Asset Prices with Multifrequency Regime-Switching and Learning: A Volatility Feedback Specification[*]

Laurent E. Calvet and Adlai J. Fisher

This version: March 21, 2004

Abstract

This paper develops a Markov-switching asset pricing economy with Epstein-Zin consumers and regime shifts in the mean and standard deviation of dividend growth. We show how to filter beliefs and solve for equilibrium asset prices under different learning environments even when the state space is very large. In an empirical application, we specialize to a volatility feedback setting where the mean of dividend news growth is constant, but volatility is stochastic and subject to shocks of heterogeneous durations. This provides a parsimonious structural econometric model for the time-series of asset returns, where skewness and excess kurtosis are endogenous. The likelihood has closed form under two learning environments of special interest. We find that relatively large numbers of volatility components give the best fit to the data, and in a comparison with the classic Campbell and Hentschel (1992) specification, the new model dominates. The unconditional feedback effect is up to ten times larger than in previous literature, and ex post conditional feedback can be much more concentrated in time. We also explore the learning implications of the model and identify a tradeoff between skewness and kurtosis as the volatility information available to investors increases. Economies with intermediate levels of information best match the data.

1. Introduction

Over the past fifteen years, the asset pricing literature has embraced a new and useful tool in Markov regime-switching. Researchers have used these discrete dynamics to help explain a number of phenomena including stock market volatility, return predictability, the relation between conditional risk and return, the shape of the yield curve, and the recent growth of the stock market.[1]

Pricing models with stochastic regime-shifts typically assume a small number of states, usually two or four. This partly stems from the view that switching transitions affect only low-frequency variations. Correspondingly, when researchers bring regime-shifts into an asset pricing framework, they relate their models to return data at monthly, quarterly or yearly intervals, and often confine the analysis to a comparison of stylized facts. At higher frequencies such as daily, the existing formulations offer limited realism and empirical usefulness.[2] Thus, while contributions related to this new tool are impressive, the full potential of regime-switching in asset pricing settings has likely not yet been realized.

Our paper develops a theoretically tractable and empirically useful asset pricing framework for Markov-switching processes with a large number of discrete latent states. The setup can be applied to either high or low frequency data, and adapts easily to different information structures or learning environments. Convenient exact solutions are available for equilibrium prices, return dynamics, and filtered *ex ante* and *ex post* beliefs. These methods remain computationally practical with several hundred states. We also develop a closed form loglinearized solution that provides economic intuition and a simple empirical alternative to exact computations.

To achieve these objectives, we must address manifestations of the "curse of dimensionality" associated with large state spaces. General regime-switching formulations require the number of parameters to grow quadratically with the cardinality of the state space. This poses obvious difficulties in working with more than a few regimes. We adopt a solution that has recently been developed in the econometrics literature (Calvet and Fisher, 2001, 2002, 2004). Markov-switching Multifractal (MSM) processes

---

[1] Applications of regime-switching to asset pricing include Abel (1994, 1999), Cagetti, Hansen, Sargent and Williams (2002), Cecchetti, Lam and Mark (1990), David (1997), Kandel and Stambaugh (1990, 1991), Lettau, Ludvigson and Wachter (2003), Turner, Startz and Nelson (1989), Veronesi (1999, 2000, 2002), Wachter (2002), and Whitelaw (2000). Applications that focus primarily on the term-structure of interest rates include Bansal, Tauchen and Zhou (2003), Bansal and Zhou (2002), Dai, Singleton, and Yang (2003), and Naik and Lee (1997).

[2] Econometricians have developed hybrid specifications such as Markov-switching GARCH (Cai, 1994; Hamilton and Susmel, 1994; Gray, 1996) and Markov-switching stochastic volatility (So, Lam, and Li, 1998; Smith, 2002) that combine regime-switching at low frequencies with alternative dynamics at higher frequencies. These can achieve reasonable fit to daily data, but their goal is generally statistical moreso than developing asset pricing implications.

are characterized by arbitrarily large state spaces and a small number of parameters. Exogenous shocks have heterogeneous durations ranging from one day to more than a decade. The model remains parsimonious because the shocks have identical marginals and the frequencies are tightly parameterized by an exponential progression. The assumed heterogeneity in news duration is consistent with economic intuition about multiple sources of fundamental news including liquidity changes, earnings cycles, business cycles, technology innovations, and demographics. MSM aggregates conveniently, allowing researchers to estimate a model on daily data and still analyze a problem at a longer horizon. The process provides a realistic description of financial data. It capture the outliers, volatility persistence and scaling of financial series, and substantially outperforms specifications such as $GARCH(1, 1)$ and Markov-switching GARCH that are known for their excellent performance in volatility forecasting (Calvet and Fisher, 2004). Finally, the model permits maximum likelihood estimation and analytical forecasting at multi-step horizons. MSM thus provides many of the econometric conveniences of standard GARCH formulations, but in a parsimonious, multifrequency, stochastic volatility setting that matches the data. It is thus natural to now embed MSM in an asset-pricing framework.

We begin with a standard Markov-switching economy. An isoelastic Epstein-Zin consumer receives an exogenous consumption stream, and prices the dividend flow provided by a stock. Dividend news growth follows a conditional lognormal path with Markov regime-shifts in drift and volatility. The imperfect correlation between consumption and dividends, which has been widely noted in the literature (e.g. Campbell and Cochrane, 1999), permits us to avoid another instance of the curse of dimensionality that is normally present in Epstein-Zin economies with learning. If consumption is identical to dividends, the stock return impacts the stochastic discount factor, and in an Epstein-Zin economy with learning the price:dividend ratio is a non-linear function over the entire simplex of beliefs. Our setup however implies that the Euler equation is linear in returns, and as a result, the price:dividend ratio is linear in investor beliefs.

We develop an empirical application of our model to volatility feedback. Exogenous changes in the volatility of dividend news have long been proposed as possible explanation for the large movements exhibited by equity returns (e.g. Pindyck, 1984; Poterba and Summers, 1985; Barsky, 1989; Abel, 1988; Campbell and Hentschel, 1992). We specialize our framework by assuming that dividend news has a constant mean and a volatility that is hit by shocks of heterogeneous frequencies. The model generates skewness and predictive asymmetry in returns, which are purely endogenous since the dividend news is conditionally Gaussian. To develop intuition, we loglinearize the Euler equation and compute an approximation to the price:dividend ratio and returns, as in Campbell and Shiller (1988) and Campbell and Viceira (2002). We can then characterize the magnitude of the volatility feedback and the sensitivity of the price:dividend

2

ratio to shock persistence.

In a Lucas (1978) tree economy with isoelastic expected utility, increased volatility reduces the price of the consumption stream only when the relative risk aversion is less than one (e.g. Abel, 1989; Whitelaw, 2000). This again relates to the fact that volatility affects not only the distribution of dividends, but also the pricing kernel and in particular the risk-free rate through the precautionary savings motive. In Epstein-Zin tree economies, feedback of the expected sign requires that the elasticity of intertemporal substitution be larger than unity (Lettau, Ludvigson, and Wachter, 2003). The inverse relation between volatility and price thus critically depends on preference restrictions that are empirically questionable. In our model, dividend news volatility does not affect consumption growth and thus has no impact on the pricing kernel. We therefore obtain a constant risk-free rate and the desired volatility feedback for all preference parameters.

We estimate by maximum likelihood the full information economy, in which investors directly observe the volatility state vector. The data consist of daily excess returns on the value-weighted CRSP index over the period 1962-2003. We estimate models with 1 to 8 components and a corresponding number of states ranging from 2^1 to 2^8. Specifications with between 6 and 8 frequencies provide statistically significant improvements in likelihood relative to models with smaller state spaces. Moreover, all models with three or more components dominate the classic Campbell and Hentschel (1992, hereafter "CH") specification based on a QGARCH(1,2) dividend news process, even though the multifrequency specification has fewer parameters. These results support the validity of our multifrequency approach.

The estimated full information process generates substantially larger feedback than previous research. Using their estimated daily process, CH report unconditional return variance that is approximately 2% larger than dividend news variance. In our specifications, feedback increases almost monotonically with the number of components and the likelihood function. Unconditional feedback is around 2% for specifications with a small number of volatility frequencies, increasing to between 10% and 20% for the preferred specifications with six to eight components. We thus obtain unconditional feedback effects that are 5 to 10 times larger than in previous literature.[3]

We analyze the unconditional moments of the full information regime-switching model and compare it with the CH specification. Using simulation methods, we find that although both models generate some degree of endogenous skewness, neither is likely to produce data that captures the first or third moments of actual returns.[4] While the

---

[3] Wu (2001) claims large volatility feedback effects based on a graphical depiction. Careful analysis of his results however shows that unconditional volatility feedback is 3.5% for his model estimated on monthly data, and is actually negative for his model estimated on weekly data. He does not estimate a model on daily data.

[4] Bias in the first moment is related to bias in the third moment, as discussed in the empirical section.

regime-switching model fits both the second and fourth moments well, the CH specification does not match the fourth moment. These misspecifications are not surprising, since both models attempt to generate skewness and kurtosis through endogenous economic mechanisms rather than a purely statistical approach. Both approaches should be viewed as structural econometric efforts to fit return data. This is challenging because higher moments are not specifically controlled by individual parameters.

In our setup, these difficulties can be partially resolved by introducing incomplete investor information and learning. We first consider an extreme case of limited information, in which the investor observes only the dividend itself and then makes inferences about the volatility state. This specification is conveniently estimated by maximum likelihood. Holding parameter values constant, daily learning generates weaker kurtosis and stronger negative skewness than the full information economy. Reduced kurtosis stems from information about state changes filtering to the investor slowly through the learning process. Skewness becomes more negative because only a single signal is available (the return). Inference about the volatility state thereby becomes correlated with inference about the return innovation. This intuition is similar to Veronesi (1999) and Lettau, Ludvigson and Wachter (2003), but in our case arises with uncertainty about the conditional distribution restricted to the second moment. In this case, when dividend growth is extremely low, the bad news about dividends is amplified by the additional bad news that volatility has probably increased. Conversely, when dividend growth is very high, good news about dividends is mitigated by the bad news that volatility may have increased. Thus, volatility feedback amplifies exceptionally bad news and dampens exceptionally good news about dividends. While this setup is interesting from a theoretical point of view, our empirical analysis of the model shows that the effects are too extreme. Excessive strengthening of endogenous skewness and weakening of endogenous kurtosis causes fit to deteriorate relative to the full information economy.

These results suggest that intermediate information environments may achieve a better compromise between skewness and kurtosis. We thus assume that the agent receives dividend news at a higher frequency than the stock returns observed by the econometrician. The investor uses the intradaily dividend news to form a "realized volatility" statistic as in Schwert (1989) and Andersen and Bollerslev (1998). More frequent observation yields more precise inference, and a range of intermediate information levels can be achieved by altering the observation frequency of dividend news. In this setup the investor filtering problem and return simulation are straightforward, but likelihood calculation would require integrating over unobserved intradaily dividend news. We therefore estimate the model using simulated method of moments, and find that approximately ten intradaily observations produce levels of endogenous skewness and endogenous kurtosis that are empirically reasonable.

Section 2 presents the asset pricing model and the equilibrium solution for a general

4

Markov structure. Section 3 specializes to a volatility feedback setup and develops intuition on a loglinearized version of the model. In Section 4, empirical results are provided for economies with full information. Learning economies are investigated in Section 5. Unless stated otherwise, all proofs are in the Appendix.

1.1. Literature Review

The paper contributes to two related strands of the asset pricing literature. First, we propose an operational model of learning when the state space is very large. Our work is thus related to the asset pricing literature on incomplete symmetric information. While early work on learning delivers only transitory effects (e.g. Detemple, 1986; Dothan and Feldman, 1986; Gennotte, 1986; Timmermann, 1993),[5] researchers have recently explored the possibility of regime-switching in latent states, which leads agents to constantly revise their conditional beliefs. For instance, David (1997), Veronesi (1999, 2000, 2002), and Lettau, Ludvigson and Wachter (2003) consider economies in which the growth rate and standard deviation of dividend growth switches through time. These papers emphasize low-frequency effects.

Second, the paper contributes to theoretical research on volatility feedback. Temporal fluctuations in volatility have long been proposed as a possible explanation for the large movements exhibited by equity returns. Pindyck (1984) and Poterba and Summers (1985) explore these issues in a decision-theoretic framework. Investigation in a general equilibrium framework was pioneered by Barsky (1989) in a two-period setting and Abel (1988) in the dynamic case. The equilibrium implications of regime-switching in the consumption process were considered by Cecchetti, Lam and Mark (1990), Kandel and Stambaugh (1990) and Whitelaw (2000). A standing problem is that an increase in volatility reduces prices and returns only for special choices of the preference parameters. We solve this difficulty by separating the consumption and dividend processes.

Our work is also closely related to empirical research on volatility feedback. Pindyck (1984) attributes the decline of the US stock market in the seventies to the increased economic uncertainty associated with high inflation and oil shocks. Poterba and Summers (1985) emphasize the importance of volatility persistence for such dynamics. Using GARCH-type processes, French, Schwert and Stambaugh (1987) and Campbell and Hentschel (1992) show that ex-post returns are negatively affected by positive innovations in volatility. Kandel and Stambaugh (1990) and Bekaert and Wu (2000) provide further support of this hypothesis.[6]

Volatility feedback has been found to contribute little to the unconditional variance of returns. For instance, Campbell and Hentschel (1992) show that feedback amplifies

---

[5] An early solution to this problem is proposed by Detemple (1991), who considers uninformed agents with non-Gaussian priors in an economy in which the fundamentals are conditionally Gaussian.

[6] See also Schwert (1989).

by 2% the volatility of dividend news. They attribute this results to a property of GARCH processes, in which the volatility of volatility increases as the fourth power of the volatility level. Because of this, the model delivers a limited feedback effect when estimated on excess return series with large outliers such as in 1929 or 1987. CH emphasize this limitations of GARCH processes. We use a multifrequency stochastic volatility model to revisit the question and find evidence of substantially stronger feedback.

A common thread between the learning and volatility feedback literature is that shocks tend to have a single, or at most two frequencies. This necessitates that each model be specialized to the frequency of the empirical phenomena that it investigates. The volatility feedback literature thus considers daily, weekly, or monthly returns. By contrast, Veronesi (2002) calibrates to yearly returns and considers horizons ranging from twenty to two hundred years. Lettau, Ludvigson and Wachter (2003) similarly consider highly persistent shocks with durations of about a decade. In our paper, we argue that the disconnect in the literature between these various effects is due to limitations of current models, but does not originate in economic theory. We propose a unified framework in which high frequency phenomena (volatility feedback) and low frequency switches (business cycle or peso effects) can be jointly modeled.

2. An Asset Pricing Model with Regime-Switching Dividends

This section develops a discrete-time stock market economy with regime-shifts in the mean and volatility of dividend growth.

2.1. An Epstein-Zin Markov-Switching Economy

We consider an exchange economy defined in discrete time on the regular grid $t = 0, 1, 2, ..., \infty$. As in Epstein and Zin (1989) and Weil (1989), the representative agent has isoelastic recursive utility

$$U_t = \left\{ (1-\delta)C_t^{\frac{1-\alpha}{\theta}} + \delta[\mathbb{E}_t(U_{t+1}^{1-\alpha})]^{\frac{1}{\theta}} \right\}^{\frac{\theta}{1-\alpha}},$$

where α is the coefficient of relative risk aversion, ψ is the elasticity of intertemporal substitution (EIS), and $\theta = (1-\alpha)/(1-\psi^{-1})$. The agent receives an exogenous consumption stream $\{C_t\}$. The log-consumption $c_t = \ln C_t$ follows a random walk with constant drift and volatility:

$$c_t - c_{t-1} = g_c + \sigma_c \varepsilon_{c,t}. \tag{2.1}$$

The shocks $\{\varepsilon_{c,t}\}$ are IID $\mathcal{N}(0,1)$. This standard specification is consistent with the large empirical evidence that consumption growth is approximately IID in postwar US consumption data (e.g. Campbell, 2003).

6

One focus of the paper is to investigate how aggregate stock returns respond to the volatility of dividend news. The volatility feedback literature suggests that the price:dividend ratio should fall when dividends become more volatile. When the stock is a claim on aggregate consumption, the integration of this effect into a general equilibrium model is plagued by several technical difficulties. Fluctuations in dividend news imply counterfactually high volatility in interest rates. Perhaps more surprisingly, the desired volatility feedback only exists for specific values of the preference parameters. In the expected utility case ($\alpha = 1/\psi$), the price:dividend ratio $Q_t \equiv P_t/C_t$ $= \mathbb{E}_t \sum_{h=1}^{+\infty} \delta^h \left(C_{t+h}/C_t\right)^{1-\alpha}$ declines with volatility only if relative risk aversion is less than unity: $\alpha < 1$ (Barsky, 1989; Abel, 1988).[7] This restriction is of course inconsistent with a large body of empirical research, which reports estimates of risk aversion that are significantly larger than one.

In the Epstein-Zin utility case, an increase in volatility reduces prices only if the elasticity of intertemporal substitution is strictly larger than 1 and relative risk aversion differs from unity: $\psi > 1$ and $\alpha \neq 1$ (Lettau, Ludvigson and Wachter, 2003).[8] The empirical validity of the EIS restriction is questionable. For instance, Campbell and Mankiw (1989), Ludvigson (1999) and Campbell (2003) show that the EIS is small and in many cases statistically indistinguishable from zero, while Attanasio and Weber (1993) and Vissing-Jørgensen (2002) report estimates of ψ larger than 1. We find it unsatisfactory that volatility feedback should crucially depend on a preference parameter unrelated to risk aversion.

