Macroeconomic Dynamics, **19**, 2015, 728–752. Printed in the United States of America.
doi:10.1017/S1365100513000588

A BAYESIAN ANALYSIS OF WEAK IDENTIFICATION IN STOCK PRICE DECOMPOSITIONS

NATHAN S. BALKE
Southern Methodist University
and
Federal Reserve Bank of Dallas

JUN MA
University of Alabama

MARK E. WOHAR
University of Nebraska at Omaha

This paper employs the state-space model to reexamine the fundamental issue in finance of whether it is the expected returns or the expected dividends growth that is primarily responsible for stock price variations. We use Bayesian methods to show that there is a substantial uncertainty about the contributions of expected returns and expected dividends to fluctuations in the price–dividend ratio when the aggregate returns and dividends data are used. The substantial uncertainty of the contributions results from the model being weakly identified. Our finding challenges the notion long held in the existing literature that it is the expected returns that contribute most to price–dividend variations.

Keywords: Weak Identification, Bayesian Analysis, Stock Price Decomposition, State-Space Model

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.

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 often cited 2007 article on weak identification. The importance of weak identification in the empirical literature has been neglected in many important studies. This paper attempts to address this shortcoming. Address correspondence to: Jun Ma, Department of Economics, Finance and Legal Studies, Culverhouse College of Commerce and Business Administration, University of Alabama, Tuscaloosa, AL 35487-0024, USA; e-mail: jma@cba.ua.edu.

© 2014 Cambridge University Press 1365-1005/14

https://doi.org/10.1017/S1365100513000588 Published online by Cambridge University Press

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 returns 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 U.S. stock returns.[1]

The relative contributions of expected returns and expected cash flows to movements in the price–dividend ratio have been an area of interest in the finance discipline for decades. A major challenge that researchers face when attempting to investigate the 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 approach 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 approach employs a state-space method 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.[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 this extant literature indicates not only that expected returns are time-varying but also that it is expected returns, rather than expected dividend growth, that contribute 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 i.i.d. On the other hand, Bansal and Yaron (2004) depart from the 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.[4] In particular, their model implies that consumption and dividend growth are predictable by the price–dividend ratio and 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 turns out to be insignificant once finite sample inference bias

https://doi.org/10.1017/S1365100513000588 Published online by Cambridge University Press

is taken into account, 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 the earlier literature are questionable. In particular, we employ state-space Bayesian methods to illustrate that the earlier empirical work in this area is subject to severe inference problems, which makes reported findings unreliable. In particular, we argue that stock price decompositions are plagued by weak identification. Although 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 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, whereas 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 its permanent component—they are both close to being white noise. As a result, most 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 preceding 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 a ZILC (zero-information-limit condition) as formulated by Nelson and Startz (2007), and as a result, when the signal is small relative to noise

https://doi.org/10.1017/S1365100513000588 Published online by Cambridge University Press

for a particular process in the state-space model, the process becomes weakly identified and the resulting uncertainty of the estimates is high.[5] In our Bayesian analysis, we find evidence consistent with weak identification of key variances and covariances. The weak identification and the ZILC condition make 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,

$$r_{t+1} = \log(P_{t+1} + D_{t+1}) - \log(P_t). \tag{1}$$

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 (1) and derived the well-known identity in terms of logarithmic variables

$$r_{t+1} \approx \kappa + d_{t+1} + \rho \cdot pd_{t+1} - pd_t, \tag{2}$$

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

By iterating (2) forward and excluding the explosive solution, one can show that the price–dividend ratio is the sum of discounted future returns and dividends growth. If we assume finite conditional expectations for both future returns and dividends growth, we can take conditional expectations of both sides and obtain the following result:

$$pd_t = \frac{\kappa}{1-\rho} + E_t \sum_{j=0}^{\infty} \rho^j (d_{t+1+j} - r_{t+1+j}). \tag{3}$$

Equation (3) states that any variation in the price–dividend ratio must come either from the variation of future dividend growth expectations or from future returns expectations. 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 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.

https://doi.org/10.1017/S1365100513000588 Published online by Cambridge University Press

3. STATE-SPACE MODEL DECOMPOSITION

Recent work by Balke and Wohar (2002) and Binsbergen and Koijen (2010) has applied state-space models to model and estimate the expectation processes directly. The state-space framework offers a nice alternative to the 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 a finite number of lags would have difficulty in capturing. The latter point comes from the fact that the state-space model typically results in moving average terms in its reduced form [see Morley et al. (2003)].

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

$$d_{t+1} = d^p_{t+1} + d^\tau_{t+1}, \tag{4}$$

$$r_{t+1} = r^p_{t+1} + r^\tau_{t+1} + r^n_{t+1}. \tag{5}$$

For simplicity, we model d^p_{t+1} and r^p_{t+1} by random walks with innovations ε^{dp}_{t+1} and ε^{rp}_{t+1}:

$$d^p_{t+1} = d^p_t + \varepsilon^{dp}_{t+1}, \tag{6}$$

$$r^p_{t+1} = r^p_t + \varepsilon^{rp}_{t+1}. \tag{7}$$

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

$$d^\tau_{t+1} = \phi_d(L) \cdot d^\tau_t + \varepsilon^{d\tau}_{t+1}, \tag{8}$$

$$r^\tau_{t+1} = \phi_r(L) \cdot r^\tau_t + \varepsilon^{r\tau}_{t+1}, \tag{9}$$

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

$$\underset{t}{E}\, d_{t+1} = d^p_t + \phi_d(L) \cdot d^\tau_t, \tag{10}$$

and the expectation of future returns is

$$\underset{t}{E}\, r_{t+1} = r^p_t + \phi_r(L) \cdot r^\tau_t. \tag{11}$$

The additional term in the returns equation, (5), r^n_{t+1}, is white noise, with $E_t\, r^n_{t+1} = 0$. To see this, write out the companion form of (8) and (9),

$$Z^i_{t+1} = A_i \cdot Z^i_t + V^i_{t+1}, \tag{12}$$

where $i = d, r$, $Z^d_{t+1} = (d^\tau_{t+1}, \ldots, d^\tau_{t-p+2})'$, $V^d_{t+1} = (\varepsilon^{d\tau}_{t+1}, \ldots, 0)'$, $Z^r_{t+1} = (r^\tau_{t+1}, \ldots, r^\tau_{t-p+2})'$, $V^r_{t+1} = (\varepsilon^{r\tau}_{t+1}, \ldots, 0)'$, and A_i is the corresponding companion matrix. Plugging (12) into (3), we can derive the log price–dividend as

https://doi.org/10.1017/S1365100513000588 Published online by Cambridge University Press

follows:

$$pd_t = (1-\rho)^{-1}\left(d_t^p - r_t^p\right) + e_1' \cdot \left[A_d \cdot (I - \rho \cdot A_d)^{-1} \cdot Z_t^d - A_r \cdot (I - \rho \cdot A_r)^{-1} \cdot Z_t^r\right], \tag{13}$$

where e_1 is the selection vector, which has 1 as its first element and 0 elsewhere. Plugging this result into (2) we obtain after some algebra

$$r_{t+1}^n = (1-\rho)^{-1}\left(\varepsilon_{t+1}^{dp} - \varepsilon_{t+1}^{rp}\right) + [1 - \varphi_d(\rho)]^{-1}\varepsilon_{t+1}^{d\tau} - [1 - \varphi_r(\rho)]^{-1}\varepsilon_{t+1}^{r\tau}. \tag{14}$$

Notice here that we use the result $e_1' \cdot (I - \rho \cdot A_i)^{-1} \cdot V_{t+1}^i = [1 - \phi_d(\rho)]^{-1}\,\varepsilon_{t+1}^{i\tau}$ for $i = d, r$.

