncratic cash flows get diversified away, using aggregate dividend growth might actually filter out some of the high-frequency/idiosyncratic fluctuations in cash flows allowing for better inference about the low frequency component in dividend growth. To ascertain whether more disaggregated portfolios suffer from the same sort of weak identification problem that we have showed plague aggregate portfolios, we apply our above Bayesian MCMC methodology to portfolios arranged by firm size, book-to-market, dividend-yield, earnings-price, and cash-flow price-ratio.

The returns with and without dividends distribution for these portfolios are obtained from Kenneth French’s website. We impute the price-dividend ratio and aggregate the monthly observation to quarterly frequency by following the same procedure as we described and applied to the aggregate market portfolio in section 3. We examine the bottom and top deciles for each of these portfolios. The sample period is from 1952 to 2011, the same as the aggregate market portfolio analyzed in previous sections. We employ the model which allows for a permanent component in both expected dividend growth and expected returns.

Figure 12 presents the results for the bottom decile while Figure 13 presents the results for the top decile. The first column of each figure presents the contribution of permanent dividend growth to fluctuations in the log price-dividend ratio, the second column presents the contribution of permanent returns, while the third column report the results for the joint contribution of permanent dividend growth and permanent returns. Similar to the aggregate results, reported earlier in the paper, the contribution of permanent dividend growth and permanent returns are estimated imprecisely individually. However, the joint contribution of returns at the firm level, while discount rate news explains most of the movements in stock returns at the aggregate level.
permanent dividend growth and permanent returns is estimated very precisely. We are thus unable to determine the relative contribution of each component individually.

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 contributes 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 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. The source of weak identification is about the relative size of the variances of innovations to the low-frequency component of dividends or returns. 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 1. Selected statistics of the posterior distribution of parameters for model with two permanent components

<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>5th percentile of posterior distribution</th>
<th>95th 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>( var(\epsilon_{dp}^d) )</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>( var(\epsilon_{dt}^d) )</td>
<td>3.6304</td>
<td>3.778</td>
<td>0.43494</td>
<td>3.126</td>
<td>4.5319</td>
</tr>
<tr>
<td>( var(\epsilon_{tp}^r) )</td>
<td>0.002087</td>
<td>0.019619</td>
<td>0.037781</td>
<td>0.001915</td>
<td>0.079824</td>
</tr>
<tr>
<td>( var(\epsilon_{tt}^r) )</td>
<td>62.907</td>
<td>2575.3</td>
<td>3985.1</td>
<td>129.15</td>
<td>9337.9</td>
</tr>
<tr>
<td>( corr(\epsilon_{dp}^d, \epsilon_{dt}^d) )</td>
<td>-0.26609</td>
<td>-0.07817</td>
<td>0.64965</td>
<td>-0.99392</td>
<td>0.98655</td>
</tr>
<tr>
<td>( corr(\epsilon_{dp}^d, \epsilon_{tp}^r) )</td>
<td>-0.67979</td>
<td>0.55852</td>
<td>0.47759</td>
<td>-0.51583</td>
<td>0.98435</td>
</tr>
<tr>
<td>( corr(\epsilon_{dt}^d, \epsilon_{tp}^r) )</td>
<td>0.66205</td>
<td>-0.00138</td>
<td>0.50914</td>
<td>-0.82478</td>
<td>0.82</td>
</tr>
<tr>
<td>( corr(\epsilon_{tp}^r, \epsilon_{tp}^r) )</td>
<td>0.30062</td>
<td>0.14527</td>
<td>0.506</td>
<td>-0.86245</td>
<td>0.87093</td>
</tr>
<tr>
<td>( corr(\epsilon_{dt}^d, \epsilon_{tt}^r) )</td>
<td>0.029226</td>
<td>0.076433</td>
<td>0.27935</td>
<td>-0.33432</td>
<td>0.61116</td>
</tr>
<tr>
<td>( corr(\epsilon_{tp}^r, \epsilon_{tt}^r) )</td>
<td>-0.94262</td>
<td>0.007345</td>
<td>0.49364</td>
<td>-0.77055</td>
<td>0.77131</td>
</tr>
<tr>
<td>( var(\epsilon_{dp}^d - \epsilon_{tp}^r) )</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 \( \text{var}(\epsilon_{i}^{dp} - \epsilon_{i}^{rp}) \)
Figure 5.
Historical decomposition of log price-dividend ratio: contribution of expected future dividends and expected future returns. Median, 5\textsuperscript{th}, and 95\textsuperscript{th} percentiles of the posterior distribution.
Figure 6.
Contribution of permanent components to log price-dividend ratio. Median, 5th, and 95th 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.
Figure 12.
Decomposition of log price dividend ratio disaggregated Fama-French portfolios: bottom deciles
Figure 13.
Decomposition of log price dividend ratio disaggregated Fama-French portfolios: top deciles
References


Appendix. MCMC Algorithm

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

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

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

\[
P(\theta | Y_T) \propto L(Y_T, \theta)h(\theta).
\]

In this section, we consider the case of very diffuse priors so that the likelihood function is the principal determinant of the posterior distribution.\(^{11}\)

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

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

\[
\begin{align*}
\theta^{(i)} &= \theta^c \quad \text{with probability } \alpha(\theta^c, \theta^{(i-1)}) \\
\theta^{(i)} &= \theta^{(i-1)} \quad \text{with probability } 1 - \alpha(\theta^c, \theta^{(i-1)}).
\end{align*}
\]

(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 \(-log(L(\mathbf{Y}_T, \theta))\) 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, \( P(\mathbf{S}_T | \theta^{(i)}, \mathbf{Y}_T) \). 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.