We solve this difficulty by assuming that the stock is not a simple claim on aggregate consumption. We consider instead a dividend process $d_t = \ln D_t$ following a random walk with Markov-switching drift and volatility:

$$d_t - d_{t-1} = \mu_d(M_t) - \frac{\sigma_d^2(M_t)}{2} + \sigma_d(M_t)\varepsilon_{d,t}.$$

The shocks $\varepsilon_{d,t}$ are IID $\mathcal{N}(0,1)$. The drift $\mu_d(M_t)$ and the volatility $\sigma_d(M_t)$ are deter-

---

[7]When future consumption becomes riskier, two opposite effects influence the price:dividend ratio

$$Q_t = \sum_{i=1}^{+\infty} \delta^i \left\{ \mathbb{E}_t \left[\left(\frac{C_{t+i}}{C_t}\right)^{-\alpha} \right] + Cov_t \left[\left(\frac{C_{t+i}}{C_t}\right)^{-\alpha} ; \frac{C_{t+i}}{C_t} \right] \right\}.$$

First, the covariances $Cov_t \left[\left(\frac{C_{t+i}}{C_t}\right)^{-\alpha} ; \frac{C_{t+i}}{C_t} \right]$ become more negative and push down the price:dividend ratio Q_t, as desired. Second, the precautionary motive increases the expected marginal utility of future consumption $\mathbb{E}_t \left[\left(\frac{C_{t+i}}{C_t}\right)^{-\alpha} \right]$ and the interest rate goes down, which tends to reduce Q_t. We can eliminate the second effect by disentangling consumption and the stock market.

[8]The Euler equation is then $Q_t^\theta = \delta^\theta \mathbb{E}_t \left[(C_{t+1}/C_t)^{1-\alpha} (1 + Q_{t+1})^\theta \right]$. When consumption growth is IID, the price dividend ratio is constant and satisfies $Q/(1+Q) = \delta \left\{ \mathbb{E} \left[(C_{t+1}/C_t)^{1-\alpha} \right] \right\}^{1/\theta}$. It decreases with volatility if $(1-\alpha)/\theta > 0$ or equivalently $\psi > 1$ and $\alpha \neq 1$.

mined by a state variable M_t, which is first-order Markov. The Itô term $\sigma_d^2(M_t)/2$ guarantees that conditional on M_t, the expected dividend growth $\mathbb{E}_t(D_t/D_{t-1}) = e^{\mu_d(M_t)}$ is only controlled by the drift term $\mu_d(M_t)$. The Gaussian noises $\varepsilon_{c,t}$ and $\varepsilon_{d,t}$ are assumed to be positively correlated and IID.

The approach separates the stock from the definition of the stochastic discount factor. While falling short of a full general equilibrium model, it uses aggregate consumption in the pricing kernel and then prices other securities. It is thus a special case of the Lucas asset pricing methodology. The separation of consumption and dividends, which is common in finance, is consistent with a variety of empirical facts.[9] First, the correlation between consumption and the stock market is generally small. In US data, the correlation between real consumption growth and real dividend growth is 0.05 at a quarterly frequency, and 0.25 at a 4-year horizon (Campbell, 2003). Second, aggregate consumption is smooth and not noticeably heteroskedastic. In contrast, the volatility of stock market returns is high and exhibits substantial fluctuations through time, and earlier research seems to confirms that dividend news share the same features (e.g. Campbell and Hentschel, 1992). Third, the disconnect between d_t and c_t is possible because corporate profits only account for only a small proportion of national income. For instance in US data, corporate profits and personal consumption respectively account for approximately 10% and 70% of national income over the period 1929-2002. Furthermore, the stock market accounts for only a small fraction of national wealth and thus has only limited effects on the volatility of aggregate consumption.

In applications, it will be convenient to assume that the Markov state M_t takes only a finite number of values $\{m^1, ..., m^d\}$. The process M_t is then a Markov-chain specified by a transition matrix $A = (a_{ij})_{1 \leq i,j \leq d}$, where $a_{ij} = \mathbb{P}(M_{t+1} = m^j | M_t = m^i)$ for all i, j. The exact specification of the drift, volatility and transition matrix remains fully general in the rest of the section. We will introduce in Section 3 a special, high-dimensional specification that will be useful for empirical applications.

2.2. Asset Pricing under Complete Information

We easily check in the Appendix that the stochastic discount factor can be written as

$$SDF_{t+1} = \delta\{\mathbb{E}[(C_{t+1}/C_t)^{1-\alpha}]\}^{\frac{1}{\theta}-1}\left(\frac{C_{t+1}}{C_t}\right)^{-\alpha}. \tag{2.2}$$

This expression is proportional to the stochastic discount factor obtained under expected utility ($\theta = 1$). This suggests that the elasticity of intertemporal substitution affects the interest rate but has no impact on the price of risk. The simple interest rate $1 + R_{ft} = 1/\mathbb{E}_t(SDF_{t+1})$ is constant through time, and the logarithmic transform $r_f = \ln(1 + R_{ft})$ satisfies the familiar relationship: $r_f = -\ln\delta + \alpha g_c - \frac{(\alpha\sigma_c)^2}{2} +$

---

[9]See for instance Campbell and Cochrane (1999).

$(1 - \theta^{-1}) \left[(1 - \alpha)g_c + \frac{(1-\alpha)^2 \sigma_c^2}{2} \right]$. The interest rate is high when agents are impatient or expect a high consumption growth.

The information available to the investor is one of the major variables of the model. To develop intuition, we begin the analysis by considering that the agent directly observe the true state of the economy M_t. This will be the case if agents observe the macroeconomic quantities determining the state or obtain M_t by engaging into fundamental research. The economy with information set $I_t = \{M_s; s \leq t\}$ is a useful benchmark, which will be called *full information* case. The econometrician has a smaller information set $I_t^0 \subset I_t$, which will be typically limited to stock returns.[10]

In equilibrium, the stock price is proportional to the current dividend: $P_t = Q(M_t)D_t$, and the price:dividend ratio $Q(M_t)$ is determined by the volatility state M_t. The gross return on the stock is given by

$$1 + R_{t+1} \equiv \frac{D_{t+1} + P_{t+1}}{P_t} = \frac{D_{t+1}}{D_t} \frac{1 + Q(M_{t+1})}{Q(M_t)}. \qquad (2.3)$$

It satisfies the pricing condition $\mathbb{E}_t \left[SDF_{t+1}(1 + R_{t+1}) \right] = 1$, or equivalently

$$k\mathbb{E} \left[\left(\frac{C_{t+1}}{C_t} \right)^{-\alpha} \frac{D_{t+1}}{D_t} \frac{1 + Q(M_{t+1})}{Q(M_t)} \middle| I_t \right] = 1,$$

where $k = \delta \{\mathbb{E}[(C_{t+1}/C_t)^{1-\alpha}]\}^{\frac{1}{\theta} - 1}$. As shown in the Appendix, the price:dividend ratio therefore solves the fixed-point equation

$$Q(M_t) = \mathbb{E}_t \left\{ e^{\mu_d(M_{t+1}) - r_f - \alpha\rho_{c,d}\sigma_c\sigma_d(M_{t+1})} [1 + Q(M_{t+1})] \right\}, \qquad (2.4)$$

where $\rho_{c,d} = Cov(\varepsilon_{c,t}, \varepsilon_{d,t}) > 0$ is the correlation between the Gaussian noises in consumption and dividends. When the process $\{\sigma_d(M_t)\}$ is persistent, a large standard deviation of dividend growth at a given date t implies a low contemporaneous price:dividend ratio $Q(M_t) = \mathbb{E}_t \sum_{n=1}^{+\infty} \left(\Pi_{h=1}^n e^{\mu_d(M_{t+h}) - r_f - \alpha\rho_{c,d}\sigma_c\sigma_d(M_{t+h})} \right)$. We thus obtain the desired volatility feedback for any choices of the relative risk aversion α and the EIS ψ.

When the state space is finite, the equilibrium price:dividend ratio can be easily computed numerically. Consider the row vector $\iota = (1, ..., 1) \in \mathbb{R}^d$, the equilibrium column vector

$$q = [Q(m^1), ..., Q(m^d)]',$$

and the matrix $B = (b_{ij})_{1 \leq i,j \leq d}$ with components

$$b_{ij} = e^{\mu_d(m^j) - r_f - \alpha\rho_{c,d}\sigma_c\sigma_d(m^j)} \mathbb{P}(M_{t+1} = m^j | M_t = m^i).$$

---

[10] The assumption that investors are more informed than the econometrician is a reasonable assumption, as for instance discussed in Cochrane (2001).

The pricing condition (2.4) can be rewritten as $q = B(\iota + q)$, or equivalently

$$q = (I - B)^{-1}B\iota'. \tag{2.5}$$

This expression fully characterizes the equilibrium prices corresponding to a given set of parameters.

We consider the log-return $r_{t+1} \equiv \ln(1 + R_{t+1})$. It is easy to show that the excess return between date t and date $t + 1$ satisfies

$$r_{t+1} - r_f = \ln \frac{1 + Q(M_{t+1})}{Q(M_t)} + \mu_d(M_{t+1}) - r_f - \frac{\sigma_d^2(M_{t+1})}{2} + \sigma_d(M_{t+1})\varepsilon_{d,t+1}. \tag{2.6}$$

The excess return is thus determined by the price:dividend ratio and the realization of the dividend growth. Movements in the price:dividend ratio are manifestations of volatility feedback, and are partly predictable. If the multipliers M_t is high, we expect that M_{t+1} will be smaller and thus expect a *high* return.

2.3. Economies with Incomplete Information and Learning

The results easily extend to incomplete information structures. We now assume that the investor observes in each period a signal $\delta_t \in \mathbb{R}^N$. The information set $I_t = \{\delta_{t'}; t' \leq t\}$ generates a conditional belief Π_t over the state space $\{m^1, ..., m^d\}$. The price:dividend ratio is now a function of the investor probability: $P_t = Q(\Pi_t)D_t$. The gross return on the stock satisfies the pricing condition $\mathbb{E}[SDF_{t+1}(1 + R_{t+1})|I_t] = 1$, or equivalently

$$k\mathbb{E}\left[\left(\frac{C_{t+1}}{C_t}\right)^{-\alpha} \frac{D_{t+1}}{D_t} \frac{1 + Q(\Pi_{t+1})}{Q(\Pi_t)} \middle| I_t\right] = 1.$$

The price:dividend ratio

$$Q(\Pi_t) = \mathbb{E}\left[\sum_{i=1}^{\infty} k^i \left(\frac{C_{t+i}}{C_t}\right)^{-\alpha} \frac{D_{t+i}}{D_t} \middle| I_t\right]$$

is the conditional expectation of exogenous variables. It is therefore linear in the current belief Π_t :

$$Q(\Pi_t) = \sum_{j=1}^{d} Q(m^j)\Pi_t^j,$$

where $Q(m^j)$ is the price:dividend ratio computed under full information. This property considerably simplifies the learning problem, and is especially important for the analysis of economies with high dimensional state spaces.

The excess return is determined by the volatility state and investor belief:

$$r_{t+1} - r_f = \ln \frac{1 + Q(\Pi_{t+1})}{Q(\Pi_t)} + \mu_d(M_{t+1}) - r_f - \frac{\sigma_d^2(M_{t+1})}{2} + \sigma_d(M_{t+1})\varepsilon_{d,t+1}. \tag{2.7}$$

When a new state occurs, it takes time for investors to learn. The market thus adjusts slowly to shocks and generates less extreme returns than in the full information economy.

2.4. Inference and Estimation

The econometrician observes excess returns but not the volatility state. Let $I_t^0 \equiv \{r_s - r_f\}_{s=1}^t$ denote the set of excess returns up to date t, and $\hat{\Pi}_t^j = \mathbb{P}\left(M_t = m^j \left| I_t^0\right.\right)$, $j \in \{1, ..., d\}$, the implied conditional probabilities over the state space.

We assume until Section 5 that the investor observes the true volatility state M_t. For any return r, consider the matrix $F(r)$ with elements $f^{ij}(r) \equiv f_{r_{t+1}}\left(r|M_t = m^i, M_{t+1} = m^j\right)$. The econometrician's conditional probabilities are computed recursively using Bayes' rule:

$$\hat{\Pi}_{t+1} = \frac{\hat{\Pi}_t \left[A * F(r_{t+1})\right]}{\hat{\Pi}_t \left[A * F(r_{t+1})\right] \iota'} \tag{2.8}$$

As shown in the Appendix, the log-likelihood of the return process is then:

$$\ln L(r_1, ..., r_T) = \sum_{t=1}^{T} \ln \left\{\hat{\Pi}_{t-1} \left[A * F(r_t)\right] \iota'\right\}. \tag{2.9}$$

The model thus generates a return process with stochastic volatility and closed-form likelihood. This permits asymptotically efficient estimation in empirical applications.

3. Volatility Feedback with High-Dimensional Regime-Switching

In this section, we develop a parsimonious high-dimensional version of our setup, which can be used to investigate volatility feedback under multifrequency shocks.

3.1. A Multifrequency Specification for Dividend News

While the standard literature (e.g. Campbell and Hentschel, 1992) assumes that volatility shocks decline at a single frequency, economic intuition and a large body of research suggests that corporate profits and dividends are hit by shocks that have heterogeneous degree of persistence. For instance, earnings may be affected by short-run effects such as weather shocks. In the medium run, the business cycle creates uncertainty that is reflected in the volatility of dividends. In the long run, trends in demography, globalization or technology probably have quite persistent effects on dividends. The heterogeneous duration of volatility shocks has for instance been recognized in the option literature (Heston, 1993). Earlier theoretical research (e.g. Veronesi, 2000), suggests that the impact on the volatility of stock returns is likely to be substantial.

The specification of multifrequency volatility shocks might seem cumbersome and lead to a model with a large number of parameters. Fortunately, a solution to these

issues is provided by a recent advance in time series econometrics, the Markov-Switching Multifractal (MSM) process of Calvet and Fisher (2001, 2002, 2004). We assume a constant drift

$$\mu_d(M_t) \equiv g_d,$$

and consider $\overline{k}$ volatility components $M_{1,t}, M_{2,t}, ..., M_{\overline{k},t}$, decaying at heterogeneous frequencies $\gamma_1, .., \gamma_{\overline{k}}$. We specify stochastic volatility as

$$\sigma_d(M_t) \equiv \bar{\sigma}_d \left(M_{1,t} M_{2,t} ... M_{\overline{k},t} \right)^{1/2},$$

where each random *multiplier* $M_{k,t}$ satisfies $\mathbb{E}(M_{k,t}) = 1$, and $\bar{\sigma}_d$ is a positive constant. We conveniently stack the multipliers into a vector

$$M_t = (M_{1,t}, M_{2,t}, ..., M_{\overline{k},t}).$$

Given $m = (m_1, .., m_{\overline{k}}) \in \mathbb{R}^{\overline{k}}$, denote by $g(m)$ the product $\prod_{i=1}^{\overline{k}} m_i$. We can now write the time t volatility as $\bar{\sigma}_d[g(M_t)]^{1/2}$. The process has a finite number $\overline{k}$ of latent volatility state variables, each of which corresponds to a different frequency.

Consistent with the previous section, we assume that M_t follows a first-order Markov process. This design facilitates the iterative construction of the process through time, and permits maximum likelihood estimation of the parameters.[11] We call M_t the *volatility state vector*.

Each component $M_{k,t}$ follows a first-order Markov process that is identical except for time scale. Assume that the volatility state vector has been constructed up to date $t-1$. For each $k \in \{1, .., \overline{k}\}$, the next period multiplier $M_{k,t}$ is drawn from a fixed distribution M with probability γ_k, and is otherwise equal to its current value: $M_{k,t} = M_{k,t-1}$. The dynamics of $M_{k,t}$ can be summarized as:

$$\begin{array}{ll} M_{k,t} \text{ drawn from distribution } M & \text{with probability } \gamma_k \\ M_{k,t} = M_{k,t-1} & \text{with probability } 1 - \gamma_k. \end{array}$$

The switching events and new draws from M are assumed to be independent across k and t. The volatility components $M_{k,t}$ thus differ in their transition probabilities γ_k but not in their marginal distribution M. These features greatly contribute to the parsimony of the model.

Following our earlier work, we close the model by using a tight parameterization of the transition probabilities γ_k. The probabilities $\gamma \equiv \left(\gamma_1, \gamma_2, ..., \gamma_{\overline{k}} \right)$ satisfy

$$\gamma_k = 1 - (1 - \gamma_1)^{\left(b^{k-1} \right)}. \tag{3.1}$$

Since $1 - \gamma_k = (1 - \gamma_1)^{\left(b^{k-1} \right)}$, we observe that the probability of a multiplier not changing is decreasing exponentially in base b powers as k increases. We complete the specification of dynamics by assigning $\gamma_k = \gamma^*$.

---

[11] This innovation, introduced in Calvet and Fisher (2001), distinguishes our construction from previous multifractal processes that are generated by recursive operations on the entire sample path.