Note that r_{t+1}^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. Because of this restriction, derived from the accounting identity that ties together the three variables d, r, and pd, one only needs to use two of them to estimate the model, whereas the third variable can be backed out. We follow Binsbergen and Koijen (2010) in choosing to model dividends growth explicitly, together with the log price–dividend ratio.

The preceding model can be put into a state-space representation as follows:

$$S_t = F(\theta)S_{t-1} + V_t,\; V_t \sim MVN(0, Q(\theta)), \tag{15}$$

where θ is the vector of structural parameters, $S_t = (d_t^p, r_t^p, d_t^\tau, r_t^\tau, d_{t-1}^\tau, r_{t-1}^\tau, \ldots, d_{t-(p-1)}^\tau, r_{t-(p-1)}^\tau)'$ the vector of unobserved state variables, $V_t = (\varepsilon_t^{dp}, \varepsilon_t^{rp}, \varepsilon_t^{d\tau}, \varepsilon_t^{r\tau}, 0_{2\times(p-1)})'$ the vector of innovations to the state variables, and the transition matrix is

$$F(\theta) = \begin{bmatrix} I_{2\times2} & 0_{2\times2} & & \cdots & & 0_{2\times2} \\ 0_{2\times2} & \phi_1 & \phi_2 & \cdots & \phi_{p-1} & \phi_p \\ & I_{2\times2} & 0_{2\times2} & \cdots & 0_{2\times2} & 0_{2\times2} \\ \vdots & 0_{2\times2} & I_{2\times2} & \ddots & & \vdots \\ & & & \ddots & \ddots & \vdots \\ 0_{2\times2} & 0_{2\times2} & \cdots & 0_{2\times2} & I_{2\times2} & 0_{2\times2} \end{bmatrix},$$

where

$$\phi_i = \begin{bmatrix} \phi_{d,i} & 0 \\ 0 & \phi_{r,i} \end{bmatrix} \quad \text{for } i = 1, 2, \ldots p.$$

The observation equation is

$$Y_t^{\text{obs}} = H(\theta)S_t, \tag{16}$$

https://doi.org/10.1017/S1365100513000588 Published online by Cambridge University Press

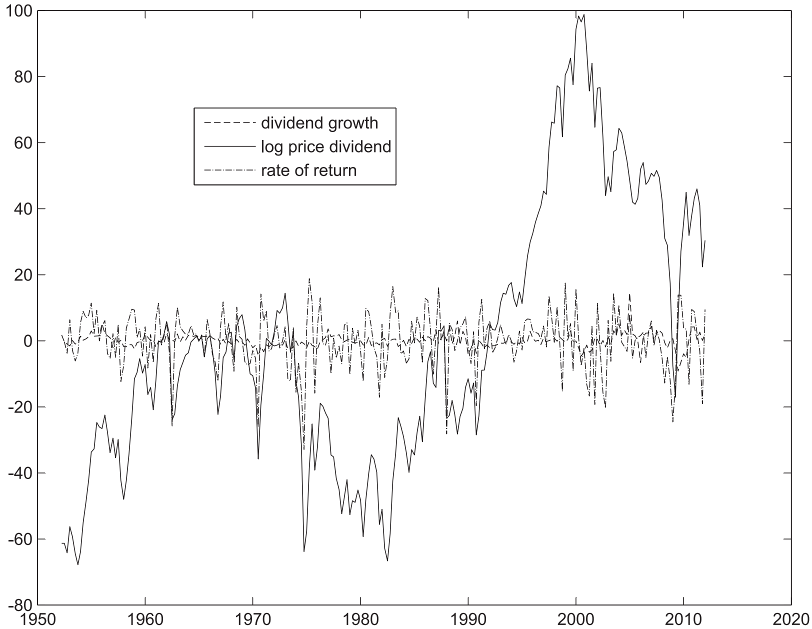

FIGURE 1. Log price–dividend ratio, dividend growth, and the rate of return, 1952–2011.

where $Y_t^{\text{obs}} = (d_t, pd_t)'$ and

$$H(\theta) = \begin{bmatrix} H_d \\ (H_d - H_r)(I - \rho F)^{-1} F \end{bmatrix},$$

with $H_d = (1, 0, 1, 0, 0_{2\times(p-1)})$ and $H_r = (0, 1, 0, 1, 0_{2\times(p-1)})$. The observation variables, dividend growth (d_t) and the log price–dividend ratio (pd_t), are from the CRSP market indices of NYSE/AMEX/NASDAQ. We follow Hansen et al.'s (2008) aggregation procedure to construct the quarterly real equity return and real dividend growth series for the period 1952–2011 (see Figure 1).[6] 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 we move on to the empirical results, a short discussion of how to interpret the unobserved components, d_t^p, r_t^p, d_t^τ, r_t^τ, might be useful. The market fundamentals interpretation of these components is that they reflect the predictable components in dividend growth (d_t^p, d_t^τ) and in returns (r_t^p, r_t^τ). 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 nonmarket fundamental effects on the log price–dividend ratio.[7]

https://doi.org/10.1017/S1365100513000588 Published online by Cambridge University Press

TABLE 1. Selected statistics of the posterior distribution of parameters for model with two permanent components

Parameter	Maximum of posterior probability function	Mean of posterior distribution	St. dev. of posterior distribution	5th percentile of posterior distribution	95th percentile of posterior distribution
$\phi_{d,1}$	0.33168	0.32766	0.066572	0.21769	0.43601
$\phi_{r,1}$	0.33531	0.022255	0.11652	−0.11624	0.23836
$\phi_{d,2}$	0.2229	0.22914	0.064297	0.12356	0.33448
$\phi_{r,2}$	0.30176	0.023044	0.10237	−0.0996	0.22675
$\mathrm{var}(\varepsilon_t^{dp})$	8.53E-05	0.017412	0.040751	2.77E-05	0.081385
$\mathrm{var}(\varepsilon_t^{d\tau})$	3.6304	3.778	0.43494	3.126	4.5319
$\mathrm{var}(\varepsilon_t^{rp})$	0.002087	0.019619	0.037781	0.001915	0.079824
$\mathrm{var}(\varepsilon_t^{r\tau})$	62.907	2575.3	3985.1	129.15	9337.9
$\mathrm{corr}(\varepsilon_t^{dp}, \varepsilon_t^{d\tau})$	−0.26609	−0.07817	0.64965	−0.99392	0.98655
$\mathrm{corr}(\varepsilon_t^{dp}, \varepsilon_t^{rp})$	−0.67979	0.55852	0.47759	−0.51583	0.98435
$\mathrm{corr}(\varepsilon_t^{dp}, \varepsilon_t^{r\tau})$	0.66205	−0.00138	0.50914	−0.82478	0.82
$\mathrm{corr}(\varepsilon_t^{d\tau}, \varepsilon_t^{rp})$	0.30062	0.14527	0.506	−0.86245	0.87093
$\mathrm{corr}(\varepsilon_t^{d\tau}, \varepsilon_t^{r\tau})$	0.029226	0.076433	0.27935	−0.33432	0.61116
$\mathrm{corr}(\varepsilon_t^{rp}, \varepsilon_t^{r\tau})$	−0.94262	0.007345	0.49364	−0.77055	0.77131
$\mathrm{var}(\varepsilon_t^{dp} - \varepsilon_t^{rp})$	0.002746	0.004402	0.000969	0.002977	0.006078

For the discussion of weak identification, it does not matter which interpretation one uses.

4. ESTIMATION RESULTS

Table 1 presents the mean, standard deviation, and 5th and 95th percentiles of the posterior distributions of the state-space model parameters as well as the maximum likelihood 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 means of the posterior distributions for both of the autoregressive coefficients for transitory returns are centered at zero. The sum of the AR(1) and AR(2) coefficients from the ML estimation is not significantly different from zero. The means of the posterior distributions for the AR(1) and AR(2) coefficients for the transitory dividend process are 0.32 and 0.23, respectively. One interesting result is that the MCMC sampler 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 that the AR(1) coefficient for returns is about 0.3 and the variance is small relative to the 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]

https://doi.org/10.1017/S1365100513000588 Published online by Cambridge University Press

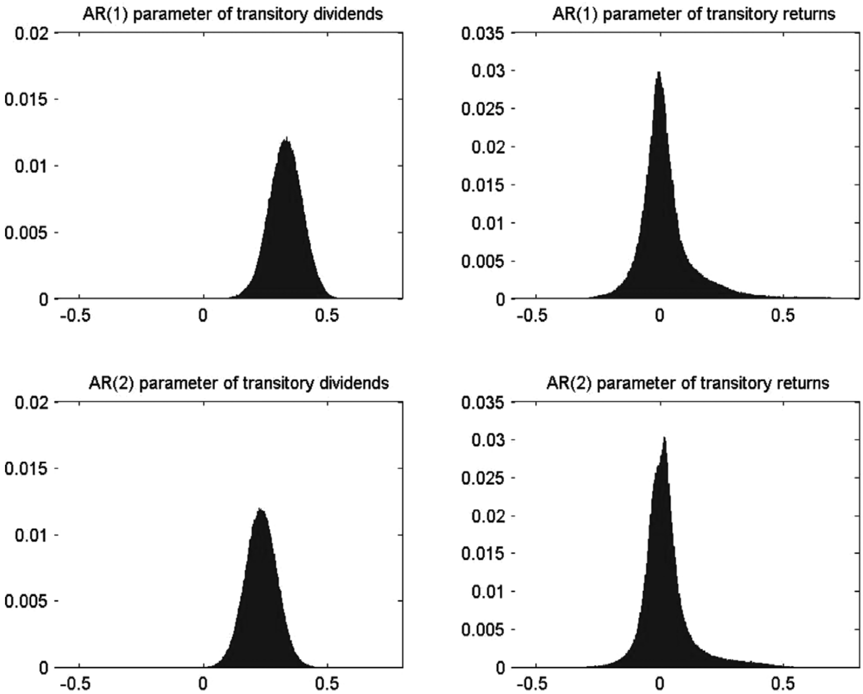

FIGURE 2. Histogram of posterior distribution of autoregressive parameters for model with permanent components in both dividends and returns.

Figure 3 displays posterior distributions 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 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).

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 compared more readily with that of $\text{var}(\varepsilon_t^{dp} - \varepsilon_t^{rp})$. Note that 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 are 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

https://doi.org/10.1017/S1365100513000588 Published online by Cambridge University Press

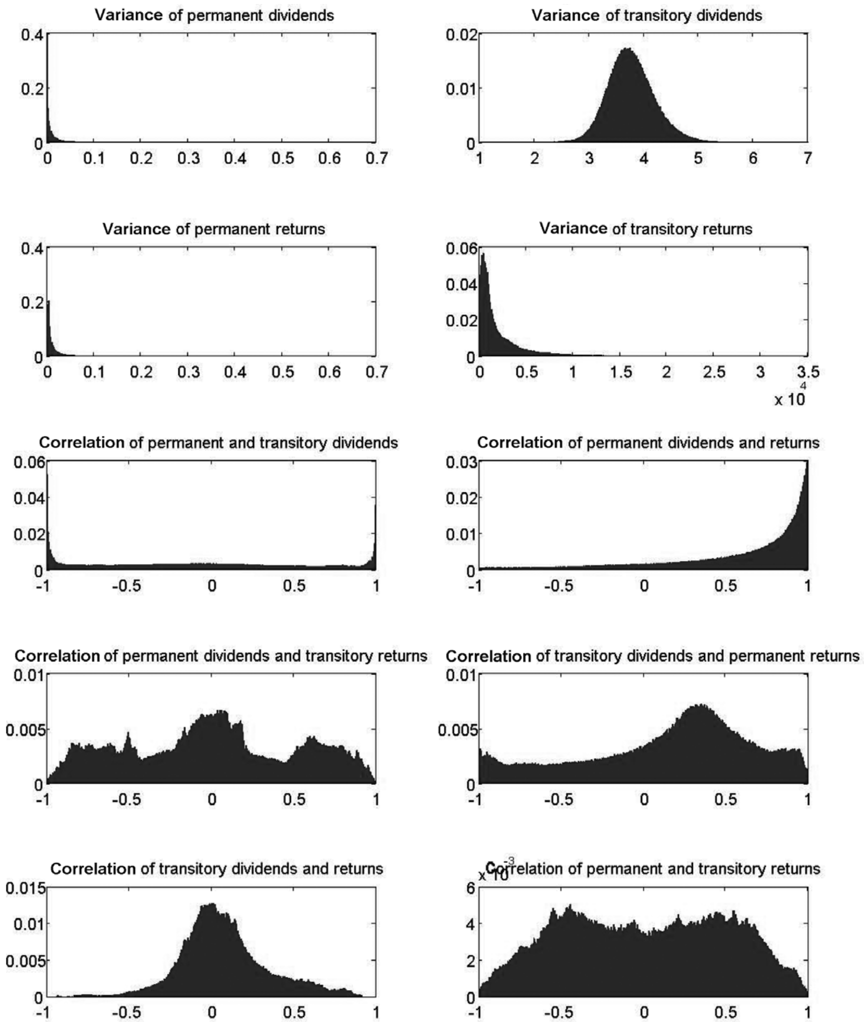

FIGURE 3. Posterior distributions of variances and correlations of innovations for state-space model with permanent components in dividend growth and returns.