3.2. Loglinearized Return Dynamics under Full Information

To develop intuition, we consider a full information economy and derive a loglinear approximation to the pricing equation. Specifically, assume that the logarithm of the price-dividend ratio $q(M_t) = \ln Q(M_t)$ satisfies

$$q(M_t) \approx \bar{q} - \sum_{k=1}^{\bar{k}} q_k(M_{k,t} - 1). \tag{3.2}$$

Following Campbell-Shiller, let $\rho = e^{\bar{q}}/(1 + e^{\bar{q}})$ denote the average ratio of the stock price to the sum of the stock price and the dividend. We know that ρ is empirically close to 1. As shown in the Appendix, fixed-point condition (2.4) implies that

$$q_k = \alpha \sigma_{c,d} q_k^*, \tag{3.3}$$

for each $k \in \{1, .., \bar{k}\}$, where $\sigma_{c,d} = \sigma_c \bar{\sigma}_d \rho_{c,d}$ and each coefficient q_k^* satisfies

$$q_k^* = \frac{(1 - \gamma_k)/2}{1 - (1 - \gamma_k)\rho}.$$

The approximate solution holds for all choices of $\gamma_1, ..., \gamma_{\bar{k}}$, and thus does not require that scaling rule (3.1) be imposed. High frequency components have a negligible effect on the price:dividend ratio: $q_k^* \to 0$ when $\gamma_k \to 1$. On the other hand for very persistent components ($\gamma_k \to 0$), the coefficient q_k^* is approximately $1/[2(1-\rho)]$, which is typically large since ρ is close to unity.

The unconditional expected return $\mu = \mathbb{E}r_t$ satisfies $\mu - r_f = \alpha \sigma_{c,d}$, as is familiar in Consumption CAPM. Volatility innovations thus have no impact the unconditional equity premium or the interest rate. Realized returns are of course affected and satisfy

$$r_{t+1} - r_f \approx (\mu - r_f) \left\{ 1 + \sum_{k=1}^{\bar{k}} q_k^*[(M_{k,t} - 1) - \rho(M_{k,t+1} - 1)] \right\} + \sigma_d(M_{t+1})\varepsilon_{d,t+1}.$$

The regimes generate large clustered outliers, as in our earlier work. We now show that they also have two important effects. First, the regimes introduce predictability and mean reversion in returns, as in Cecchetti, Lam, and Mark (1990). Second, unexpected volatility increases are accompanied by negative returns (volatility feedback).[12]

Given the agent's information, the *predictable* component of the return between t and $t + 1$ is

$$\mathbb{E}_t r_{t+1} - r_f \approx (\mu - r_f) \left[1 + \sum_{k=1}^{\bar{k}} (1 - \gamma_k)(M_{k,t} - 1)/2 \right]. \tag{3.4}$$

---

[12] For instance the return innovation defined in (3.5) satisfies $Cov_t(r_{t+1}; M_{k,t+1}) = q_k \rho Var(M) < 0$.

The conditional return is thus the persistence-weighted sum of volatility components. Multipliers with higher durations command a higher expected return than more transitory components. We note that the formula contrasts with the relationships obtained in traditional volatility models, where the conditional return is typically a function of volatility itself (e.g. Merton, 1980; CH, 1992). Another feature of our model is that returns exhibit mean reversion: $\mathbb{E}_t r_{t+n} - r_f \approx (\mu - r_f)[1 + \sum_{k=1}^{\bar{k}}(1 - \gamma_k)^n (M_{k,t} - 1)/2] \rightarrow \mu - r_f$ as $n \rightarrow \infty$. Note, however, that the convergence of $\mathbb{E}_t r_{t+n}$ to the mean may be non-monotonic. For instance if $M_{1,t} > 1$ and $M_{\bar{k},t} < 1$, volatility is expected to increase in the short run and decrease in the long run, implying similar movements in conditional returns.

The *unpredictable* return $u_{t+1} = r_{t+1} - \mathbb{E}_t r_{t+1}$ satisfies

$$u_{t+1} \approx -\rho \sum_{k=1}^{\bar{k}} q_k (M_{k,t+1} - \mathbb{E}_t M_{k,t+1}) + \sigma_d(M_{t+1})\varepsilon_{d,t+1}. \tag{3.5}$$

An unexpected increase in a volatility component reduces the price: dividend ratio and the return on the stock. Similarly, the return innovation is positive when the volatility component is smaller than expected. As previously, the effect of an innovation on a multiplier depends on its frequency. This mechanism suggest that volatility and returns are negatively correlated, and generates skewness in the distribution of returns.

The model permits us to revisit the "no news is good news" effect discussed in CH. Consider component k and assume that no news has arrived between date t and date $t+1$: $M_{k,t+1} = M_{k,t}$. If the component is initially low ($M_{k,t} < 1$), volatility remains at a low level and no news is then good news for the stock market: $-\rho q_k(M_{k,t+1} - \mathbb{E}_t M_{k,t+1}) = \rho q_k \gamma_k(1 - M_{k,t+1}) > 0$. On the other hand if volatility is initially high ($M_{k,t} > 1$), no arrival is bad news for stock returns: $-\rho q_k(M_{k,t+1} - \mathbb{E}_t M_{k,t+1}) < 0$. In contrast to CH, the model implies that the absence of an arrival can be either bad news or good news for the stock market depending on the volatility state.

Investor anticipation tends to make returns more volatile than dividend news volatility. The stock market amplification of exogenous shocks is quantified by the *unconditional volatility feedback*

$$\frac{Var(r_{t+1} - r_f)}{\bar{\sigma}_d^2} - 1 \approx (\alpha \sigma_c \rho_{c,d})^2 Var(M) \sum_{k=1}^{\bar{k}} q_k^{*2}[2\rho\gamma_k + (\rho - 1)^2].$$

Note that this quantity increases with the duration and size of the volatility components. Volatility feedback may thus help explain the findings of Campbell and Shiller (1988) and Campbell (1991) that returns are considerably more variable than revisions of dividend forecasts.

The conditional variance of returns is

$$Var_t(r_{t+1}) \approx \mathbb{E}_t \sigma_{t+1}^2 + \rho^2 \sum_{k=1}^{\bar{k}} q_k^2 Var_t(M_{k,t+1}),$$

where $\sigma_{t+1} = \sigma_d(M_{t+1})$ and $Var_t(M_{k,t+1}) = \gamma_k[Var(M) + (1 - \gamma_k)(M_{k,t+1} - 1)^2]$. This implies that the conditional expected return $\mathbb{E}_t r_{t+1}$ and the conditional variance $Var_t(r_{t+1})$ are positively correlated.

CH attribute their weak estimates of volatility feedback to the fact in GARCH-type processes, the volatility of volatility increases very rapidly (as a fourth power) of the volatility level. This precludes the estimation of large effects, and makes it difficult for the CH model to capture the dynamics around the crashes of 1929 and 1987. Our volatility specification does not exhibit this undesirable property. For instance when $\bar{k} = 1$, we know that $Var_t(\sigma_{t+1}^2) = \bar{\sigma}_d^4 Var_t(M_{1,t})$ and thus

$$Var_t(\sigma_{t+1}^2) = \bar{\sigma}_d^4 \gamma_k[Var(M) + (1 - \gamma_k)(M_{1,t} - 1)^2]$$

The volatility of volatility is thus a non-monotonic, U-shaped function of the volatility level. Since $M_{1,t} \in (0, 2)$ in application, the volatility of volatility is symmetric around $M_{1,t} = 1$. Since the volatility state is mean-reverting, volatility is more subject to abrupt adjustments when it is further away from the mean. In the presence of several frequencies, $Var_t(\sigma_{t+1}^2)$ is then a sum of U-shaped functions of the multipliers, but cannot be expressed as a function of σ_t alone. These properties suggest that our model does not suffer from the same shortcomings as GARCH, and may yield larger estimates of the volatility feedback.

4. Empirical Results with Symmetric Dividends and Full Information

This section begins our empirical investigation of volatility feedback in U.S. equity markets. We specialize to the case of a symmetric dividend process and full information. This specification contributes significantly to explaining extreme returns and excess kurtosis in stock market data.

4.1. Excess Return Data

We estimate our model on daily excess returns of the value-weighted CRSP index. As a proxy for the risk-free rate, we impute daily returns on 30-day T-bills from the monthly return on the same instrument as reported by CRSP. Our data spans from July 1962 to December 2002, for a total of 10,194 observations. The data are plotted in Figure 1, and show the thick tails, low-frequency volatility cycles, and negative skewness that characterize aggregate stock market returns.

Table 1 reports moments of the excess return series, for the entire sample and four evenly spaced subsamples of the data. We observe that all moments show substantial variability across subsamples. This contradicts standard formulations based on innovations from simple GARCH or stochastic volatility models. These models have short memory, and sample moments tend to converge quickly to population moments. Substantial variability across subsamples is, however, consistent with the multifrequency regime-switching investigated in this paper.

4.2. Maximum Likelihood Estimates

The full-information model with symmetric dividends is specified by the number $\bar{k}$ of frequencies and the six parameters $m_0, \bar{\sigma}_d, b, \gamma_{\bar{k}}, g_d - r_f$, and $\alpha\sigma_{c,d} = \alpha\sigma_c\bar{\sigma}_d\rho_{c,d}$. As is standard in the literature, (e.g. Campbell and Shiller, 1988), we restrict one parameter by relating the price:dividend ratio in our estimated model to the price:dividend ratio in the data. Consider the value ρ defined by

$$\ln\rho \quad = \quad \mathbb{E}\ln\frac{Q(M_t)}{1 + Q(M_t)} \tag{4.1}$$

$$= \quad \mathbb{E}\ln\frac{P_t}{P_t + D_t}. \tag{4.2}$$

Given values for the other five parameters, a unique value of $\alpha\sigma_{c,d}$ ensures that the average price-dividend ratio $Q(M_t)$ matches the empirical value of ρ defined by (4.2).[13]

We thus estimate the restricted model with parameters

$$\psi \equiv (m_0, \bar{\sigma}_d, b, \gamma_{\bar{k}}, g_d - r_f) \in \mathbb{R}_+^5.$$

Maximization of the likelihood function (2.9) gives the parameter estimates and standard errors reported in Table 2. The columns of the table correspond to $\bar{k}$ varying from 1 to 10. The first column corresponds to a standard regime-switching model with only two possible states for volatility, as has been investigated previously in many settings, including volatility feedback (Kim, Morley, and Nelson, 2002). As $\bar{k}$ increases the number of states increases at the rate $2^{\bar{k}}$. For the largest value $\bar{k} = 8$, there are over 250 volatility states.

First examining the likelihood function, we observe a dramatic improvement over the standard two-state Markov specification as $\bar{k}$ increases. When $\bar{k}$ is low $(\bar{k} = 1, 2, 3)$, the incremental increase in likelihood is over 100 points and thus very substantial. For moderate values $(\bar{k} = 4, 5)$, adding components still increases the fit significantly.

---

[13]The expression $Q(M_t) = \mathbb{E}_t\sum_{n=1}^{+\infty}e^{n(g_d-r_f)-\alpha\sigma_{c,d}[\sqrt{g(M_{t+1})}+...+\sqrt{g(M_{t+n})}]}$ implies that the price:dividend ratio $Q(M_t)$ monotonically decreases from $+\infty$ to 0 as $\alpha\sigma_{c,d}$ increases from $-\infty$ to $+\infty$. Thus for every $\mu - r_f$ and $\rho < 1$, equation (4.1) has a unique solution. The loglinearized solution suggests that $\alpha\sigma_{c,d}$ is of the same order as $\mu - r_f$.

Finally, at $\bar{k} = 6$ the likelihood increase becomes more marginal, and between $\bar{k} = 7$ and $\bar{k} = 8$ the function exhibits a slight decline. Thus, as with exchange rates (Calvet and Fisher, 2004), a substantial number of heterogeneous volatility components are useful in fitting the dynamics of equity returns.

Table 2 also shows that the multiplier value $\hat{m}_0$ decreases as $\bar{k}$ increases, and is estimated with good precision. This is a sensible result, since a larger number of volatility components allows each individual component to do less work in explaining aggregate volatility variations. The estimates of $\bar{\sigma}_d$ show no apparent pattern across $\bar{k}$ and some degree of variability. This can be viewed as a strength of our model, as it is consistent with the idea that the long-run average of volatility is difficult to identify even in very large samples. This type of result is not possible with standard short-memory GARCH and stochastic volatility models, but it is in keeping with the results from Table 1 that sample second moments can vary considerably across subsamples. The dividend growth rate $g_d - r$ is positive and estimated with good precision. This parameter helps to control mean returns. Finally, the frequency parameters $\gamma_{\bar{k}}$ and b are both estimated with reasonable precision and show interesting patterns across $\bar{k}$. First, the switching probability $\gamma_{\bar{k}}$ of the highest frequency volatility component is fairly stable across specifications, occurring approximately once every 15 to 30 days. The parameter b, which controls spacing between frequencies, drops initially with $\bar{k}$ but then stabilizes at a value of about 2 for specification with $\bar{k} = 6$ and larger. Thus, as frequencies are added to the model, they primarily extend the low frequency range of the volatility components.

Given a set of parameter estimates, we can calculate unconditional moments of the model. For expected returns, note that since M_t is stationary, $\ln \rho \equiv -\mathbb{E} \ln\{[1 + Q(M_{t+1})]/Q(M_t)\}$, and thus

$$\mathbb{E}(r_t - r_f) = g_d - r_f - \ln \rho - \bar{\sigma}_d^2/2.$$

The unconditional volatility feedback is given by

$$\frac{Var(r_t - r_f)}{\bar{\sigma}_d^2} - 1 = \frac{1}{\bar{\sigma}_d^2} Var \left[\ln \frac{1 + Q(M_{t+1})}{Q(M_t)} - \frac{\bar{\sigma}_d^2}{2} g(M_{t+1}) \right].$$

This statistic was first calculated by CH, and is of particular interest because large values could help to explain thick tails and high volatility in stock market returns.

Table 2 reports the first four unconditional moments of returns under each specification. The estimated equity premium is too large relative to the data, and the skewness is not negative enough. These are related findings. One can interpret the estimated mean of a conditionally symmetric stochastic volatility model in terms of weighted least squares. When the inferred volatility at time t is low, the weighting of the time t return should be increased because the signal to noise ratio is higher. When inferred volatility

is high (as will generally be the case when an outlier from the tail of a skew distribution is drawn) the weighting should be decreased. Thus, if skewness in returns is not adequately modeled, the mean estimate will be biased towards the mode. Now examining the second moment, it is in most specifications comparable to the data. Finally, the model captures well excess kurtosis in returns when the value of $\bar{k}$ is large. The main problem revealed by this analysis is thus that the model has difficulty capturing large negative skewness.

Finally, we examine the unconditional volatility feedback of each specification reported in the last row of Table 2. For the best performing models with $\bar{k} \geq 6$, the effect is between $10 - 20\%$ of total variance, or about 5 to 10 times larger than reported by CH.

4.3. Volatility Decomposition of CRSP Excess Returns

This section analyzes conditional beliefs about the volatility state of stock market dividend news. We use a decomposition of the state space into marginals at different frequencies, which allows us to make inferences about the contribution of each frequency to overall volatility movements. The previous section shows that the likelihood function of the full information model tends to increase with the number of components. We therefore now focus on the specification with $\bar{k} = 8$ components.

Recall that in the full information specification, investors have access to the information set $I_t \equiv \{M_s; s \leq t\}$ of exact states up to and including time t. As empiricists, we would like to infer as much as possible about the investor's information given the more limited data on stock returns available to us.

We have already defined one information set $I_t^0 \equiv \{r_s - r_f; s \leq t\}$ available to the econometrician. This is the set of excess returns up to and including date t, and using this information to make inferences about states produces the beliefs $\hat{\Pi}_t^j \equiv \mathbb{P}\left(M_t = m^j \,\middle|\, I_t^0\right)$, $j \in \{1, ..., d\}$, where the number of states is given by $d = 2^{\bar{k}}$. We call these the *predictive* or *ex ante* beliefs, because they do not allow the econometrician access to forward-looking data from $t + 1$ and beyond.

In certain situations, the empiricist may also want to make inferences about the dividend news state using the larger information set of all returns I_T^0. For instance, if we want the best estimate of the portion of returns on any date t that are attributable to feedback effects, it is generally useful to condition on all returns. We thus define the *smoothed* or *ex post* probabilities $\hat{\Psi}_t^j \equiv \mathbb{P}\left(M_t = m^j \,\middle|\, I_t^0\right)$. Kim (1994) develops a smoother for standard Hamilton-type regime-switching specifications in which the conditional density of returns depends only on the current state M_t. In our model the conditional mean also depends on the previous state M_{t-1}, due to feedback from volatility changes. We show how to calculate smoothed beliefs under this expanded

regime-switching environment in the Appendix.

When $\bar{k} = 8$, the belief vectors $\hat{\Pi}_t$ and $\hat{\Psi}_t$ have dimension $2^8 = 256$. An interesting decomposition is to examine the marginals $\hat{\Pi}_t^{M(1)}, .., \hat{\Pi}_t^{M(\bar{k})}$ and $\hat{\Psi}_t^{M(1)}, ..., \hat{\Psi}_t^{M(\bar{k})}$, where $\hat{\Pi}_t^{M(k)} \equiv \mathbb{P}\left(M_{k,t} = m_0 \,\middle|\, I_t^0\right)$ and $\hat{\Psi}_t^{M(k)} \equiv \mathbb{P}\left(M_{k,t} = m_0 \,\middle|\, I_T^0\right)$ for $k = 1, .., \bar{k}$ and $t = 1, .., T$. These quantities are calculated and presented in Figure 2. We first examine the *ex ante* beliefs depicted in the eight panels on the left hand side of the figure. Each panel corresponds to one of the eight volatility components, increasing in frequency from low to high from top to bottom of the figure. For the lowest frequency $k = 1$, the marginal beliefs at first drop from 0.5 to about 0.1 then recover to a value of about 0.4 over the first five years of the sample. The marginal probability is then relatively flat or drifting downward slightly until the 1987 crash, at which point there is a large jump in beliefs to a probability of almost 1.0. The probability stays at this elevated level for the remainder of the sample. The econometrician thus infers under the full information model that an increase in the lowest frequency volatility component is likely to have occurred at the time of the crash.