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, whereas the bottom panel displays the contribution of expected future returns. Although 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

https://doi.org/10.1017/S1365100513000588 Published online by Cambridge University Press

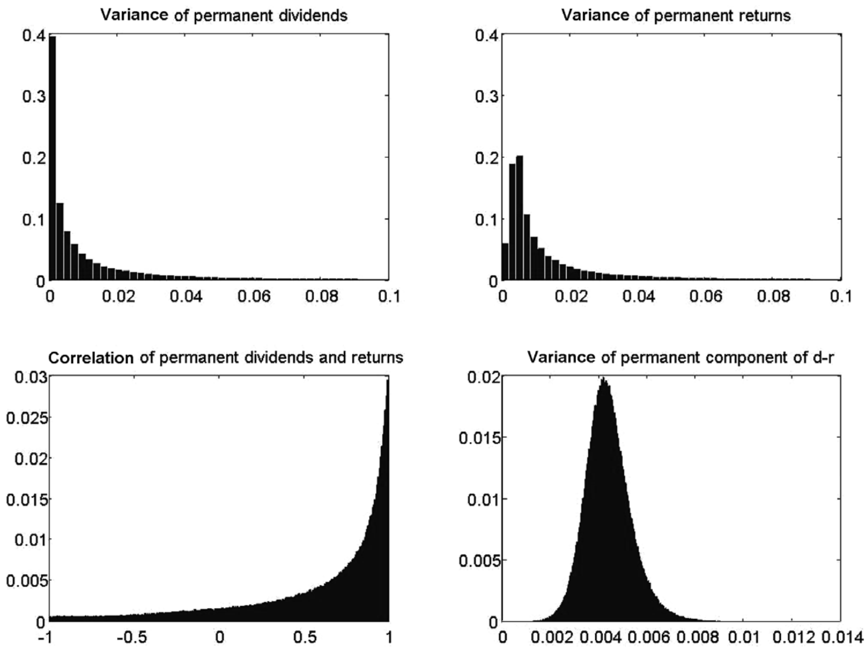

FIGURE 4. Posterior distribution of the components of var($\varepsilon_t^{dp} - \varepsilon_t^{rp}$).

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 contributions of expected future dividends and expected future returns are 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 relatively small autoregressive parameters and hence do not provide much information about future dividends or returns. Second, although 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 the fact that there was substantial uncertainty about the individual variances of (and correlation between) innovations in the permanent components but much

https://doi.org/10.1017/S1365100513000588 Published online by Cambridge University Press

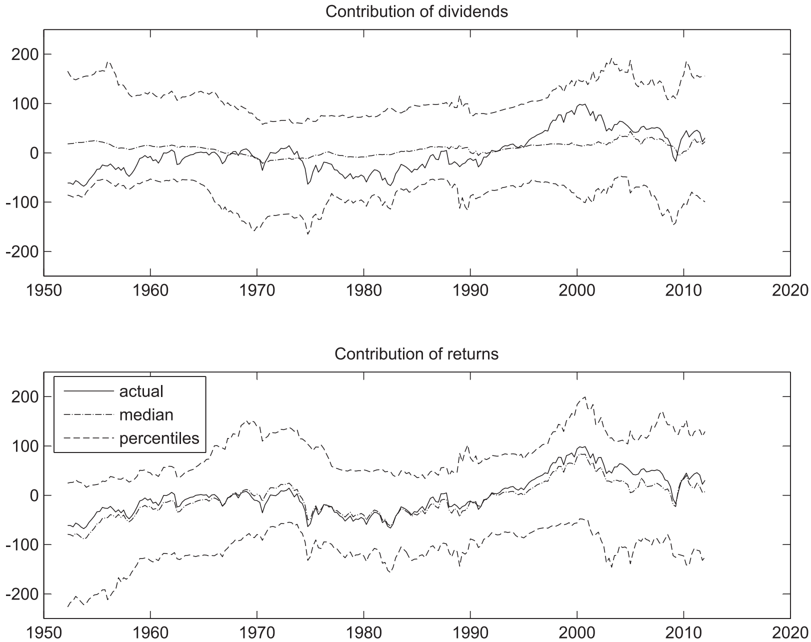

FIGURE 5. Historical decomposition of log price–dividend ratio: contribution of expected future dividends and expected future returns. Median and 5th and 95th percentiles of the posterior distribution.

less uncertainty about the value of $\text{var}(\varepsilon_t^{dp} - \varepsilon_t^{rp})$. Overall, Figure 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 more 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 and the 5th and 95th percentiles of the permanent component for dividend growth and returns. What stands out from Figure 7 is that the permanent components for both dividend growth and returns are 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 for the relative contributions of dividends and returns to fluctuations in the price–dividend ratio. Because the

https://doi.org/10.1017/S1365100513000588 Published online by Cambridge University Press

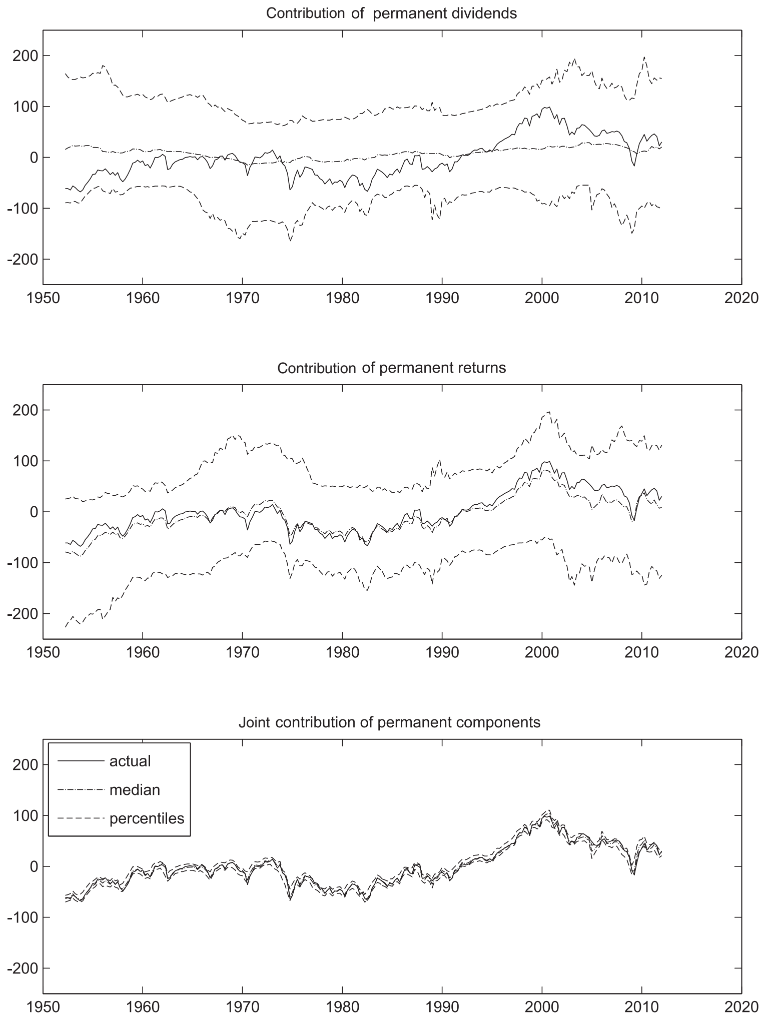

FIGURE 6. Contribution of permanent components to log price–dividend ratio. Median and 5th and 95th percentiles of the posterior distribution.