Examining the remaining panels for the *ex ante* beliefs, we see that the second lowest frequency $k = 2$ shows patterns fairly similar to the lowest frequency $k = 1$. For the rest of the components, cycles in marginal beliefs become increasingly shorter in duration as k increases, consistent with intuition. Also, the strength of beliefs show an interesting pattern across frequencies. For low values of k, the conditional distribution of the volatility state spends considerable time at the extreme values of zero and one. By contrast, at high frequencies beliefs move up and down rapidly, but rarely reach their boundaries.

We now inspect the smoothed beliefs depicted on the right-hand side of Figure 2. Conditioning on all returns causes beliefs to move more sharply, and the empiricist now makes a strong distinction between the lowest $(k = 1)$ and second lowest $(k = 2)$ frequency volatility components. The smoothed belief for the first component begins the sample having an almost zero probability of being in the high state, and moves discretely to a value of nearly one on the crash date. By contrast, the beliefs for the second component stay near one for the entire sample. The remaining panels show similar results. The smoothed beliefs generally follow the same patterns as the *ex ante* beliefs, but the smoothed beliefs tend to move less frequently, more sharply, and to spend more time near the extreme boundaries of zero and one.

We finally note that careful inspection of either set of beliefs shows that on the date of the crash, all components are believed with near certainty to have reached their highest state. Here, dividend news volatility takes the value $\sigma_{d,t} = \bar{\sigma}_d \left(m_0^8\right)^{1/2} = 2.48\%$ per day. In Section 4.5 we determine the inferred realization of dividend news on the crash date after removing the conditional mean and feedback components of returns.

19

4.4. Estimated Conditional Moments and Conditional Feedback

Our full information regime-switching model permits investigation of conditional moments and feedback using the predictive beliefs $\hat{\Pi}_t$. For example, we can easily calculate the conditional mean $E\left(r_{t+1} - r_f | I_t^0\right)$ as a linear function of $\hat{\Pi}_t$. The Appendix details these computations for the first four conditional moments and feedback. Our empirical analysis again focuses on the specification with $\bar{k} = 8$ volatility components.

In the first two panels of Figure 3, we plot the conditional mean and volatility of excess returns. These are positively correlated, both showing small peaks in the early 1970's with higher levels in 1987 and around 2000. The conditional moments move less following the crash than is typical with GARCH specifications. For example, the ratio of the maximum to minimum value is about 2 for the conditional mean, and about 5 for conditional volatility. The inferred moments following the crash are similar to levels achieved in the last five years of the sample. Thus, conditional mean and variance respond to the crash as an important but not entirely anomalous event.

We now turn to the conditional feedback decomposition in the last three panels of Figure 3. One way to measure volatility feedback is the difference $Var_t\left(r_{t+1}\right) - Var_t\left(d_{t+1}\right)$. We call this quantity the *absolute conditional feedback*, because it gives the raw (non-standardized) variance contribution of risk-aversion induced components of returns. CH instead introduce the now traditional measure

$$\frac{Var_t\left(r_{t+1}\right) - Var_t\left(d_{t+1}\right)}{Var_t\left(d_{t+1}\right)}.$$

We call this *proportional conditional feedback* because it normalizes by the conditional dividend variance. This is a natural statistic to use in the CH environment since, as they discuss at length, the absolute feedback in their model grows as a fourth power of volatility. Standardizing by conditional variance thus helps to produce a less wildly varying statistic. Even after this normalization, proportional feedback is an amplified version of conditional volatility in their model. CH discuss this as a potentially undesirable feature that prevents obtaining higher levels of volatility feedback.

Our regime-switching model delivers a positive but much weaker relation between absolute feedback and volatility. As surmised by CH this is accompanied by a correspondingly higher contribution to variance from feedback effects. The last three panels of Figure 3 consecutively show the conditional absolute feedback, dividend news volatility, and proportional feedback. As in CH, absolute feedback moves with conditional volatility, and we compute a correlation coefficient of 75.8%. The variations in absolute feedback nonetheless do not appear extreme. We calculate the ratio between the maximum and minimum of absolute feedback and find a value less than two. This suggests that absolute feedback in our model is a dampened rather than magnified version of conditional variance. To confirm this, we regress the log of absolute feedback on the

log of dividend volatility $\sigma_{d,t}$. We find a regression coefficient of 0.223 with a standard error of 0.001 and an R^2 of 73.6%.

Thus, in contrast with CH where absolute feedback grows as a fourth power of conditional volatility, our absolute feedback is approximately proportional to a fourth root of volatility. This explains the inverse relationship between conditional variance and proportional feedback shown in the last two panels of Figure 3. This relationship is implied when, as suggested by CH, we achieve a considerably reduced amplification of conditional volatility in absolute feedback.

4.5. CRSP Excess Return Decomposition

We develop a decomposition of excess returns into a conditional expectation, feedback innovation, and a residual that is equal in conditional expectation to dividend news. This provides a convenient *ex post* quantification of the impacts of volatility feedback in our sample.

We first consider the fully informed investor. It is convenient to omit the Itô term in returns to obtain the approximation

$$r_{t+1} - r_f \approx g_d - r_f + \ln \frac{1 + Q(M_{t+1})}{Q(M_t)} + \sigma_d(M_{t+1})\varepsilon_{d,t+1}. \tag{4.3}$$

We note that the first three terms of (4.3) are equal to $\mathbb{E}\left(r_{t+1}|M_t, M_{t+1}\right)$. At time $t+1$ or later, the investor can thus implement the decomposition

$$r_{t+1} - r_f \approx \mathbb{E}\left(r_{t+1} - r_f|M_t\right) + \left[\mathbb{E}\left(r_{t+1}|M_t, M_{t+1}\right) - \mathbb{E}\left(r_{t+1}|M_t\right)\right] + \sigma_d(M_{t+1})\varepsilon_{d,t+1}.$$

This separates realized returns into a portion that is anticipated at time t, an innovation due to feedback effects from volatility changes, and the dividend news arrival.

Even after the entire sample is observed, the empiricist has a smaller information set $I_T^0 \subset I_T$, and thus derives an analogous decomposition with less precision. For any information sets A and B and random variable X, let $(\mathbb{E}_A - \mathbb{E}_B)(X) \equiv \mathbb{E}(X|A) - \mathbb{E}(X|B)$. We then obtain

$$r_{t+1} - r_f \approx \mathbb{E}_{\hat{\Psi}(t)}(r_{t+1} - r_f) + \left(\mathbb{E}_{\hat{\Psi}(t+1)} - \mathbb{E}_{\hat{\Psi}(t)}\right)\ln[1 + Q(M_{t+1})] + \hat{e}_{d,t+1}, \tag{4.4}$$

where

$$\hat{e}_{d,t+1} \equiv \mathbb{E}[\sigma_d(M_{t+1})\varepsilon_{d,t+1}|I_T^0] \tag{4.5}$$

is the econometrician's *ex post* estimate of realized dividend news. By the law of iterated expectations, $\hat{e}_{d,t+1}$ has mean zero.

We implement this decomposition on CRSP excess returns in Figure 4, which has in its first panel the excess return series $\{r_t - r_f\}$. The remaining panels show consecutively the three terms of (4.4): the smoothed mean return, the smoothed volatility

feedback component, and the smoothed dividend news estimate. We first note that the smoothed mean return does not appear substantially different from the *ex ante* mean return shown in Figure 3. Smoothed feedback, however, differs greatly from the *ex ante* predictions. Under the predictive beliefs in Figure 3, the variance of absolute feedback moves positively with conditional volatility, but is fairly stable across time. Conditioning on the larger information set I_T^0, the empiricist makes very different inferences about the contribution of feedback at different points in time. To understand this result, we consider the predictability of volatility innovations. In our regime-switching formulation, changes in the individual volatility states are unpredictable. Thus, even given precise information about the volatility state M_t, the econometrician has limited ability forecast feedback. When an exogenous change in a low-frequency component does occur, it appears immediately in the $t + 1$ return because investors have full information. The return r_{t+1} is thus very informative to the empiricist about changes in low-frequency volatility, and smoothed beliefs therefore give considerably greater refinement in the estimation of feedback effects. In particular, our *ex post* analysis attributes over half of the 1987 crash to volatility feedback.

We calculate moment statistics of the econometrician's *ex post* estimates of dividend news realizations, and find a variance of 0.693, skewness coefficient -0.121, and kurtosis 8.39. These are not necessarily unbiased signals of the higher moments of the true dividend news process. For example, applying Jensen's inequality to (4.5) shows that the variance of $\hat{e}_{d,t+1}$ is a lower bound for the variance of dividend news. Nonetheless, it is interesting to note that relative to the actual return data, the residual variance is 13% smaller, skewness is 89% smaller, and leptokurtosis is 78% smaller. The feedback decomposition thus explains a large portion of the higher moments of returns.

4.6. Comparison with Campbell and Hentschel (1992)

This section compares our full-information regime-switching model with the classic CH volatility feedback model. The CH specification is obtained by assuming a QGARCH(1,2) process for dividend news, and then loglinearlizing the identity relating returns to capital gains and dividends. Under these assumptions, returns have dynamics:

$$
\begin{aligned}
h_{t+1} &= \mu + \gamma\sigma_t^2 + \kappa\eta_{d,t+1} - \lambda\left(\eta_{d,t+1}^2 - \sigma_t^2\right), \\
\sigma_t^2 &= \varpi + \alpha_1\left(\eta_{d,t} - b\right)^2 + \alpha_2\left(\eta_{d,t-1} - b\right)^2 + \beta\sigma_{t-1}^2 \\
\lambda &= \frac{\gamma\rho\left(\alpha_1 + \rho\alpha_2\right)}{1 - \rho\left(\alpha_1 + \rho\alpha_2 + \beta\right)} \\
\eta_{d,t-1} &\sim N\left(0, \sigma_t^2\right),
\end{aligned}
$$

where h_{t+1} is the time $t + 1$ stock return, $\kappa = 1 + 2\lambda b$, the value ρ is obtained by calibration to the empirical price:dividend ratio, and the seven parameters to be estimated

are $\mu, \gamma, \varpi, \alpha_1, \alpha_2, b, \beta$.

The model contains sensible economic intuition. Expected returns $\mu + \gamma\sigma_t^2$ increase in conditional volatility, and the parameter γ can be interpreted as risk aversion under particular assumptions. Part of the volatility feedback effect appears in the linear third term as $\kappa\eta_{d,t+1} - \eta_{d,t+1}$, and the other portion of feedback is due to the quadratic fourth term. When returns are either very high or very low, the investor knows that volatility in the following period will be higher, hence there is an immediate drop in current returns.

The conditional variance σ_t^2 follows the QGARCH(1, 2) process of Engle (1990) and Sentana (1995). As in a standard GARCH(1,2), volatility is determined by lagged volatility and two lags of squared returns. QGARCH is distinct because it permits $b \neq 0$, and through this channel we obtain predictive asymmetry. If $b > 0$, then a positive dividend innovation will have greater impact on $t+1$ volatility than a negative innovation of the same size.

We estimate this model on the excess returns to the value weighted CRSP index, and report in Table 3A the parameter estimates. Both α_1 and α_2 are statistically significant, and α_2 has negative sign. This is because index data, even though value weighted, contains non-trading effects from less liquid stocks. The sum $\alpha_1 + \alpha_2 + \beta$ is close to one, as is typical of GARCH models estimated on daily data. The parameter γ is low relative to its possible interpretation as risk-aversion. Finally, the negative estimate of ϖ would be problematic in a standard GARCH model, but under QGARCH is not anomalous. Overall, the parameter estimates are consistent with those reported by CH.

Turning to Panel B of Table 3, we compare the in-sample fit of the CH specification to our multifrequency specification with $\bar{k} = 8$ volatility components. Although the multifrequency specification has two fewer parameters, it has a likelihood value over one hundred points larger. We calculate the BIC criterion to adjust for the number of parameters in each specification, and calculate a test statistic using the methodology of Vuong (1989). We also calculate a HAC adjusted version of the Vuong test using the methodology described in Calvet and Fisher (2004). These results show that the difference in likelihood is highly significant. In fact, for all multifrequency specifications with three or more volatility components, the likelihood dominates QGARCH dividend news. Thus, the multifrequency specification provides a better description of the data.

We next examine the volatility feedback effect of the estimated CH model. Conditional feedback is quantified by $\kappa^2 + 2\lambda^2\sigma_t^2$. As in CH, the volatility feedback effect is in the range of 1-2%, and thus quite small relative to the multifrequency model.

Table 4 examines how well each of the models matches the first four moments of the data. This is a challenging test, since these are not simply statistical models with free parameters to directly adjust characteristics of returns. Instead, both rely on endogenous economic mechanisms to fit higher moments of the data. We simulate $1,000$ paths

of the same length as the data from each process and calculate the first four moments of each path. We compute the mean and standard deviation across paths of each moment. and the percentage of times that the simulated moment exceeds the corresponding empirical moment. From this analysis, we see that both models overestimate the mean and underestimate negative skewness. These effects are related, as pointed out by CH and French, Schwert, and Stambaugh (1987), since failing to account for skewness in a conditionally heteroskedastic sample will likely bias the estimate of the mean toward the mode. Both models produce accurate estimates of the second moment, whereas in the fourth moment the multifrequency model does well while the CH model does not.

We thus identify difficulty matching the first and third moments of the data as common problems in feedback models.[14] The CH model does somewhat better on this point, because it adopts the QGARCH formulation with predictive asymmetry. In Section 5, we explore an alternative, more powerful, and perhaps economically more appealing method of generating endogenous skewness through learning.

5. Learning and Endogenous Skewness

The previous section made the extreme assumption that investors are always perfectly informed. If fundamental research is costly, we should instead expect some degree of imperfect information. We now develop a framework that permits a range of assumptions about the learning environment when the investor is less than perfectly informed.

5.1. Investor Information and Bayesian Updating

The other extreme from perfect information is to assume that the investor observes dividends at the same frequency the econometrician observes returns. We will see that under certain assumptions, the econometrician can then fully back out the dividend information used by the investor. This suggests that in this environment, investors have no information advantage relative to the empiricist. This is economically unappealing in the sense that it corresponds to a world where fundamental research has unbounded cost. Nonetheless it is a useful special case to begin with. When the investor receives only the daily dividend information

$$d_{t+1} - d_t = g_d + \sigma_{t+1}\varepsilon_{d,t+1} - \sigma_{t+1}^2/2$$

we call the environment a *daily learning* economy.

An economically more appealing setting allows the modelling of a range of intermediate information between full information and daily learning. Since the amount of information will be specified exogenously, this falls short of a more elaborate model in

---

[14]Wu (2001) comments on the difficulty of matching the third moment as well.

which investors endogenously choose when to incur costs to refine their information. Nonetheless, permitting a range of intermediate information structures captures in a reduced form way the intuition that when learning is costly, investors will undertake a positive but finite level of fundamental research. To achieve this, we assume for simplicity that volatility σ_{t+1} is constant over a day and that the agent observes $N \geq 1$ intraday pieces of dividend news:

$$\delta_{n,t+1} = \frac{g_d}{N} + \frac{\sigma_{t+1}}{\sqrt{N}} z_{n,t+1} - \frac{\sigma_{t+1}^2}{2N}, \quad (n = 1, .., N).$$

The shocks $\{z_{n,t+1}\}$ are IID Gaussian $\mathcal{N}(0,1)$, and the daily dividend news is the sum of intraday innovations: $d_{t+1} - d_t = \sum_{i=1}^{N} \delta_{n,t+1}$. We call this setup the *intraday learning* economy, and note that it nests daily learning.

Belief updating is straightforward for the investor under intraday learning. The investor receives the signal vector $\delta_{t+1} \equiv (\delta_{1,t+1}; ...; \delta_{N,t+1})$ and has information set $I_t = \{\delta_{t'}; t' \leq t\}$. Let Π_t denote her conditional belief over the state space at the end of period t. She updates her belief at the end of day $t+1$ contingent on the signal δ_{t+1}. Bayes' rule implies

$$\Pi_{t+1}^j \propto f(\delta_{t+1}|M_{t+1} = m^j)\mathbb{P}(M_{t+1} = m^j|I_t).$$

Conditional on $M_{t+1} = m^j$, the signal $\delta_{t+1} = (\delta_{1,t+1}; ...; \delta_{N,t+1})$ is Gaussian:

$$f(\delta_{t+1}|M_{t+1} = m^j) = \left[\frac{2\pi\sigma_d^2(m^j)}{N}\right]^{-N/2} \prod_{n=1}^{N} \exp\left[-\frac{\left(\delta_{n,t+1} - \frac{g_d}{N} + \frac{\sigma_d^2(m^j)}{2N}\right)^2}{2\sigma_d^2(m^j)/N}\right].$$

(5.1)

These expressions define the investor updating rule $\Pi_{t+1} = \Pi^*(\Pi_t; \delta_{t+1})$. We can easily simulate the intraday learning return process, as discussed in the Appendix.

When the number of intraday signals N is very large, the conditional density (5.1) trivially simplifies to a function of $\sum_{n=1}^{N}\left[(\delta_{n,t+1})^2\right]$. The sum of squared innovations, or "realized volatility of dividend news," is then a sufficient statistic for all intraday observations. In the limit as $N \to \infty$, the volatility level $\sigma_d(M_{t+1})$ is directly observable but the true volatility state M_{t+1} remains latent. The investor is thus less knowledgeable with $N = \infty$ intraday samples of dividend news than in the full information case. The investor trivially receives the least information in the daily news economy where $N = 1$.

The state of a learning economy consists of the volatility vector M_{t+1} and the investor belief Π_{t+1}. The econometrician observes excess returns: $I_t^0 = \{r_s - r_f; \quad s \leq t\}$, and can only compute the conditional distribution of the state (Π_t, M_t) given I_t^0. The full inference problem is therefore computationally expensive with large state spaces. Several solutions to this problem are now proposed.

5.2. Estimation of Daily Learning Model and Empirical Results

A simple solution is available in daily dividend news economies ($N = 1$). The investor then observes the one-dimensional signal $d_{t+1} - d_t$, and the econometrician may be able to extract it from the observed return r_{t+1}. The investor and the econometrician then have the same information sets, which considerably simplifies econometric inference and estimation.

Given her belief Π_t and the signal $d_{t+1} - d_t$, the investor has new belief $\Pi_{t+1} = \Pi^*(\Pi_t; d_{t+1} - d_t)$, which determines the capital gain or loss $\Phi(\Pi_t; d_{t+1} - d_t) \equiv \ln\{1 + Q[\Pi^*(\Pi_t; d_{t+1} - d_t)]\} - \ln Q(\Pi_t)$. The return between date t and date $t+1$ can then be written as

$$r_{t+1} = d_{t+1} - d_t + \Phi(\Pi_t; d_{t+1} - d_t). \tag{5.2}$$

If Π_t is known to the econometrician (i.e., $\hat{\Pi}_t = \Pi_t$), equation (5.2) implicitly defines the signal $d_{t+1} - d_t$ as a function of r_{t+1} and Π_t. In practice, the econometrician can guess a value φ, and successively compute the corresponding dividend growth $r_{t+1} - \varphi$ and capital gain/loss $\Phi(\Pi_t; r_{t+1} - \varphi)$. The true value of φ thus solves the fixed point

$$\varphi = \Phi(\Pi_t; r_{t+1} - \varphi).$$

This method thus permits maximum likelihood estimation, as discussed in the Appendix.

Maximum likelihood estimates of the daily learning economy are reported in Table 5. The parameter estimates are relatively similar to the full information parameter estimates, with the notable exception that the mean dividend growth rate is substantially lower (and often negative). This matches our intuition. In the full information model we hypothesized that the model resulted in an excessive mean return because it failed to adequately capture skewness. Now with a model that we expect to generate more skewness, we can calculate the mean return as between ± 0.001 per day across models, which is excessively low. The actual skewness generated by these parameter estimates is not substantially negative, however. To understand this counterintuitive result, we calculate the skewness coefficient of full information and daily learning models across a variety of parameter estimates and compare skewness holding the parameters constant. In unreported results, we find that for moderate levels of the equity premium the degree of skewness under daily learning is in fact excessive. This likely relates to the fact that the likelihood values achieved under daily learning are substantially lower than with full information. We speculate that the intraday learning model will best be able to capture moments of the data.

5.3. Estimation of Intraday Learning Model and Empirical Results

When $N > 1$, the investor receives a multidimensional signal $\delta_t \in \mathbb{R}^N$ and maximum likelihood estimation would then require the econometrician to integrate over the dis-

tribution of the unobservables. We instead estimate the intraday learning model by Simulated Method of Moments (SMM). Consider the true parameter vector ψ_0 that generates the excess return data $R^e = \{r_t - r_f\}_{t=1}^T$. Following Ingram and Lee (1991) and Duffie and Singleton (1993), it is common to simulate a single path of considerable length for a given parameter vector. One then computes the sample moments of the simulated series and a HAC weighting matrix to adjust for possible autocorrelation in the time-series of moment conditions. We slightly depart from this methodology because the moment conditions may exhibit low-frequency variations, which can be difficult to correct with a HAC weighting matrix. We instead draw multiple independent paths of identical length to the data, as discussed in the Appendix.

We use seven moment conditions motivated by the results obtained with full information economies. We consider: (a) the first four moments of excess returns: $\mathbb{E}[(r_t - r_f)^p]$, $p \in \{1, .., 4\}$, (b) the average size of excess returns $\mathbb{E}|r_t - r_f|$, and (c) two autocovariances: $\mathbb{E}|(r_t - r_f)(r_{t+h} - r_f)|$, $h \in \{1, 2\}$, which quantify short-run volatility persistence.

The SMM estimation results are reported in Table 6. For a number of components $\bar{k} \in \{5, 6, 8\}$, the preferred specification includes nine intraday observations. The exception is $\bar{k} = 7$, whose preferred specification has fifteen observations. This finding validates the idea that intermediate information structures best match the moments of equity returns.

All models generate substantial levels of skewness and kurtosis. These values are larger and thus closer to the data than the statistics provided by the estimated CH specification. The criterion function suggests that the model with 5 frequencies provides the worst overall fit. The criterion has an asymptotic $\chi^2(1)$ distribution, and the value 5.004 reported for $\bar{k} = 5$ is rejected at the 5% level. For larger values of $\bar{k}$, the reported criteria are smaller and well within the bound of the 95% confidence region. The model with 8 frequencies produces the most accurate mean and the most substantial skewness coefficient, consistent with the intuition that adding frequencies increases realism. The first moment is 0.0158% compared to CRSP sample average return of 0.0163%. A related finding is that the model endogenously generates a substantial level of skewness, which is equal to -0.43. The average absolute value and square of returns are remarkably close to the data. The model also captures well the short-run autocovariances in the size of returns, and thus short-run autocorrelation in volatility. Finally, the level of kurtosis is substantial at 8.9.

When we compare our model across frequencies, we observe that improvements in the skewness coefficient are monotonic with $\bar{k}$. The level of kurtosis is relatively stable between 8.15 and 10.15 with no particular trend. This is related to the fact that as we add more frequencies, the coefficient m_0 and the frequencies are varying. We also observe in Table 1 that kurtosis exhibits low frequency variations, and is thus measured with considerable uncertainty.

The intraday learning model eliminates the main biases of the full information and daily learning versions of our model. As compared to CH, the intraday learning economy provides a much better fit of mean return, skewness and kurtosis. The first moment is now essentially unbiased, and skewness is almost twice as large as in CH. Moreover, endogenous skewness is generated by an economically appealing learning mechanism, as opposed to asserting predictive asymmetry exogenously in the dividend process, as with QGARCH. As anticipated, the intraday learning model achieves the best balance between endogenous skewness and kurtosis.

6. Conclusion

This paper develops an asset pricing framework for economies with high dimensional regime-switching. Previous literature views stochastic regime shifts as an almost exclusively low-frequency phenomenon. We show that a pure Markov-switching economy can fit daily data well, better than the classic Campbell and Hentschel (1992) volatility feedback specification with QGARCH dividend news. Pure regime-switching models provide a viable and often improved alternative to standard approaches based on autoregressive dynamics. Using our methods, the role of regime-switching in asset pricing can, and should, be expanded.

The key to using pure switching models with high frequency data is our multifrequency approach. It is natural that different economic fundamentals such as liquidity imbalances, earnings cycles, business cycles, technological change, and demographics should all have different durations. Standard GARCH and stochastic volatility models do not permit this conveniently. We easily estimate models with up to eight different frequencies and over 250 states, and find that the likelihood function increases almost monotonically in the number of frequencies.

Our framework also easily accommodates learning. We estimate extreme models in which the investor has full information, or alternatively information no better than the econometrician. We then generalize our framework to permit intermediate levels of learning by giving the investor access to intraday dividend news. The degree of learning strongly affects endogenous skewness and kurtosis, and intermediate learning models best fit the data. This notion is economically compelling. Fundamental research is costly but not impossible. Investors should thus have imperfect information, but an advantage relative to the empiricist who relies on CRSP. One interesting extension of our model would be to consider endogenous information acquisition.

One caveat, and another direction for future work, is that like Campbell and Hentschel (1992), we do not believe that dividend volatility is the only source of variation in the price:dividend ratio. Ours is, to the best of our knowledge, the first piece of asset pricing research with high-dimensional regime-switching. It is therefore useful to fully

explore the framework in a single direction before moving to more empirically accurate assumptions. We nonetheless find it encouraging that the model generates considerable realism through volatility feedback alone. The theoretical framework also accommodates changes in the mean of dividend news growth, which can be the basis of future empirical work. Other extensions are also envisioned and will be the object of future research.

7. Appendix A. Pricing and Inference

7.1. First-Order Conditions

As shown by Epstein and Zin (1989), a utility-maximizing agent with budget constraint $W_{t+1} = (W_t - C_t)(1 + R_{t+1})$ has stochastic discount factor

$$SDF_{t+1} = \left[\delta \left(\frac{C_{t+1}}{C_t} \right)^{-\frac{1}{\psi}} \right]^{\theta} \left[\frac{1}{1 + R_{t+1}} \right]^{1-\theta},$$

where R_{t+1} is the simple net return on the optimal portfolio.

In our setup, the representative agent can be viewed as holding a long-lived claim on the aggregate consumption stream $\{C_t\}_{t=0}^{\infty}$. The price of the tree is $P_{c,t} = p_c C_t$, and the return on the tree is $1 + R_{c,t+1} = (1 + 1/p_c) C_{t+1}/C_t$. The stochastic discount factor is thus

$$
\begin{aligned}
SDF_{t+1} &= \delta^{\theta} (1 + 1/p_c)^{\theta-1} \left(\frac{C_{t+1}}{C_t} \right)^{-\theta/\psi - 1 + \theta} \\
&= \delta^{\theta} (1 + 1/p_c)^{\theta-1} \left(\frac{C_{t+1}}{C_t} \right)^{-\alpha}.
\end{aligned}
$$

The condition $\mathbb{E}_t[SDF_{t+1}(1+R_{c,t+1})] = 1$ implies that $\delta^{\theta}(1+1/p_c)^{\theta} \, \mathbb{E}[(C_{t+1}/C_t)^{1-\alpha}] = 1$ or equivalently

$$1 + 1/p_c = \delta^{-1} \{ \mathbb{E}[(C_{t+1}/C_t)^{1-\alpha}] \}^{\frac{1}{\theta}}.$$

We conclude that equation (2.2) holds.

7.2. Bayesian Updating and Likelihood

We consider the Bayesian updating problem solved by the econometrician in the full information economy. The conditional probability $\hat{\Pi}_{t+1}^{j} = \mathbb{P} \left(M_{t+1} = m^j \, | I_t^0, r_{t+1} \right)$ satisfies

$$\hat{\Pi}_{t+1}^{j} \propto f_{r_{t+1}} \left(r_{t+1} | M_{t+1} = m^j, I_t^0 \right) \mathbb{P} \left(M_{t+1} = m^j \, | I_t^0 \right).$$

We recognize that $f_{r_{t+1}} \left(r_{t+1} | M_{t+1} = m^j, I_t^0 \right)$ can be rewritten as

$$
\begin{aligned}
& \sum_{i=1}^{d} f_{r_{t+1}} \left(r_{t+1} | M_{t+1} = m^j, M_t = m^i \right) \mathbb{P} \left(M_t = m^i \, | M_{t+1} = m^j, I_t^0 \right) \\
= & \sum_{i=1}^{d} f_{r_{t+1}} \left(r_{t+1} | M_{t+1} = m^j, M_t = m^i \right) \frac{\mathbb{P} \left(M_{t+1} = m^j \, | M_{t+1} = m^i \right) \mathbb{P} \left(M_t = m^i \, | I_t^0 \right)}{\mathbb{P} \left(M_{t+1} = m^j \, | I_t^0 \right)}.
\end{aligned}
$$

The updated probability can now be written as

$$\hat{\Pi}_{t+1}^{j} \propto \sum_{i=1}^{d} f_{r_{t+1}} \left(r_{t+1} | M_{t+1} = m^j, M_t = m^i \right) \hat{\Pi}_t^i a_{ij},$$

which implies (2.8).

We now compute the log-likelihood function $\ln L\left(r_1,...,r_T\right)=\sum_{t=1}^{T}\ln f\left(r_t\,|r_1,...,r_{t-1}\right)$. Bayes' rule implies

$$
\begin{aligned}
f(r_t\,|r_1,...,r_{t-1}) &= \sum_{i=1}^{d}\sum_{j=1}^{d}\mathbb{P}(M_{t-1}=m^i,M_t=m^j|r_1,...,r_{t-1})f(r_t|M_{t-1}=m^i,M_t=m^j)\\
&= \sum_{i=1}^{d}\sum_{j=1}^{d}\hat{\Pi}_{t-1}^{i}a_{ij}f^{ij}\left(r_t\right),
\end{aligned}
$$

and therefore $f(r_t\,|r_1,...,r_{t-1})=\hat{\Pi}_{t-1}[A*F\left(r_t\right)]\iota'$.

7.3. Bayesian Smoothing

We derive the econometrician's smoothed belief $\hat{\Psi}_t^i\equiv\mathbb{P}\left(M_t=m^i\,\big|I_T^0\right)$ in the full information economy. Since M_t is first-order Markov, we know that

$$
\begin{aligned}
\hat{\Psi}_t^i &= \sum_{j=1}^{d}\mathbb{P}\left(M_{t+1}=m^j\,\big|I_T^0\right)\mathbb{P}\left(M_t=m^i\,\big|M_{t+1}=m^j,I_T^0\right)\\
&= \sum_{j=1}^{d}\mathbb{P}\left(M_{t+1}=m^j\,\big|I_T^0\right)\mathbb{P}\left(M_t=m^i\,\big|M_{t+1}=m^j,I_{t+1}^0\right)\\
&= \sum_{j=1}^{d}\frac{\mathbb{P}\left(M_{t+1}=m^j\,\big|I_T^0\right)}{\mathbb{P}\left(M_{t+1}=m^j\,\big|I_{t+1}^0\right)}\mathbb{P}\left(M_t=m^i,M_{t+1}=m^j\,\big|I_{t+1}^0\right),
\end{aligned}
$$

and therefore

$$
\hat{\Psi}_t^i=\sum_{j=1}^{d}\frac{\hat{\Psi}_{t+1}^j}{\hat{\Pi}_{t+1}^j}\mathbb{P}\left(M_t=m^i,M_{t+1}=m^j\,\big|I_{t+1}^0\right). \tag{7.1}
$$

The probability $\mathbb{P}\left(M_t=m^i,M_{t+1}=m^j\,\big|I_{t+1}^0\right)=\mathbb{P}\left(M_t=m^i,M_{t+1}=m^j\,\big|I_t^0,r_{t+1}\right)$ can be rewritten as

$$
\frac{f_{r_{t+1}}\left(r_{t+1}|M_{t+1}=m^j,M_t=m^i\right)\mathbb{P}(M_t=m^i,M_{t+1}=m^j|I_t^0)}{f_{r_{t+1}}\left(r_{t+1}|I_t^0\right)}=\frac{f^{ij}\left(r_{t+1}\right)\hat{\Pi}_t^i a_{ij}}{f_{r_{t+1}}\left(r_{t+1}|I_t^0\right)}
$$

and thus

$$
\mathbb{P}\left(M_t=m^i,M_{t+1}=m^j\,\big|I_{t+1}^0\right)=\frac{\hat{\Pi}_t^i a_{ij}f^{ij}\left(r_{t+1}\right)}{\sum_{m=1}^{d}\sum_{n=1}^{d}\hat{\Pi}_t^m a_{mn}f^{mn}\left(r_{t+1}\right)} \tag{7.2}
$$

By (7.1) and (7.2), we conclude that the smoothed probability is

$$
\hat{\Psi}_t^i=\hat{\Pi}_t^i\sum_{j=1}^{d}a_{ij}\frac{\hat{\Psi}_{t+1}^j}{\hat{\Pi}_{t+1}^j}\left[\frac{f^{ij}\left(r_{t+1}\right)}{L_{t+1}}\right],
$$

where $L_{t+1} = f_{r_{t+1}} \left(r_{t+1} | I_t^0 \right)$.

This smoothing rule slightly differs from the one derived by Kim (1994) for traditional Hamilton models, in which the signal observed by the econometrician (e.g. the excess return $r_{t+1} - r_f$) depends on the current state M_{t+1} but not on the past state M_t. To illustrate this point, note that our model implies

$$
\begin{aligned}
\hat{\Pi}_{t+1}^j &= \mathbb{P} \left(M_{t+1} = m^j \left| I_t^0, r_{t+1} \right. \right) \\
&= \mathbb{P} \left(M_{t+1} = m^j \left| I_t^0 \right. \right) f_{r_{t+1}} \left(r_{t+1} | I_t^0, M_{t+1} = m^j \right) / L_{t+1}.
\end{aligned}
$$

The smoothed probability thus satisfies

$$
\hat{\Psi}_t^i = \hat{\Pi}_t^i \sum_{j=1}^d a_{ij} \frac{\hat{\Psi}_{t+1}^j}{\mathbb{P} \left(M_{t+1} = m^j \left| I_t^0 \right. \right)} \left[\frac{f^{ij} \left(r_{t+1} \right)}{f_{r_{t+1}} \left(r_{t+1} | I_t^0, M_{t+1} = m^j \right)} \right]
$$

If the past state M_t has no effect on the density of r_{t+1}, the term in square brackets equals one and the smoothed belief then reduces to the Hamilton-Kim formulation. The expressions are otherwise different.