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.

https://doi.org/10.1017/S1365100513000588 Published online by Cambridge University Press

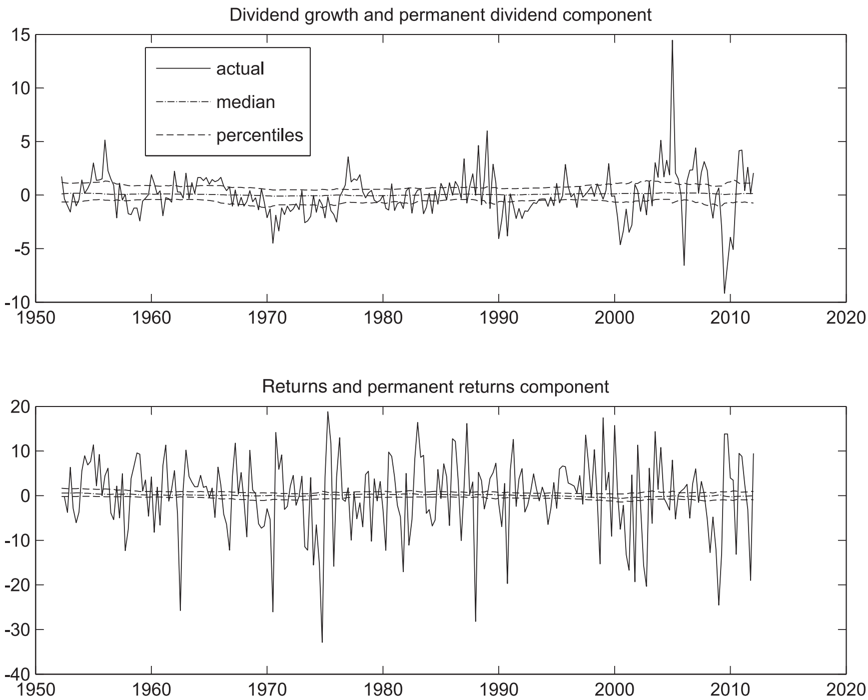

FIGURE 7. Actual dividend growth and returns along with permanent components of dividend growth and returns: median and 5th and 95th percentiles.

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 $\mathrm{var}(\varepsilon_t^{dp})$ for the model with no permanent component in returns looks more like the posterior distribution of $\mathrm{var}(\varepsilon_t^{dp} - \varepsilon_t^{rp})$ 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.

https://doi.org/10.1017/S1365100513000588 Published online by Cambridge University Press

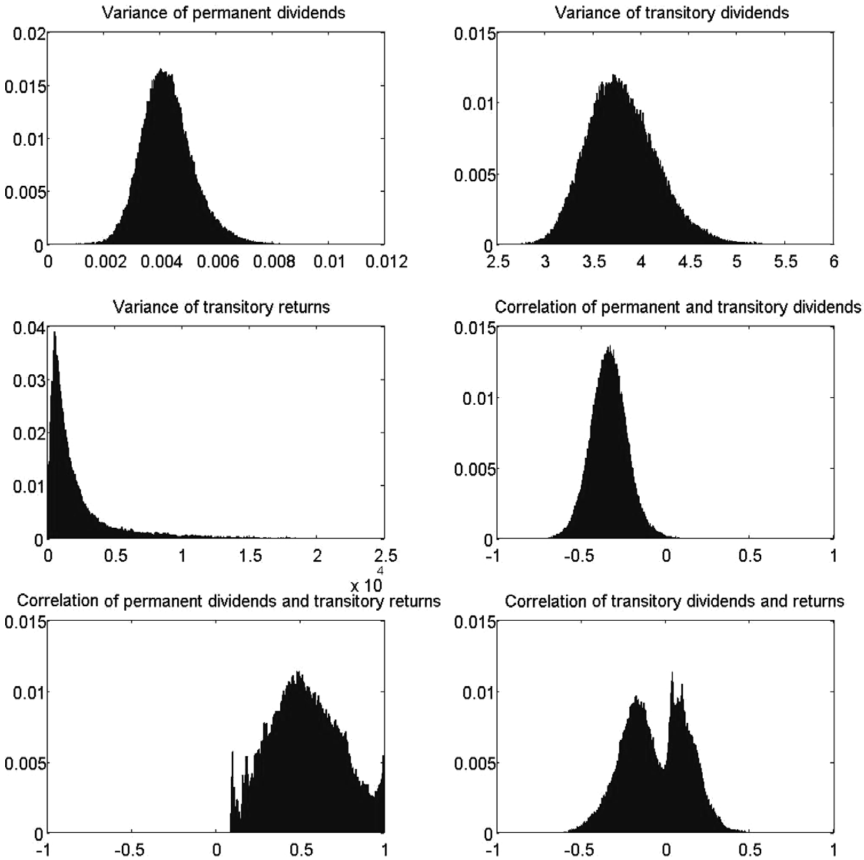

FIGURE 8. Posterior distribution of variances and correlations of innovations in state-space model with a permanent component in dividend growth but no permanent component in returns.

Similarly, the posterior distribution of innovations in the single permanent component [in Figure 9,var(ε_t^{rp})] looks much like the posterior distribution of var($\varepsilon_t^{dp} - \varepsilon_t^{rp}$) in the model with two permanent components.

Although both models' parameters appear to be estimated with precision, which of the two models do the data favor? Figure 10 plots the cumulative density functions 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 view these models as almost equally likely.

Although the data do not appear to favor one of the models over the other, the two models have very different implications for stock price decomposition.

https://doi.org/10.1017/S1365100513000588 Published online by Cambridge University Press

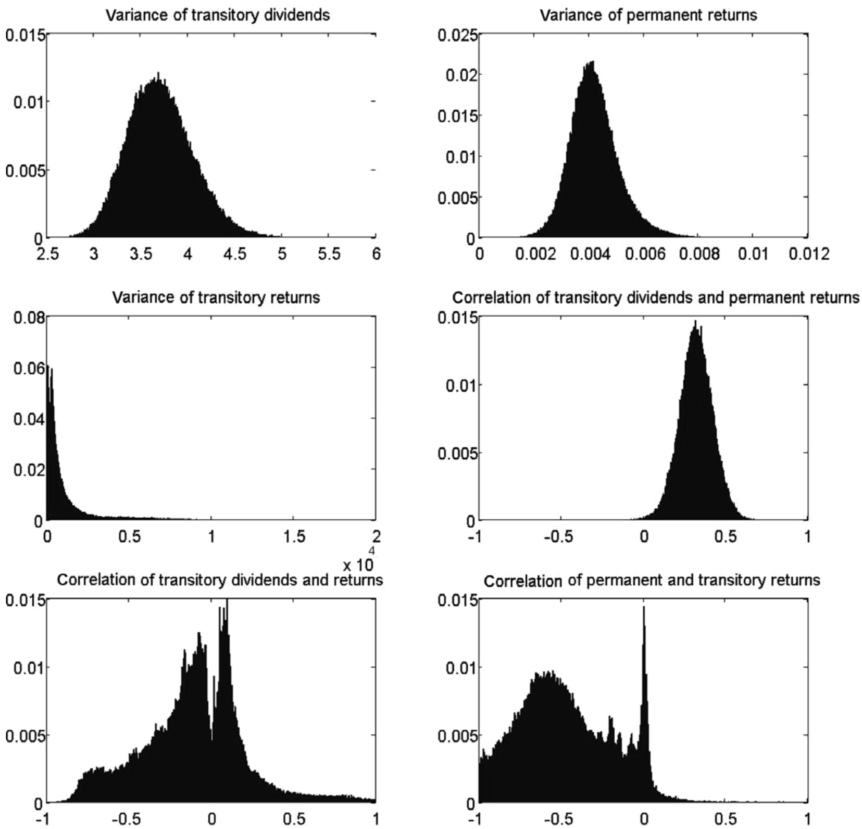

FIGURE 9. Posterior distribution of variances and correlations of innovations in a state-space model with a permanent component in returns but no permanent component in dividend growth.