7.4. Moments of Returns in Full Information Economy

Consider the $1 \times d$ row vectors $\hat{\Pi}_t = (\hat{\Pi}_t^1, .., \hat{\Pi}_t^d)$, $\iota = (1, .., 1)$, $v = [\sigma_d^2(m^1), ..., \sigma_d^2(m^d)]$, and the $d \times d$ matrix $C = (c_{ij})$ with elements $c_{ij} = g_d - r_f - \sigma_d^2(m^j)/2 + \ln \frac{1 + Q(m^j)}{Q(m^i)}$. With this notation, the excess return can be rewritten as

$$
r_{t+1} - r_f = c_{ij} + \sigma_d(m^j) \varepsilon_{d,t+1}
$$

The conditional first moment is thus $\mathbb{E}_t(r_{t+1} - r_f) = \sum_{i,j} \hat{\Pi}_t^i a_{ij} c_{ij}$, or in matrix notation

$$
\mathbb{E}_t(r_{t+1} - r_f) = \hat{\Pi}_t (A * C) \iota'.
$$

Consider the matrix $\tilde{C}_{t+1}$ with elements $\tilde{c}_{ij}(t+1) = c_{ij} - \mathbb{E}_t(r_{t+1} - r_f)$. The unexpected return is

$$
r_{t+1} - r_f - \mathbb{E}_t(r_{t+1} - r_f) = \sigma_d(m^j) \varepsilon_{d,t+1} + \tilde{c}_{ij}(t+1).
$$

The conditional variance is therefore $Var_t(r_{t+1} - r_f) = \mathbb{E}_t[\sigma_d^2(m^j)] + \mathbb{E}_t[\tilde{c}_{ij}^2(t+1)]$. We easily check that $\mathbb{E}_t[\sigma_d^2(M_{t+1})] = \hat{\Pi}_t A v'$ and $\mathbb{E}_t[\tilde{c}_{ij}^2(t+1)] = \sum_{i,j} \hat{\Pi}_t^i a_{ij} \tilde{c}_{ij}^2(t+1)$, which implies

$$
Var_t(r_{t+1} - r_f) = \hat{\Pi}_t A v' + \hat{\Pi}_t (A * \tilde{C}_{t+1} * \tilde{C}_{t+1}) \iota'.
$$

The two addends quantify: (1) the conditional variance of dividend news, and (2) the variability of the price dividend ratio. If we neglect the Itô term, the conditional variance of dividends is approximated as

$$
Var_t(d_{t+1} - d_t) \approx \mathbb{E}_t[\sigma_d^2(M_{t+1})] = \hat{\Pi}_t A v',
$$

and the conditional volatility feedback can be computed as

$$\frac{Var_t(r_{t+1} - r_f)}{Var_t(d_{t+1} - d_t)} \approx 1 + \frac{\hat{\Pi}_t(A * \tilde{C}_{t+1} * \tilde{C}_{t+1})\iota'}{\hat{\Pi}_t A v'}.$$

For any matrix Z, it is henceforth convenient to denote by $Z^{(p)}$ the pth Hadamard power $Z * ... * Z = (z_{ij}^p)$.

The conditional third centered moment $\mathbb{E}_t[(r_{t+1} - \mathbb{E}_t r_{t+1})^3]$ is $\mathbb{E}_t[3\tilde{c}_{ij}(t+1)\sigma_d^2(m^j) + \tilde{c}_{ij}^3(t+1)]$, or equivalently

$$\mathbb{E}_t[(r_{t+1} - \mathbb{E}_t r_{t+1})^3] = 3\hat{\Pi}_t(A * \tilde{C}_{t+1})v' + \hat{\Pi}_t(A * \tilde{C}_{t+1}^{(3)})\iota'.$$

The fourth conditional centered moment $\mathbb{E}_t[(r_{t+1} - \mathbb{E}_t r_{t+1})^4]$ is $\mathbb{E}_t[\sigma_d^4(m^j)\varepsilon_{d,t+1}^4 + 6\tilde{c}_{ij}^2(t+1)\sigma_d^2(m^j) + \tilde{c}_{ij}^4(t+1)]$, or equivalently $\sum_{i,j} \hat{\Pi}_t^i a_{ij}[3\sigma_d^4(m^j) + 6\tilde{c}_{ij}^2(t+1)\sigma_d^2(m^j) + \tilde{c}_{ij}^4(t+1)]$. In matrix notation,

$$\mathbb{E}_t[(r_{t+1} - \mathbb{E}_t r_{t+1})^4] = 3\hat{\Pi}_t A(v * v)' + 6\hat{\Pi}_t(A * \tilde{C}_{t+1}^{(2)})v' + \hat{\Pi}_t(A * \tilde{C}_{t+1}^{(4)})\iota'.$$

7.5. Ex-Post Decomposition

We condition the return equation (4.3) with respect to the economometrician's information set I_T^0:

$$r_{t+1} - r_f \approx \mathbb{E}[\mathbb{E}(r_{t+1} - r_f|I_t)|I_T^0] + \mathbb{E}\{\ln[1 + Q(M_{t+1})] - \mathbb{E}_t \ln[1 + Q(M_{t+1})]|I_T^0\} + \hat{e}_{d,t+1}.$$

The first term can be rewritten as

$$\mathbb{E}[\mathbb{E}(r_{t+1} - r_f|I_t)|I_T^0] = \sum_{j=1}^{d} \mathbb{P}(M_t = m^j|I_T^0) \, \mathbb{E}[\mathbb{E}(r_{t+1} - r_f|I_t)|M_t = m^j, I_T^0].$$

Since $\mathbb{E}(r_{t+1} - r_f|I_t)$ is a deterministic function of the state M_t, the conditional expectation simplifies to

$$\mathbb{E}[\mathbb{E}(r_{t+1} - r_f|I_t)|I_T^0] = \sum_{j=1}^{d} \hat{\Psi}_t^j \, \mathbb{E}(r_{t+1} - r_f|M_t = m^j),$$

or equivalently $\mathbb{E}[\mathbb{E}(r_{t+1} - r_f|I_t)|I_T^0] = \mathbb{E}_{\hat{\Psi}(t)}(r_{t+1} - r_f)$

We similarly infer that

$$\mathbb{E}\{\ln[1 + Q(M_{t+1})] - \mathbb{E}_t \ln[1 + Q(M_{t+1})]|I_T^0\}$$

$$= \mathbb{E}\{\ln[1 + Q(M_{t+1})]|I_T^0\} - \sum_{j=1}^{d} \hat{\Psi}_t^j \mathbb{E}\{\ln[1 + Q(M_{t+1})]|M_t = m^j\}$$

$$= \sum_{i=1}^{d}(\hat{\Psi}_{t+1}^i - \sum_{j=1}^{d} \hat{\Psi}_t^j a_{ij}) \ln[1 + Q(m^i)],$$

and thus $\mathbb{E}\{\ln[1 + Q(M_{t+1})] - \mathbb{E}_t \ln[1 + Q(M_{t+1})]|I_T^0\} = \left(\mathbb{E}_{\hat{\Psi}(t+1)} - \mathbb{E}_{\hat{\Psi}(t)}\right) \ln[1 + Q(M_{t+1})].$

33

7.6. Learning Economies

Simulating the Return Process. The return process is simulated recursively by keeping track of the volatility state M_t and the investor conditional probability Π_t at the end of every period t. Assuming that these quantities are know, we draw N IID standard normals $\{z_{n,t}\}_{n=1}^N$ and sample M_{t+1} from M_t using the transition matrix A. We then compute $\{\delta_{n,t}\}_{n=1}^N$ $(n = 1, .., N)$ and the daily dividend growth $d_{t+1} - d_t = \sum_{n=1}^N \delta_{n,t}$. We finally infer the new investor belief $\Pi_{t+1} = F(\Pi_t; \delta_{1,t+1}; ...; \delta_{N,t+1})$ and the excess stock return (2.7). We are ready to iterate again.

SMM Estimation. For any candidate parameter vector ψ, we simulate J paths of length T, which are denoted by $Y_j(\psi) = \{Y_{j,t}(\psi)\}_{t=1}^T$, $j \in \{1, .., J\}$. For each path, a vector of sample moments $h[Y_j(\psi)]$ is calculated. We arrange the simulated paths in a $J \times T$ matrix $Y(\psi) = [Y_1(\psi), ..., Y_J(\psi)]'$, and define

$$H\left[Y(\psi), R^e\right] = h(R^e) - \frac{1}{J}\sum_{j=1}^J h[Y_j(\psi)].$$

The function H quantifies how well the model fits the empirical sample moments. In particular, we can define an objective function

$$G\left[Y(\psi), R^e, W\right] = H'WH$$

for any positive definite weighting matrix W. Maximizing the objective function G with respect to the parameter vector ψ provides a simulated method of moments estimator $\widehat{\psi}_{SMM}(W)$ for the process.

In practice, we start with the parameter estimates obtained under full information, compute the corresponding weighting matrix

$$W_{\widehat{\psi}} = \left\{\sum_{j=1}^J H\left[Y(\widehat{\psi}), Y_j(\widehat{\psi})\right] H\left[Y(\widehat{\psi}), Y_j(\widehat{\psi})\right]'/J\right\}^{-1},$$

and obtain a first-stage SMM estimate. This procedure is repeated once to produce an efficient SMM estimate. It is easy to check that as J and T go to infinity, $H'WH$ converges to a χ^2 distribution.

8. Appendix B - Loglinearized Return Dynamics

8.1. Approximate Solution

By (2.4), the price:dividend ratio satisfies

$$\ln Q(M_t) = g_d - r_f - \alpha\sigma_{c,d} + \ln \mathbb{E}_t \left\{ [1 + Q(M_{t+1})]e^{-\alpha\sigma_{c,d}[\sqrt{g(M_{t+1})}-1]} \right\}.$$

We assume that $\sigma_{c,d}$ is close to 0 and that the marginal distribution M is concentrated around 1, and look for a linear approximate solution to this fixed-point equation. The conditional expectation

$$\mathbb{E}_t \left\{ [1 + e^{\bar{q} - \sum_{k=1}^{\bar{k}} q_k(M_{k,t+1}-1)}]e^{-\alpha\sigma_{c,d}[\sqrt{g(M_{t+1})}-1]} \right\} \tag{8.1}$$

is approximately $\mathbb{E}_t \left\{ [1 + e^{\bar{q}} - e^{\bar{q}} \sum q_k(M_{k,t+1} - 1)] \left[1 - \frac{\alpha\sigma_{c,d}}{2} \sum(M_{k,t+1} - 1)\right] \right\}$, or

$$(1 + e^{\bar{q}})\mathbb{E}_t \left[1 - \sum \left(\frac{e^{\bar{q}}}{1 + e^{\bar{q}}}q_k + \frac{\alpha\sigma_{c,d}}{2}\right)(M_{k,t+1} - 1)\right].$$

Since $\rho = \frac{e^{\bar{q}}}{1+e^{\bar{q}}}$ and $\mathbb{E}_t(M_{k,t+1} - 1) = (1 - \gamma_k)(M_{k,t} - 1)$, we infer that (8.1) is approximately equal to $(1 + e^{\bar{q}}) \left[1 - \sum_{k=1}^{\bar{k}}(1 - \gamma_k)\left(\rho q_k + \frac{\alpha\sigma_{c,d}}{2}\right)(M_{k,t} - 1)\right]$. The linearized version of the Euler equation is thus

$$\bar{q} - \sum_{k=1}^{\bar{k}} q_k(M_{k,t} - 1) = g_d - r_f - \alpha\sigma_{c,d} + \ln(1 + e^{\bar{q}}) - \sum_{k=1}^{\bar{k}}(1 - \gamma_k)\left(\rho q_k + \frac{\alpha\sigma_{c,d}}{2}\right)(M_{k,t} - 1),$$

implying

$$\bar{q} = \ln(1 + e^{\bar{q}}) + g_d - r_f - \alpha\sigma_{c,d},$$
$$q_k = (1 - \gamma_k)\left(\rho q_k + \frac{\alpha\sigma_{c,d}}{2}\right).$$

The first equation can be rewritten as $\rho = e^{g_d - \mu}$. We infer from the second equation that q_k satisfies (3.3).

We next derive the log-linearized return on the stock. Linearize $\ln[1 + Q(M_{t+1})] \approx \ln[1 + e^{\bar{q} - \sum_{k=1}^{\bar{k}} q_k(M_{k,t+1}-1)}]$ around the unconditional mean $(1, 1, ..., 1)$:

$$\ln[1 + Q(M_{t+1})] \approx \ln(1 + e^{\bar{q}}) - \frac{e^{\bar{q}}}{1 + e^{\bar{q}}} \sum_{k=1}^{\bar{k}} q_k(M_{k,t+1} - 1).$$

Combining this result with (3.2), we infer

$$\ln \frac{1 + Q(M_{t+1})}{Q(M_t)} \approx -\ln\rho - \sum_{k=1}^{\bar{k}} q_k[\rho(M_{k,t+1} - 1) - (M_{k,t} - 1)].$$

Since $\mu = g_d - \ln \rho$, we conclude that

$$r_{t+1} \approx \mu + \bar{\sigma}_d \sqrt{g(M_{t+1})} \varepsilon_{d,t+1} - \sum_{k=1}^{\bar{k}} q_k [\rho(M_{k,t+1} - 1) - (M_{k,t} - 1)].$$

We note that $\mathbb{E}_t r_{t+1} = (\mu - r_f) \sum_{k=1}^{\bar{k}} (1 - \gamma_k)(M_{k,t} - 1)/2$.

8.2. Volatility Decomposition

Since $\varepsilon_{d,t+1}$ is independent from the multipliers, the unconditional variance of returns is

$$Var(r_{t+1}) \approx \bar{\sigma}_d^2 + \sum_{k=1}^{\bar{k}} q_k^2 Var[\rho(M_{k,t+1} - 1) - (M_{k,t} - 1)].$$

We note that $\mathbb{E}[(M_{k,t+1} - 1)(M_{k,t} - 1)] = (1 - \gamma_k)Var(M)$ and conclude that

$$Var(r_{t+1}) \approx \bar{\sigma}_d^2 + Var(M) \sum_{k=1}^{\bar{k}} q_k^2 [\rho^2 + 1 - 2\rho(1 - \gamma_k)].$$

8.3. Unconditional Skewness

There are two possible sources of skewness in the model: (1) the negative correlation between drift and volatility; and (2) the possible negative skewness of the multiplier. Let

$$\Lambda_{t+1} = \sum_{k=1}^{\bar{k}} q_k [(M_{k,t} - 1) - \rho(M_{k,t+1} - 1)].$$

The return r_{t+1} has third centered moment

$$\mathbb{E}[(r_{t+1} - \mu)^3] = 3Cov(\sigma_{t+1}^2, \Lambda_{t+1}) + \mathbb{E}(\Lambda_{t+1}^3).$$

We easily show that this implies

$$\mathbb{E}[(r_{t+1} - \mu)^3] = -3\sigma_d^2 Var(M) \sum_{k=1}^{\bar{k}} q_k(\gamma_k + \rho - 1) \tag{8.2}$$

$$+ \mathbb{E}[(M - 1)^3] \sum_{k=1}^{\bar{k}} q_k^3 [\gamma_k(1 - \rho^3) + (1 - \gamma_k)(1 - \rho)^3].$$

We begin by considering the component $3\sigma_d^2 Var(M) \sum_{k=1}^{\bar{k}} q_k(\gamma_k + \rho - 1)$. The addend $\rho + \gamma_k - 1$ is positive for reasonably high frequencies, but can be negative at very low frequencies. The overall effect is generally ambiguous. The second component on the

RHS of (8.2) quantifies the effect of the skewness of the multiplier. We observe that a negatively skewed multiplier leads to a more negatively skewed return process r_t.

To clarify the intuition underlying this result, we consider a symmetric binomial multiplier, which takes the values $m_0 > 1$ and $m_1 = 2 - m_0 < 1$ with equal probability. Assume for simplicity that the process has a unique component ($\bar{k} = 1$). Volatility is assumed to be high ($M_{1,t} = m_0$) in periods $t \in \{1, .., t^*\}$, low ($M_{1,t} = m_1$) in periods $t \in \{t^* + 1, t^{**}\}$, and high again in period $t^{**} + 1$. The return in the initial periods $t \leq t^*$ is

$$r_t = \mu + \Delta + \sigma_d \sqrt{m_0} \varepsilon_{d,t} \tag{8.3}$$

where $\Delta = (1 - \rho)(m_0 - 1)q_1 > 0$. Volatility falls between dates t^* and $t^* + 1$, and

$$r_{t^*+1} = \mu + \frac{1 + \rho}{1 - \rho} \Delta + \sigma_d \sqrt{m_1} \varepsilon_{d,t^*+1} \tag{8.4}$$

is typically much larger than previous returns. The multiplier is then constant until date t^{**}, implying

$$r_t = \mu - \Delta + \sigma_d \sqrt{m_1} \varepsilon_{d,t}. \tag{8.5}$$

Finally, we switch back to m_0 at date t^{**} and obtain the low return

$$r_{t^{**}+1} = \mu - \frac{1 + \rho}{1 - \rho} \Delta + \sigma_d \sqrt{m_0} \varepsilon_{d,t+1}. \tag{8.6}$$

There are thus two countervailing forces affecting the skewness of the return process. On one hand, transitions generate returns with identical means but different variances. As seen in (8.6), the transition from low to high volatility is associated with a low mean and a high variance, and can thus induce very negative excess returns. In contrast, the transition from high to low variance implies returns with high means but low variance, as seen in (8.4). In transition periods, very large positive outliers are thus less likely to be observed than very negative ones, which induces negative skewness in the unconditional return distribution. On the other hand, periods with constant volatility tend to induce positive skewness: by (8.5) and (8.3) returns have both a higher mean and a higher variance in high volatility periods than in low volatility ones. The switching probability γ_k. determines the dominant effect. At sufficiently high frequencies ($\gamma_k > 1 - \rho$), transition periods occur often and returns are negatively skewed, as seen in (8.2).

8.4. State Dependence in Sign of Conditional Skewness

The example also indicates that the conditional skewness of returns depends on the volatility state. Consider a component with a sufficiently low frequency. When volatility is high in a given period, the conditional mean is slightly lower than $\frac{1+\rho}{1-\rho}\Delta$. Volatility either stays high next period and generates a modest negative return; or switches to a

low level and generates a very positive return. Conditional skewness is then positive. Conversely in a market with low volatility, returns are either slightly positive (no switch) or very negative (switch).

We illustrate this intuition by computing the conditional skewness of the return process. By (3.5), the return innovation is $u_{t+1} = \tilde{\Lambda}_{t+1} + \sigma_{t+1}\varepsilon_{d,t+1}$, where $\tilde{\Lambda}_{t+1} = -\rho \sum_{k=1}^{\bar{k}} q_k(M_{k,t+1} - \mathbb{E}_t M_{k,t+1})$. The conditional third moment of the innovation is therefore

$$\mathbb{E}_t(u_{t+1}^3) = 3Cov_t(\sigma_{t+1}^2; \tilde{\Lambda}_{t+1}) + \mathbb{E}_t(\tilde{\Lambda}_{t+1}^3).$$

First, we know that $Cov_t(\sigma_{t+1}^2; \tilde{\Lambda}_{t+1}) = -\rho(\mathbb{E}_t\sigma_{t+1}^2)\sum_{k=1}^{\bar{k}} q_k \frac{Var_t(M_{k,t+1})}{\mathbb{E}_t M_{k,t+1}} < 0$. This is a direct consequence of volatility feedback. Second, we also note that $\mathbb{E}_t(\tilde{\Lambda}_{t+1}^3) = -\rho^3 \sum_{k=1}^{\bar{k}} q_k^3 \mathbb{E}_t[(M_{k,t+1} - \mathbb{E}_t M_{k,t+1})^3]$. Each addend can be expressed as a cubic function of the volatility component $M_{k,t}$, implying

$$\mathbb{E}_t(\tilde{\Lambda}_{t+1}^3) = \rho^3 \sum_{k=1}^{\bar{k}} q_k^3 \gamma_k \left\{ \begin{array}{c} 3(1-\gamma_k)Var(M)(M_{k,t}-1) \\ +(1-\gamma_k)(1-2\gamma_k)(M_{k,t}-1)^3 - \mathbb{E}[(M-1)^3] \end{array} \right\}.$$

If M has a symmetric distribution, we know that $\mathbb{E}_t(\tilde{\Lambda}_{t+1}^3)$ has a zero unconditional expectation: $\mathbb{E}[\mathbb{E}_t(\tilde{\Lambda}_{t+1}^3)] = 0$. The addends have the same sign as $3Var(M)(M_{k,t}-1) + (1-2\gamma_k)(M_{k,t}-1)^3$. When $\gamma_k < 1/2$, conditional skewness is thus positive in high volatility states ($M_{k,t} > 1$) and negative otherwise.

References

[1] Abel, Andrew (1988), Stock Prices Under Time-Varying Dividend Risk: An Exact Solution in an Infinite Horizon General Equilibrium Model, *Journal of Monetary Economics* **22**, 375-395.

[2] Abel, Andrew (1994), Exact Solutions for Expected Rates of Return Under Markov-Regime Switching: Implications for the Equity Premium, *Journal of Money, Credit and Banking* **26**, 345-361.

[3] Abel, Andrew (1999), Risk Premia and Term Premia in General Equilibrium, *Journal of Monetary Economics* **43**, 3-33.

[4] Andersen, Torben, and Tim Bollerslev (1998), Answering the Skeptics: Yes, Standard Volatility Models Do Provide Accurate Forecasts, *International Economic Review* **39**, 885-905.

[5] Attanasio, Orazio, and Guglielmo Weber (1993), Consumption Growth, The Interest Rate, and Aggregation, *Review of Economic Studies.* **60**, 631-649.

[6] Backus, David, and Gregory, A.(1993), Theoretical Relations Between Risk Premiums and Conditional Variances, *Journal of Business and Economic Statistics* **11**, 177-185.

[7] Bansal, Ravi, and Christian Lundblad (2002), Market Efficiency, Asset Returns, and the Size of the Risk Premium in Global Equity Markets, *Journal of Econometrics* **109**, 195-237.

[8] Bansal, Ravi, V. Khatchatrian, and Amir Yaron (2003), Interpretable Asset Markets?, Working Paper, Duke University.

[9] Bansal, Ravi, George Tauchen, and Hao Zhou (2003), Regime-Shifts, Risk Premiums in the Term Structure, and the Business Cycle, forthcoming in *Journal of Business and Economic Statistics.*

[10] Bansal, Ravi, and Hao Zhou (2002), Term Structure of Interest Rates with Regime Shifts, *Journal of Finance* **57**, 1997-2043.

[11] Barsky, Robert (1989), Why Don't the Prices of Stocks and Bonds Move Together?, *American Economic Review* **79**, 1132-1145.

[12] Bekaert, Geert, and Guojun Wu (2000), Asymmetric Volatility and Risk in Equity Markets, *Review of Financial Studies* 13:1, 1-42.

[13] Bollerslev, Tim (1986), Generalized Autoregressive Conditional Heteroskedasticity, *Journal of Econometrics* **31**, 307-327.

[14] Cagetti, Marco, Lars Peter Hansen, Thomas Sargent and Noah Williams (2002), Robustness and Pricing with Uncertain Growth, *Review of Financial Studies* **15**, 363-404.

[15] Cai, Jun (1994), A Markov Model of Switching-Regime ARCH, *Journal of Business and Economic Statistics* 12, 309-316.

[16] Calvet, Laurent E., and Adlai J. Fisher (2001), Forecasting Multifractal Volatility, *Journal of Econometrics* **105**, 27-58.

[17] Calvet, Laurent E., and Adlai J. Fisher (2002), Multifractality in Asset Returns: Theory and Evidence, *Review of Economics and Statistics* **84**, 381-406.

[18] Calvet, Laurent E., and Adlai J. Fisher (2004), How to Forecast Long-Run Volatility: Regime-Switching and the Estimation of Multifractal Processes, *Journal of Financial Econometrics* **2**, 49-83.

[19] Calvet, Laurent E., Adlai J. Fisher, and Benoit Mandelbrot (1997), Cowles Foundation Discussion Paper No. 1164-1166, Yale University. Papers available from http://www.ssrn.com.

[20] Campbell, John Y. (1991), A Variance Decomposition for Stock Returns, *Economic Journal* **101**, 157-179.

[21] Campbell, John Y. (2003), Consumption-Based Asset Pricing, in G. Constantinides and M. Harris eds., *Handbook of the Economics of Finance*, Amsterdam: North-Holland.

[22] Campbell, John Y., and John Cochrane (1999), By Force of Habit: A Consumption-Based Explanation of Aggregate Stock Market Behavior, *Journal of Political Economy* **107**, 205-251.

[23] Campbell, John Y., and Ludger Hentschel (1992), No News is Good News: An Asymmetric Model of Changing Volatility in Stock Returns, *Journal of Financial Economics* **31**, 281-318.

[24] Campbell, John Y., and N. Gregory Mankiw (1989), Consumption, Income and Interest Rates: Reinterpreting the Time Series Evidence, in *NBER Macroeconomics Annual: 1989*, O. Blanchard and S. Fischer eds., pp. 185-216. Cambridge, MA: MIT Press.

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

[26] Campbell, John Y., and Luis Viceira (2002), *Strategic Asset Allocation: Portfolio Choice for Long-Term Investors*, Oxford University Press.

[27] Cecchetti, Stephen, Pok-Sang Lam and Nelson Mark (1990), Mean Reversion in Equilibrium Asset Prices, *American Economic Review* **80**, 398-418.

[28] Dai, Qiang, Singleton, Kenneth J., and Wei Yang (2003), Regime Shifts in a Dynamic Term Structure Model of U.S. Treasury Bond Yields, Working Paper, New York University.

[29] Cochrane, John (2001), *Asset Pricing*, Princeton University Press.

[30] David, Alexander (1997), Fluctuating Confidence in Stock Markets: Implications for Returns and Volatility, *Journal of Financial and Quantitative Analysis* **32**, 427-462.

[31] Detemple, Jérôme B. (1986), Asset Pricing in a Production Economy with Incomplete Information, *Journal of Finance* **41**, 383-391.

[32] Detemple, Jérôme B. (1991), Further Results on Asset Pricing with Incomplete Information, *Journal of Economic Dynamics and Control* **15**, 425-453.

[33] Dothan, Michael U., and David Feldman (1986), Equilibrium Interest Rates and Multiperiod Bonds in a Partially Observable Economy, *Journal of Finance* **41**, 369-382.

[34] Engle, Robert (1990), Discussion: Stock Market Volatility and the Crash of '87, *Review of Financial Studies* **3**, 103-106.

[35] Epstein, Larry, and Stanley Zin (1989), Substitution, Risk Aversion and the Temporal Behavior of Consumption and Asset Returns: A Theoretical Framework, *Econometrica* **57**, 937-968.

[36] French, Kenneth, G. William Schwert, and Robert F. Stambaugh (1987), Expected Stock Returns and Volatility, *Journal of Financial Economics* **19**, 3-29.

[37] Gennotte, Gérard (1986), Optimal Portfolio Choice under Incomplete Information, *Journal of Finance* **41**, 733-749.

41

[38] Glosten, Lawrence, Ravi Jagannathan, and David Runkle (1993), On the Relation between the Expected Value and the Volatility of the Nominal Excess Return on Stocks, *Journal of Finance* **48**, 1779-1801.

[39] Gray, Stephen (1996), Modeling the Conditional Distribution of Interest Rates as a Regime-Switching Process, *Journal of Financial Economics* **42**, 27-62.

[40] Hamilton, James (1989), A New Approach to the Economic Analysis of Nonstationary Time Series and the Business Cycle, *Econometrica* **57**, 357-84.

[41] Hamilton, James (1990), Analysis of Time Series Subject to Change in Regime, *Journal of Econometrics* **45**, 39-70.

[42] Hamilton, James, and Raul Susmel (1994), Autoregressive Conditional Heteroskedasticity and Changes in Regime, *Journal of Econometrics* **64**, 307-333.

[43] Kandel, Shmuel, and Robert F. Stambaugh (1990), Expectations and Volatility of Consumption and Asset Returns, *Review of Financial Studies* **3**, 207-232.

[44] Kandel, Shmuel, and Robert F. Stambaugh (1991), Asset Returns and Intertemporal Preferences, *Journal of Monetary Economics* **27**, 39-71.

[45] Kim, Chang-Jin (1994), Dynamic Linear Models with Markov-Switching, *Journal of Econometrics* **60**, 1-22.

[46] Kim, Chang-Jin, James Morley and Charles Nelson (2002), Is There a Positive Relationship Between Stock Market Volatility and the Equity Premium?, forthcoming in *Journal of Money, Credit and Banking*.

[47] Lettau, Martin, Sidney Ludvigson, and Jessica Wachter (2003), The Declining Equity Premium: What Role Does Macroeconomic Risk Play?, Working Paper, New York University.

[48] Lucas, Robert (1978), Asset Prices in an Exchange Economy, *Econometrica* **46**, 1429-1445.

[49] Merton, Robert C. (1980), On Estimating the Expected Return on the Market: An Explanatory Analysis, *Journal of Financial Economics* **8**, 323-361.

[50] Naik, Vasant and Moon Lee (1997), Yield Curve Dynamics with Discrete Shifts in Economic Regimes: Theory and Estimation, Working Paper, University of British Columbia.

[51] Nelson, Daniel (1991), Conditional Heteroskedasticity in Asset Returns: A New Approach, *Econometrica* **59**, 347-370.

[52] Pindyck, Robert (1984), Risk, Inflation, and the Stock Market, *American Economic Review* **74**, 334-351.

[53] Poterba, James, and Lawrence Summers (1985), The Persistence of Volatility and Stock Market Fluctuations, *American Economic Review* **75**, 1142-1151.

[54] Schwert, G. William (1989), Why Does Stock Market Volatility Change Over Time?, *Journal of Finance* **44**, 1115-1153.

[55] Sentana, Enrique (1995), Quadratic ARCH Models, *Review of Economic Studies* **62**, 639-661.

[56] Smith, Daniel (2002), Markov-Switching and Stochastic Volatility Diffusion Models of Short-Term Interest Rates, *Journal of Business and Economic Statistics* **20**, 183-197.

[57] So, Mike, K. Lam, and W. K. Li (1996), A Stochastic Volatility Model with Markov-Switching, *Journal of Business and Economic Statistics* **16**, 244-253.

[58] Timmermann, Allan (1993), How Learning in Financial Markets Generates Excess Volatility and Predictability of Stock Returns, *Quarterly Journal of Economics* **108**, 1135-1145.

[59] Turner, Christopher, Richard Startz, and Charles Nelson (1989), A Markov Model of Heteroskedasticity, Risk and Learning in the Stock Market, *Journal of Financial Economics* **25**, 3-22.

[60] Veronesi, Pietro (1999), Stock Market Overreaction to Bad News in Good Times: A Rational Expectations Equilibrium Model, *Review of Financial Studies* **12**, 975-1007.

[61] Veronesi, Pietro (2000), How Does Information Quality Effect Stock Returns?, *Journal of Finance* **55**, 807-837.

[62] Veronesi, Pietro (2002), The Peso Problem Hypothesis and Stock Market Returns, Working Paper, University of Chicago, forthcoming in *Journal of Economic Dynamics and Control*.

[63] Vissing-Jørgensen, Annette (2002), Limited Asset Market Participation and the Elasticity of Intertemporal Substitution, Journal of Political Economy **110**, 825-853.

[64] Wachter, Jessica (2002), Habit Formation and Returns on Stocks and Bonds, Working paper, NYU Stern School of Business.

43

[65] Weil, Philippe (1989), The Equity Premium Puzzle and the Risk-Free Rate Puzzle, *Journal of Monetary Economics* **24**, 401-421.

[66] Whitelaw, Robert (1994), Time Variations and Covariations in the Expectation and Volatility of Stock Market Returns, *Journal of Finance* **49**, 515-541.

[67] Whitelaw, Robert (2000), Stock Market Risk and Return: An Equilibrium Approach, *Review of Financial Studies* **13**, 521-547.

[68] Wu, Guojun (2001), The Determinants of Asymmetric Volatility, *Review of Financial Studies* **14**, 837-859.

TABLE 1. – MOMENTS OF CRSP EXCESS RETURNS

	Entire Sample	By Subperiod			
		1	2	3	4
mean	0.016	0.022	-0.002	0.028	0.017
standard deviation	0.892	0.627	0.882	0.928	1.072
skewness	-1.10	0.12	0.06	-3.66	-0.23
kurtosis	27.79	7.66	4.62	76.99	7.15

Notes: This table reports statistics of the first four moments of CRSP value weighted excess returns. The statistics are reported for the entire sample and for four evenly spaced subsamples. There is considerable variability in all four moments across subsamples.

TABLE 2. – Regime-Switching Model with Full Information

	$\bar{k} = 1$	2	3	4	5	6	7	8
$\hat{m}_0$		1.608	1.532	1.487	1.442	1.383	1.374	1.339
		(0.010)	(0.010)	(0.010)	(0.012)	(0.013)	(0.014)	(0.013)
$\hat{\sigma}$		0.984	1.070	1.312	1.062	0.954	0.801	0.771
		(0.024)	(0.023)	(0.016)	(0.033)	(0.034)	(0.042)	(0.038)
$\hat{g}_d - r$		0.046	0.045	0.043	0.043	0.045	0.042	0.043
		(0.007)	(0.007)	(0.007)	(0.007)	(0.007)	(0.006)	(0.007)
$\hat{\gamma}k$		0.055	0.031	0.035	0.046	0.048	0.049	0.067
		(0.008)	(0.005)	(0.007)	(0.007)	(0.011)	(0.009)	(0.006)
$\hat{b}$		5.333	2.142	2.646	2.624	2.107	2.104	1.979
		(0.991)	(0.269)	(0.332)	(0.191)	(0.160)	(0.097)	(0.085)
$\ln L$		35459.06	35612.32	35644.40	35676.52	35684.52	35688.38	35686.98
$E[r_t - r]$		0.054	0.052	0.048	0.050	0.053	0.052	0.053
$Var[r_t - r]^{1/2}$		0.995	1.086	1.342	1.117	1.002	0.887	0.850
$Skew[r_t - r]$		-0.049	-0.055	-0.053	-0.056	-0.053	-0.052	-0.050
$Kurt[r_t - r]$		2.00	2.29	3.30	9.43	6.54	32.43	27.52
feedback		1.024	1.030	1.048	1.107	1.103	1.227	1.217

Notes: This table shows parameter estimates for the full-information regime-switching model for a number of volatility components ranging from one to eight.

TABLE 3. – COMPARISON WITH CH (1992)

A. Campbell-Hentschel Model Parameter Estimates

$\omega \times 10^7$	α_1	α_2	β	$b \times 10^3$	$\mu \times 10^4$	γ
-2.56	0.136	-0.076	0.932	3.73	3.01	0.35
(0.93)	(0.01)	(0.01)	(0.01)	(0.27)	(0.66)	(0.09)

B. Likelihood Comparison

	No. of Parameters	$\ln L$	BIC	BIC p-value vs. Multifractal Vuong (1989)	HAC Adj
Multifractal	5	35686.98	-6.9961		
QGARCH	7	35565.36	-6.9714	< 0.001	< 0.001

Notes: Panel A shows parameter estimates from the Campbell-Hentschel (1992) volatility feedback model based on a QGARCH specification. Panel B gives a comparison of the in-sample fit versus the multifrequency regime-switching specification. The Bayesian Information Criterion is given by $BIC = T^{-1}(-2\ln L + NP\ln T)$. The last two columns in Panel B give p-values from a test that the QGARCH dividend specification dominates the multifractal specification by the BIC criterion. The first value uses the Vuong (1989) methodology, and the second value adjusts the test for heteroskedasticity and autocorrelation. A low p-value indicates that the QGARCH specification would be rejected in favor of the multifrequency model.