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 ratio. For the model with a permanent returns component and no permanent component for dividends, it is the expectations of future returns that explain 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

https://doi.org/10.1017/S1365100513000588 Published online by Cambridge University Press

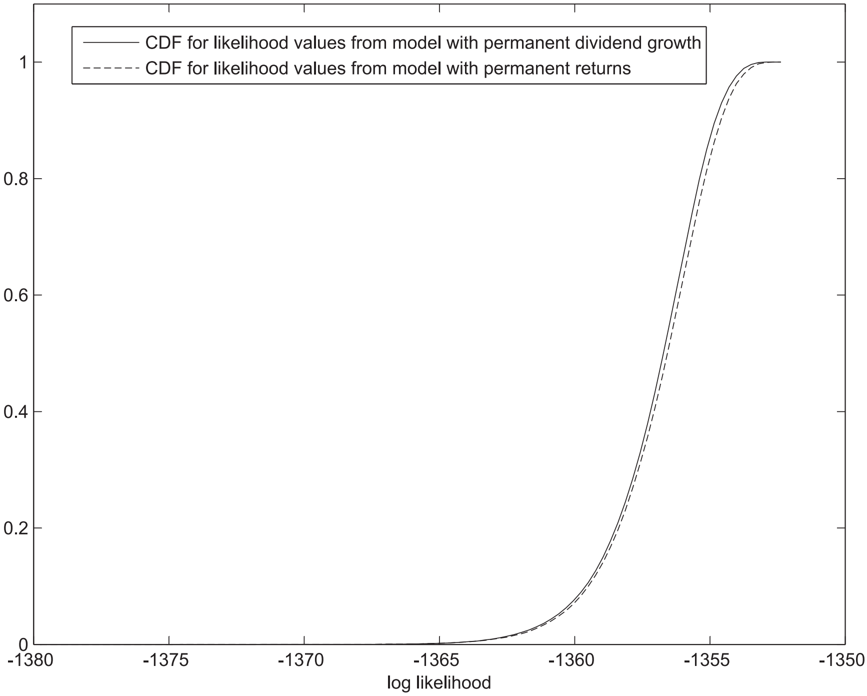

FIGURE 10. Cumulative density function for likelihood values from MCMC draws for models with a single permanent component.

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), in returns alone (middle panel), or in both dividends and returns (bottom panel). Yet, when one takes account of uncertainty across models, the data are 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 preceding sections essentially applied to the market portfolio. As expected returns largely reflect 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.[10] In terms of decompositions of the price–dividend ratio,

https://doi.org/10.1017/S1365100513000588 Published online by Cambridge University Press

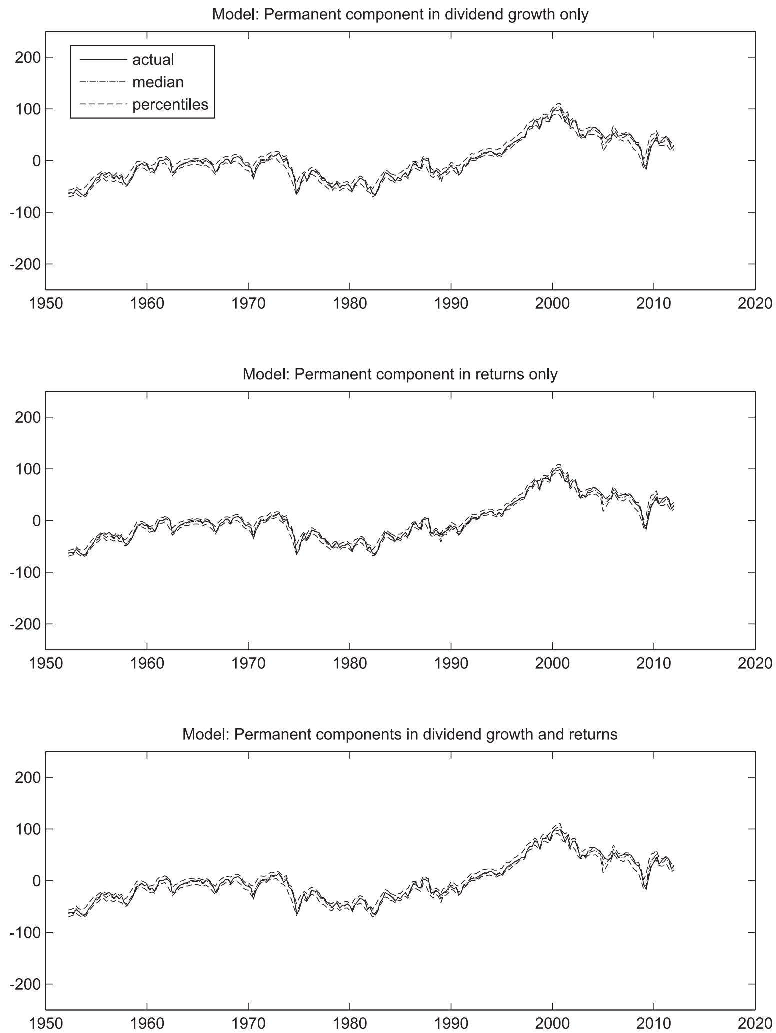

FIGURE 11. Contribution of permanent components to log price–dividend. Median and 5th and 95th percentiles of the posterior distribution.

however, it is not clear whether examining disaggregated portfolios helps to solve the weak identification problem pointed out in the preceding sections. Although 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

https://doi.org/10.1017/S1365100513000588 Published online by Cambridge University Press