TABLE 4. – Moment Comparison

Mean	Standard Deviation	Skewness	Kurtosis
	Data		
0.016	0.892	-1.10	27.79
	Full-Information Regime-Switching		
0.053	0.850	-0.050	27.52
(0.01)	(0.10)	(0.40)	(14.45)
[0.000]	[0.673]	[0.002]	[0.429]
	Campbell-Hentschel QGARCH		
0.033	0.885	-0.266	6.31
(0.01)	(0.08)	(0.20)	(3.34)
[0.028]	[0.582]	[0.009]	[0.994]

Notes: This table shows a moment-based comparison of the full information regime-switching feedback model against the CH (1992) QGARCH specification. For each model, we simulate a path the same length as the data 1,000 times. We calculate the first four moment statistics for each sample. The first line for each model gives the mean moments, and in parentheses in the second line the standard deviation across simulations. In brackets we report the simulated p-value of the data, which is the percentage of times the corresponding moment of the data exceeds the simulated moment of the model. We observe that the regime-switching model does not capture well the first and third moments, while the QGARCH model does not capture well the first, third, and fourth moments.

TABLE 5. – Regime-Switching Model with Daily Learning

	$\bar{k}=1$	2	3	4	5	6	7	8
$\hat{m}_0$		1.5867	1.4950	1.4450	1.4011	1.3633	1.3372	1.3083
$\hat{\sigma}$		0.947	0.947	0.942	0.947	0.947	0.947	0.942
$\hat{g_d} - r$		0.0017	-0.0061	-0.0083	-0.0083	-0.0083	-0.0083	-0.0083
$\hat{\gamma k}$		0.0600	0.0600	0.0600	0.0867	0.0600	0.100	0.1000
$\hat{b}$		5.3630	3.959	3.144	3.144	2.089	2.133	1.917
$\ln L$		-35402.25	-35528.50	-35562.82	-35581.20	-35594.16	-35599.02	-35603.83

Notes: This table shows maximum-likelihood parameter estimates for the regime-switching model with daily learning. [Preliminary results. Standard errors to be completed.]

TABLE 6. – SMM Estimates of Intraday Learning Model

	$\bar{k} = 1$	2	3	4	5	6	7	8
$\hat{m}_0$					1.426	1.408	1.375	1.375
$\hat{\sigma}$					0.925	0.827	0.762	0.858
$\hat{g}_d - r$					0.0126	0.0078	0.0093	0.0078
$\hat{\gamma}k$					0.697	0.898	0.863	0.936
$\hat{b}$					9.334	9.056	5.283	3.957
$\hat{N}$					9	9	15	9
$G(W)$					5.504	1.437	1.453	2.592
p-val					0.019	0.231	0.228	0.107
$E[r_t - r_f]$					0.0211	0.0171	0.0200	0.0158
$Var[r_t - r_f]$					0.876	0.701	0.635	0.757
$Skew[r_t - r_f]$					-0.302	-0.281	-0.183	-0.433
$Kurt[r_t - r_f]$					9.66	8.15	10.15	8.88

Notes: This table shows simulated method of moments parameter estimates for the regime-switching model with intraday learning. We report the optimized value of the SMM objective function, which is asymptotically distributed as a Chi-square distribution with one degree of freedom. The associated p-value and the estimated first four moments of returns are also reported. [Preliminary results. Standard errors to be completed.]
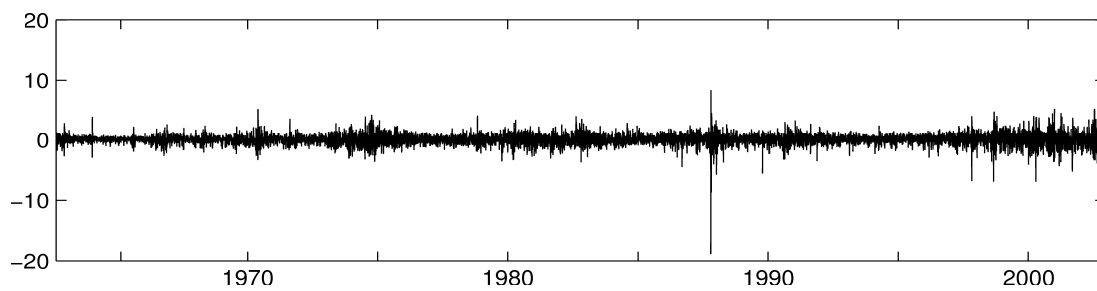
Figure 1: Daily CRSP Excess Returns. This figure shows daily CRSP excess returns from July, 1963 to December, 2002. The index series is value weighted, and the risk free rate is proxied by the return on 30 day U.S. Treasury bills.
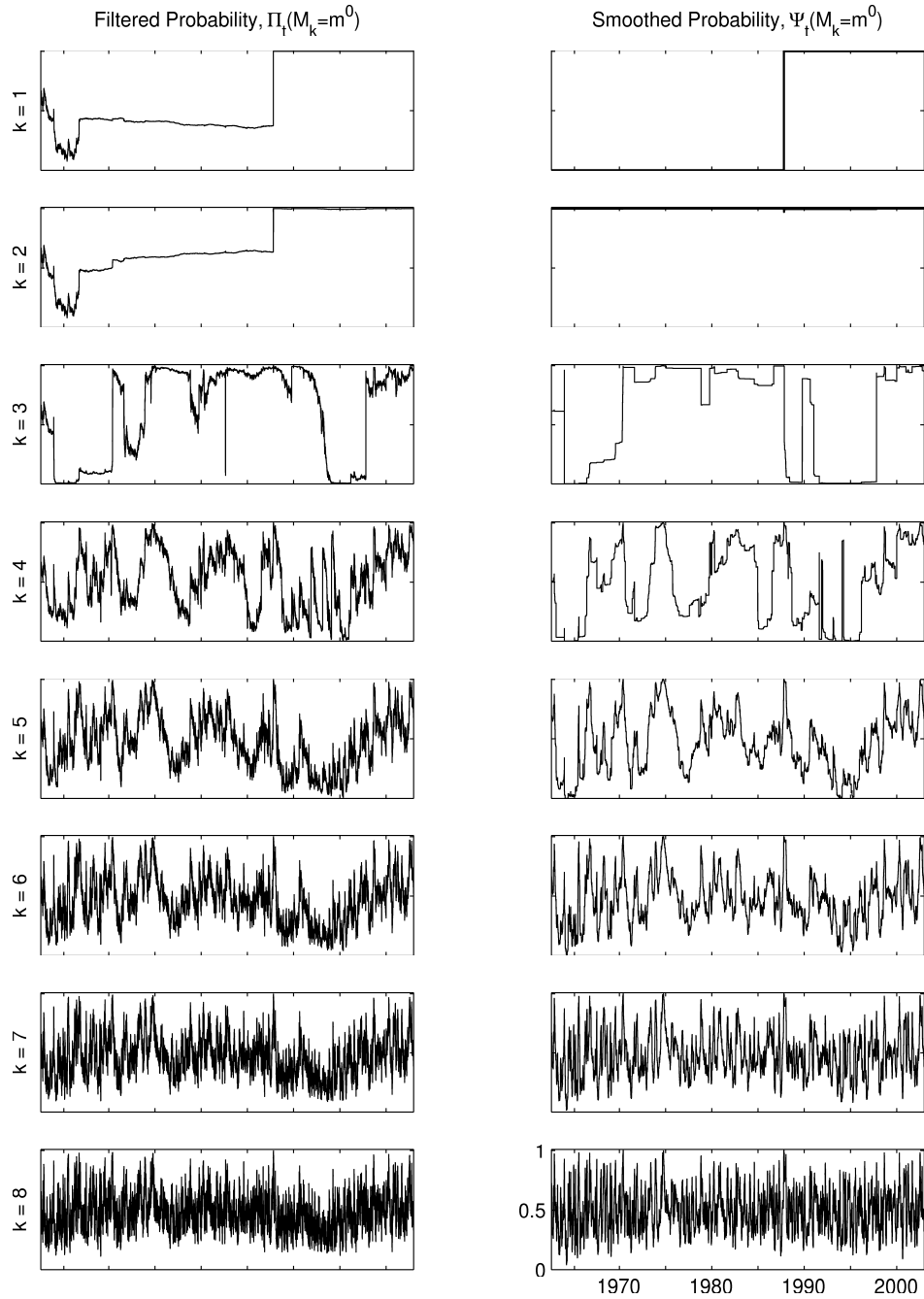
Figure 2: Volatility Component Conditional Beliefs. This figure shows *ex ante* and *ex post* conditional beliefs for the values of each volatility component in the full information regime-switching specification with $\bar{k} = 8$ components. The filtered probabilities Π_t are in the left-hand column, and the smoothed probabilites Ψ_t are in the right-hand column. Volatility components progress from low ($k = 1$) to high ($k = 8$) frequency from top to bottom of the figure.

52

Conditional Mean, $E_t(r_{t+1})$

0.1

(%) 0.05

0

Conditional Volatility, $[Var_t(r_{t+1})]^{1/2}$

3

2

(%)

1

0

Absolute Conditional Feedback: $Var_t(r_{t+1}) - Var_t(d_{t+1})$

0.2

0.1

0

Conditional Dividend News Variance: $Var_t(d_{t+1})$

8

6

4

2

0

Proportional Conditional Feedback, $[Var_t(r_{t+1}) - Var_t(d_{t+1})]/Var_t(d_{t+1})$

250
200
150
(%) 100
50
0

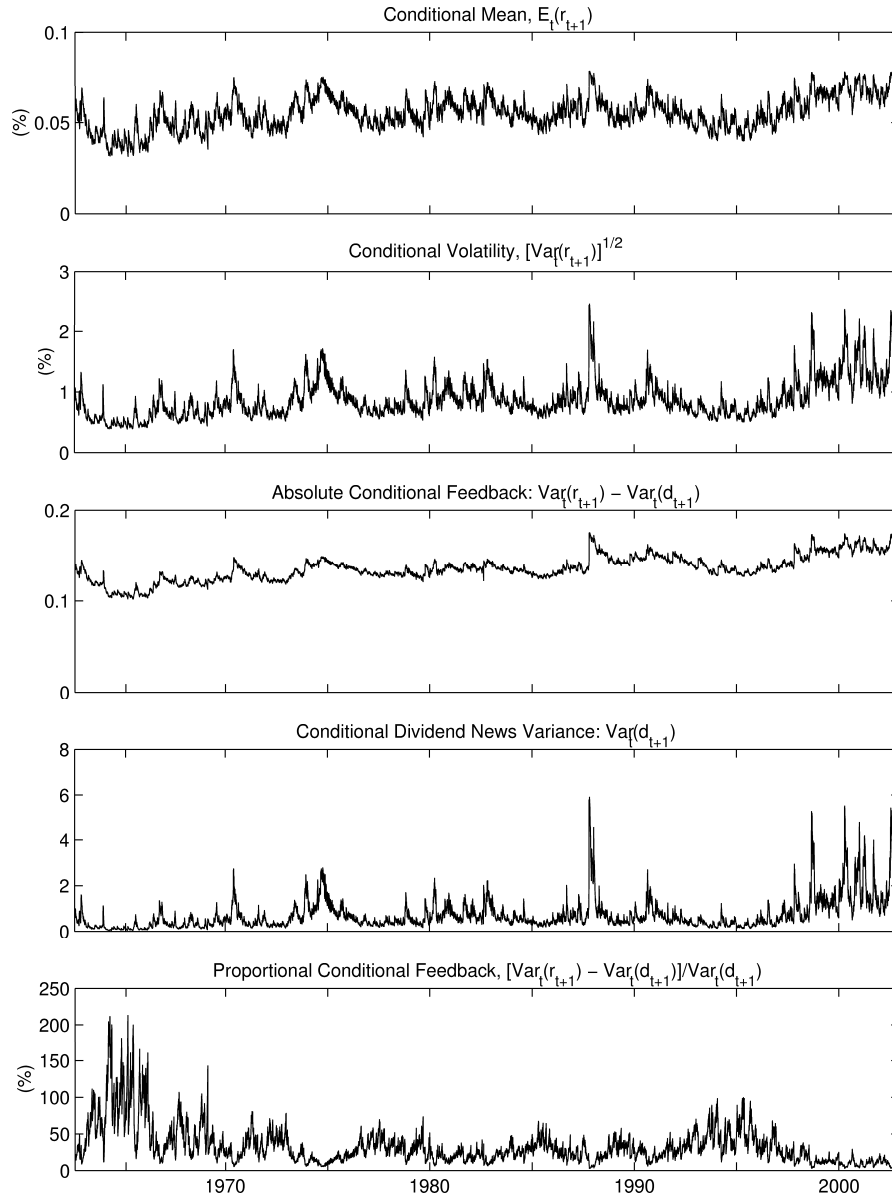1970   1980   1990   2000

Figure 3: Ex Ante Conditional Mean, Volatility, and Feedback. This figure shows conditional moments of value weighted CRSP excess returns under the full information regime switching specification with $\bar{k} = 8$ volatility components. Conditioning information is the *ex ante* information set of returns up to and including date t. The first panel shows conditional mean, and the second conditional variance. The final three panels give a decomposition of conditional variance. The "absolute conditional feedback" is the non-normalized amount of return variance contributed by feedback effects. The "conditional dividend news variance" is the denominator or normalizing factor for standard feedback calculations as in Campbell and Hentschel (1992). The final panel shows "proportional conditional feedback," which is the standard calculation of volatility feedback. This is obtained by dividing the values in the third panel by the fourth panel.
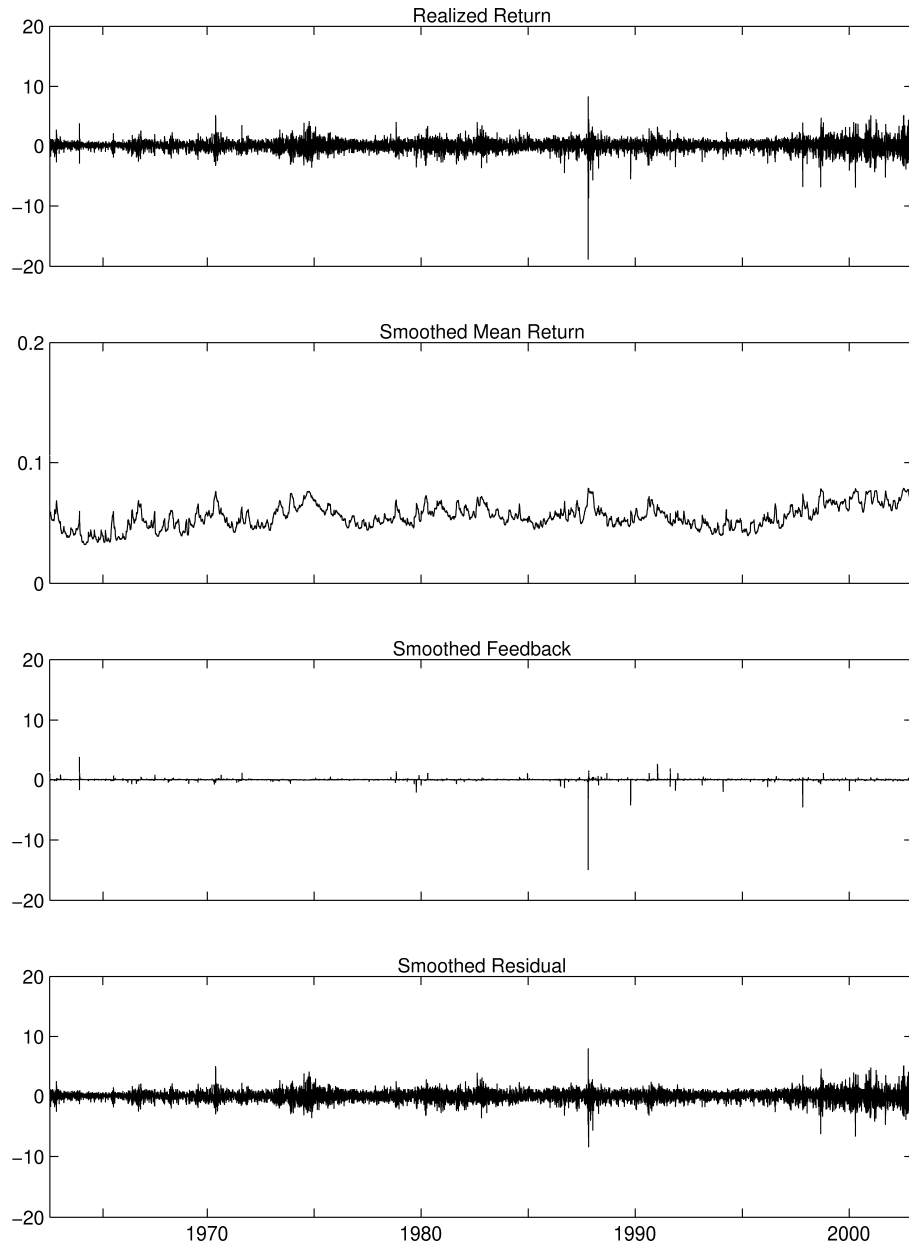
Figure 4: Ex Post Return Decomposition. This figure shows an *ex post* decomposition of realized returns using the full information regime-switching feedback model. The decomposition uses the smoothed beliefs Ψ_t obtained by using the conditioning information set of all returns. The first panel shows actual returns. The second panel shows the mean return at time $t + 1$ conditional on the beliefs Ψ_t. The third panel shows the estimated amount of returns due to volatility feedback at time $t + 1$ conditional on all of the data. The final panel is the residual, or the realized return less the second and third panels.