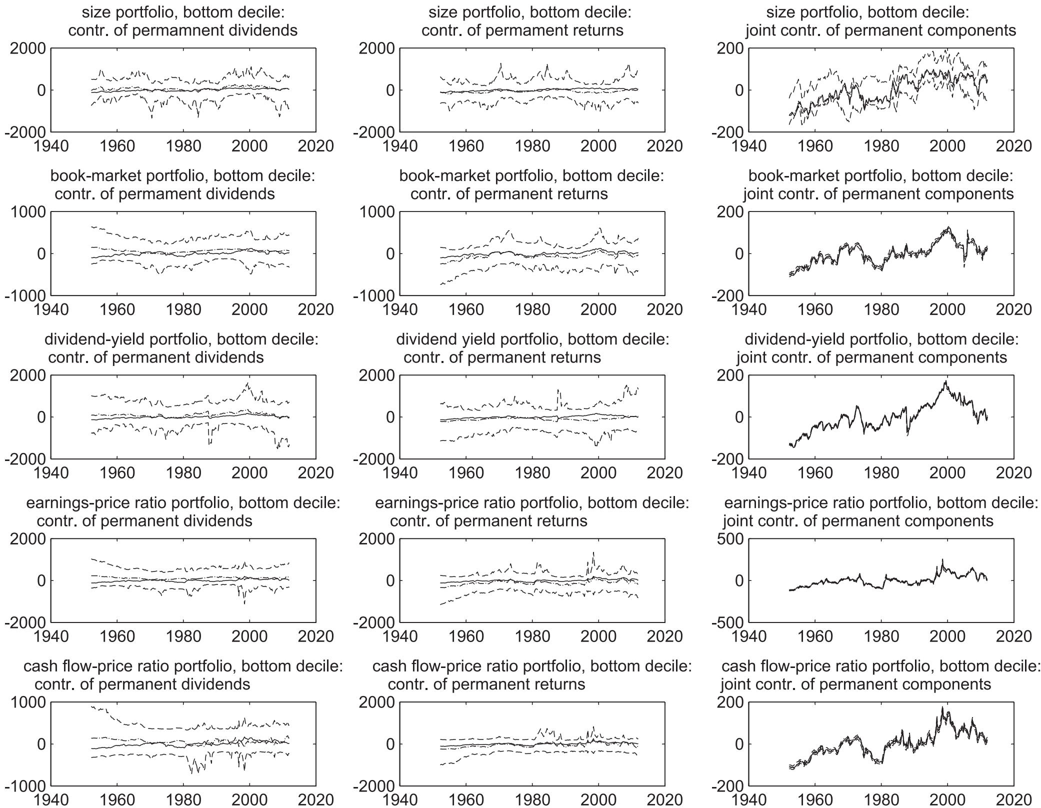

FIGURE 12. Decomposition of log price–dividend ratio disaggregated Fama–French portfolios: bottom deciles.

https://doi.org/10.1017/S1365100513000588 Published online by Cambridge University Press

https://doi.org/10.1017/S1365100513000588 Published online by Cambridge University Press

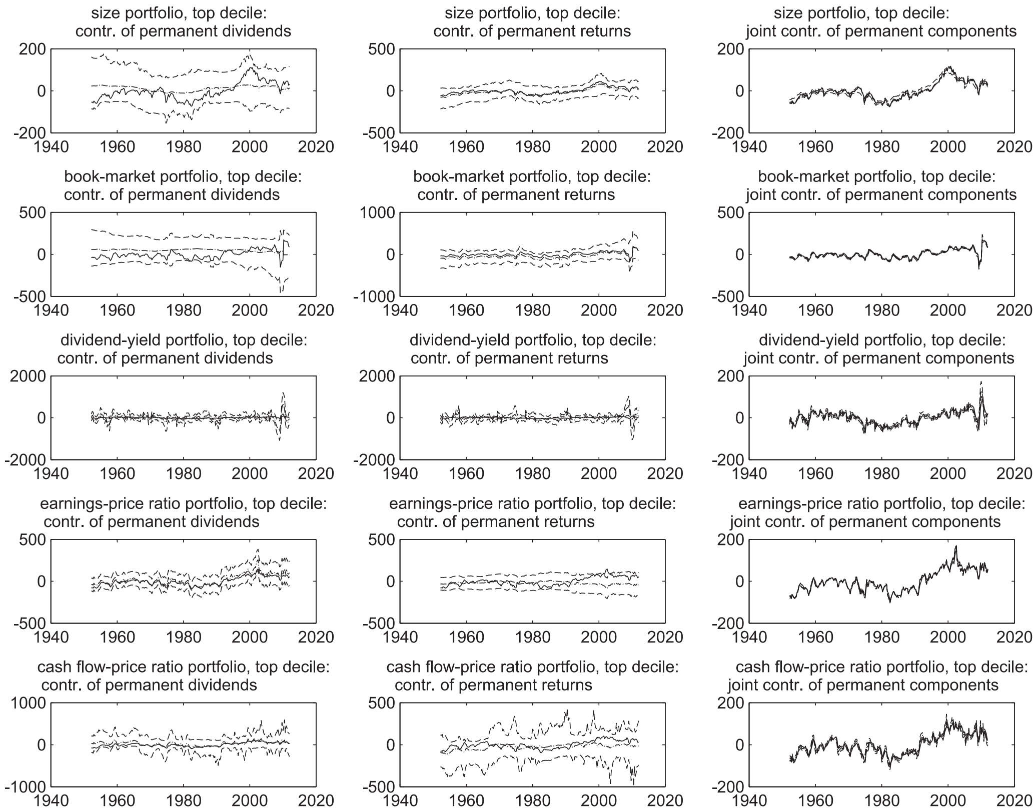

FIGURE 13. Decomposition of log price–dividend ratio disaggregated Fama–French portfolios: top deciles.

portfolios, we apply our Bayesian MCMC analysis 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.[11] 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 for the aggregate market portfolio analyzed in preceding sections. We employ the model that allows for a permanent component in both expected dividend growth and expected returns.

Figure 12 presents the results for the bottom decile, whereas 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, and the second column presents the contribution of permanent returns, whereas the third column report the results for the joint contribution of permanent dividend growth and permanent returns. Similarly to the aggregate results, reported earlier in the paper, the contributions of permanent dividend growth and permanent returns are estimated imprecisely individually. However, the joint contribution of permanent dividend growth and permanent returns is estimated very precisely. We are thus unable to determine the relative contribution of each component individually.

7. CONCLUSION

The relative contributions of expected returns and expected cash flows to movements in the price–dividend ratio have 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 that 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 infer 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 the size of the variances of innovations relative to the low-frequency component of dividends or returns. What the data support 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.

https://doi.org/10.1017/S1365100513000588 Published online by Cambridge University Press

NOTES

1. Cuthbertson et al. (1999) find that similar results hold for the U.K. stock market.

2. For the VAR literature see Campbell and Shiller (1988a, 1988b), Campbell (1991), Cochrane (1992, 2008), Shiller and Beltratti (1992), and Campbell and Ammer (1993). Note, however, that Larrain and Yogo (2008) find when a 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.

3. For the state-space approach see Balke and Wohar (2002), Binsbergen and Koijen (2010), and Ma and Wohar (2012).

4. 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.

5. Ma et al. (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.

6. Hansen, Heaton, and Li's data sources and procedures for computing return and dividend series are available on Nan Li's website: http://www.bschool.nus.edu.sg/staff/biznl/.

7. Balke and Wohar (2009) consider the case of an explicitly nonfundamental component in the log price–dividend ratio in the form of a bursting and reoccurring rational bubble. This differs from the expected future returns components in that a rational bubble imposes fairly strong restrictions on the stochastic process governing the bubble component.

8. The posterior distribution is based on a 750,000 sample of every fifth draw from the Metropolis–Hastings (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 $N(0, 1{,}000)$ and the variances of innovations in (15) are distributed $U(0, 10{,}000)$, whereas 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 in an inferior mode, as the chain traversed a large part of the parameter space, especially for the correlations between innovations.

10. Cochrane (2008) argues that much of the cash flow variance is idiosyncratic in nature, whereas the expected return variance is common (i.e., related to systematic risk). Studies by Vuolteenaho (2002), Callen and Segal (2004), and Chen et al. (2013) have found that cash flow news is the primary factor explaining the variance in stock returns at the firm level, whereas discount rate news explains most of the movements in stock returns at the aggregate level.

11. The link to the website is http://mba.tuck.dartmouth.edu/pages/faculty/ken.french/data_library.html.

12. Formally, for each of the models the individual autoregressive parameters have a prior distribution of the joint truncated normal $N(0, 1{,}000)$ and the variances of innovations in (15) are distributed $U(0, 10{,}000)$ whereas 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 these data, the acceptance in the Metropolis–Hastings sampler depends only on the likelihoods.

REFERENCES

Balke, Nathan S. and Mark E. Wohar (2002) Low frequency movements in stock prices: A state-space decomposition. *Review of Economics and Statistics* 84, 649–667.

Balke, Nathan S. and Mark E. Wohar (2006) What drives stock prices? Identifying the determinants of stock price movements. *Southern Economic Journal* 74, 55–78.

Balke, Nathan S. and Mark E. Wohar (2009) Market fundamentals versus rational bubbles in stock prices: A Bayesian perspective. *Journal of Applied Econometrics* 24, 35–75.

https://doi.org/10.1017/S1365100513000588 Published online by Cambridge University Press

Bansal, Ravi and Amir Yaron (2004) Risks for the long-run: A potential resolution of the asset pricing puzzles. *Journal of Finance* 59, 1481–1509.

Binsbergen, Jules H. van and Ralph S.J. Koijen (2010) Predictive regressions: A present-value approach. *Journal of Finance* 65, 1439–1471.

Callen, Jeffrey L. and Dan Segal (2004) Do accruals drive firm-level stock returns? A variance decomposition analysis. *Journal of Accounting Research* 43, 377–412.

Campbell, John Y. (1991) A variance decomposition for stock returns. *Economic Journal* 101, 157–179.

Campbell, John Y. and John Ammer (1993) What moves the stock and bond markets? A variance decomposition for long-term asset returns. *Journal of Finance* 48, 3–37.

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

Campbell, John Y. and Robert Shiller (1988a) The dividend-price ratio and expectations of future dividends and discount factors. *Review of Financial Studies* 1, 195–228.

Campbell, John Y. and Robert Shiller (1988b) Stock prices, earnings, and expected dividends. *Journal of Finance* 43, 661–676.

Carter, Christopher K. and Robert Kohn (1994) On Gibbs sampling for state-space models. *Biometrika* 81, 541–553.

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

Chen, Long, Zhi Da, and Xinlei Zhao (2013) What drives stock price movements? *Review of Financial Studies* 26, 841–876.

Chen, Long and Xinlei Zhao (2009) Return decomposition. *Review of Financial Studies* 22, 5213–5249.

Cochrane, John (1992) Explaining the variance of price-dividend ratios. *Review of Financial Studies* 5, 243–280.

Cochrane, John (2008) The dog that did not bark: A defense of return predictability. *Review of Financial Studies* 21, 1533–1575.

Cuthbertson, Keith, Simon Haves, and Dirk Nitzsche (1999) Explaining movements in UK stock prices. *Quarterly Review of Economics and Finance* 39, 1–19.

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

Hansen, Lars P., John C. Heaton, and Nan Li (2008) Consumption strikes back? Measuring long-run risk. *Journal of Political Economy* 116, 260–302.

Kim, Chang-Jin and Charles R. Nelson (1999) *State-Space Models with Regime Switching: Classical and Gibbs-Sampling Approaches with Applications.* Cambridge, MA: MIT Press.

Larrain, Borja and Motohiro Yogo (2008) Does firm value move too much to be justified by subsequent changes in cash flow? *Journal of Financial Economics* 87, 200–226.

Ma, Jun (2013) Long-run risks and its implications for the equity premium puzzle: New evidence from a multivariate framework. *Journal of Money, Credit and Banking* 45, 121–145.

Ma, Jun and Charles R. Nelson (2012) The Superiority of the LM test in a Class of Models Where the Wald Test Performs Poorly; Including Nonlinear Regression, ARMA, GARCH, and Unobserved Components. Working paper, University of Alabama and University of Washington.

Ma, Jun, Charles R. Nelson, and Richard Startz (2007) Spurious inference in the GARCH(1,1) model when it is weakly identified. *Studies in Nonlinear Dynamics and Econometrics* 11, 1–27.

Ma, Jun and Mark E. Wohar (2012) Expected Returns and Expected Dividend Growth: Time to Rethink an Established Empirical Literature. Working paper, University of Alabama and University of Nebraska at Omaha.

Morley, James, Charles R. Nelson, and Eric Zivot (2003) Why are unobserved component and Beveridge-Nelson trend-cycle decompositions of GDP so different? *Review of Economics and Statistics* 85, 235–243.

Nelson, Charles R. and Myung J. Kim (1993) Predictable stock returns: The role of small sample bias. *Journal of Finance* 48, 641–661.

https://doi.org/10.1017/S1365100513000588 Published online by Cambridge University Press

Nelson, Charles R. and G. William Schwert (1977) On testing the hypothesis that the real rate of interest is constant. *American Economic Review* 67, 478–486.
Nelson, Charles R. and Richard Startz (2007) The zero-information-limit condition and spurious inference in weakly identified models. *Journal of Econometrics* 138, 47–62.
Shiller, Robert and Andrea Beltratti (1992) Stock prices and bond yields. *Journal of Monetary Economics* 30, 25–46.
Vuolteenaho, Tuomo (2002) What drives firm-level stock returns? *Journal of Finance* 43, 233–264.

APPENDIX: MCMC ALGORITHM

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

$$\log(L(\mathbf{Y}_T, \theta)) = \sum_{t=1}^{T} \{-0.5\log[\det(H(\theta)' P_{t|t-1} H(\theta))]$$
$$-0.5[Y_t - H(\theta)S_{t|t-1}]'[H(\theta)' P_{t|t-1} H(\theta)]^{-1}[Y_t - H(\theta)S_{t|t-1}]\}, \quad \textbf{(A.1)}$$

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, $P(\theta|\mathbf{Y}_T)$, is

$$P(\theta|\mathbf{Y}_T) \propto L(\mathbf{Y}_T, \theta)h(\theta). \quad \textbf{(A.2)}$$

In this Appendix, we consider the case of very diffuse priors for which the likelihood function is the principal determinant of the posterior distribution.[12]

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, θ. 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 θ^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, $\theta^{(i)}$:
$\theta^{(i)} = \theta^c$ with probability $\alpha[\theta^c, \theta^{(i-1)}]$;
$\theta^{(i)} = \theta^{(i-1)}$ 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 ν is drawn from a multivariate t-distribution with 50 degrees of freedom and a covariance matrix Σ. We set Σ to be a scaled value of the Hessian matrix of $-\log[L(Y_T, \theta)]$ evaluated at the maximum likelihood estimates. We

https://doi.org/10.1017/S1365100513000588 Published online by Cambridge University Press

choose the scaling so that between 25 and 40% 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.

https://doi.org/10.1017/S1365100513000588 Published online by Cambridge University Press