

NBER WORKING PAPER SERIES

MULTIFREQUENCY NEWS AND STOCK RETURNS

Laurent E. Calvet
Adlai J. Fisher

Working Paper 11441
<http://www.nber.org/papers/w11441>

NATIONAL BUREAU OF ECONOMIC RESEARCH
1050 Massachusetts Avenue
Cambridge, MA 02138
June 2005

Calvet: Department of Finance and Economics, HEC School of Management, 78351 Jouy-en-Josas Cedex, France, and NBER, calvet@hec.fr. Fisher: Sauder School of Business, University of British Columbia, 2053 Main Hall, Vancouver, BC V6T 1Z2, adlaifisher@sauder.ubc.ca. We received helpful comments from Andrew Abel, John Campbell, John Geanakoplos, Bruno Solnik, Robert Stambaugh, Amir Yaron, Jessica Wachter, and seminar participants at CREST, Paris I Panthéon-Sorbonne, and the Wharton School. We are very appreciative of financial support provided for this project by the HEC Foundation, the UBC Bureau of Asset Management, and the Social Sciences and Humanities Research Council of Canada. The views expressed herein are those of the author(s) and do not necessarily reflect the views of the National Bureau of Economic Research.

©2005 by Laurent E. Calvet and Adlai J. Fisher. All rights reserved. Short sections of text, not to exceed two paragraphs, may be quoted without explicit permission provided that full credit, including © notice, is given to the source.

Multifrequency News and Stock Returns
Laurent E. Calvet and Adlai J. Fisher
NBER Working Paper No. 11441
June 2005
JEL No. G12, C22

ABSTRACT

Recent research documents that aggregate stock prices are driven by shocks with persistence levels ranging from daily intervals to several decades. Building on these insights, we introduce a parsimonious equilibrium model in which regime-shifts of heterogeneous durations affect the volatility of dividend news. We estimate tightly parameterized specifications with up to 256 discrete states on daily U.S. equity returns. The multifrequency equilibrium has significantly higher likelihood than the classic Campbell and Hentschel (1992) specification, while generating volatility feedback effects 6 to 12 times larger. We show in an extension that Bayesian learning about stochastic volatility is faster for bad states than good states, providing a novel source of endogenous skewness that complements the "uncertainty" channel considered in previous literature (e.g., Veronesi, 1999). Furthermore, signal precision induces a tradeoff between skewness and kurtosis, and economies with intermediate investor information best match the data.

Laurent E. Calvet
Department of Finance and Economics
HEC School of Management
78351 Jouy-en-Josas Cedex
France
and NBER
calvet@hec.fr

Adlai J. Fisher
Sauder School of Business
University of British Columbia
2053 Main Mall
Vancouver, BC V6T 1Z2
Canada
adlai.fisher@sauder.ubc.ca

1. Introduction

The recent asset pricing literature suggests that stock prices are driven by shocks with very heterogeneous degrees of persistence. For example, market returns are predictable at a range of business-cycle horizons; some variables provide best forecasts over intervals of twelve months or less while others increase in power out to five years and more.¹ Complementary studies emphasize even more persistent sources of variations in returns, including technological innovation (e.g., Greenwood and Jovanovic, 1999), exogenous demographic changes (e.g., Abel, 2003), and low frequency movements in consumption or dividend growth (Bansal and Yaron, 2004).²

At the other end of the spectrum, high-frequency returns permit a large number of observations and thus potentially more precise econometric inference. Researchers have correspondingly related daily and intraday price movements to weather news (Roll, 1984), macroeconomic announcements (Andersen, Bollerslev, Diebold, and Vega, 2004), internet bulletin boards (Antweiler and Frank, 2004), analyst reports³ and corporate announcements.⁴ The existing literature thus documents persistence levels in stock market news ranging from daily intervals to several decades.

Our research builds on this evidence by developing an equilibrium framework with news shocks at many different frequencies. Why has the existing literature not addressed this agenda? We offer two explanations.

First, it might seem plausible that some financial questions are best addressed in isolation at a single horizon. Following this logic, equity premium studies largely focus on annual data, while market efficiency research mainly uses higher frequencies. Recent evidence suggests, however, that fluctuations at different frequencies can interact. Lochstoer (2004) thus shows that a slowly-evolving generational state variable controls business-cycle variation in risk premia. Similarly, Andersen, Bollerslev, Diebold, and Vega (2004) provide evidence that the high-frequency impact of macroeconomic news depends on the state of the business cycle.

A second, and more pragmatic, impediment to multifrequency research is that complexity grows quickly with the number of components. Recent contributions (e.g., Bansal and Yaron, 2004; Lochstoer, 2004) demonstrate not only the empirical advantage of using two persistence levels, but also that formal estimation becomes more difficult and high-dimensional calibration more necessary with a more complex setup.

¹See, for example, Lettau and Ludvigson (2001).

²Endogenous mechanisms for extremely persistent fluctuations include habit formation (Campbell and Cochrane, 1999), durable consumption goods (Yogo, 2004a), and the dynamics of capital accumulation with adjustment costs (Jermann, 1998).

³See, e.g., Womack (1996).

⁴An extensive literature studies stock price reactions to corporate announcements including quarterly earnings, dividend policy, securities issuances and changes of control. See, e.g., MacKinlay (1997).

Our paper proposes a new direction to address these problems by developing a parsimonious equilibrium framework based on recent advances in multifrequency econometrics. An Epstein-Zin consumer receives an exogenous consumption stream, and prices a flow of correlated dividends with regime-switching in the mean and volatility of their growth rates.⁵ The model thus follows the recent trend to model dividends and consumption as correlated but not identical processes (e.g., Campbell and Cochrane, 1999). Exact solutions for equilibrium prices, return dynamics, and filtered beliefs are available. Unlike previous Lucas tree economies considered in the literature (e.g., Bansal and Yaron, 2004; Lettau, Ludvigson, and Wachter, 2004), our setup implies that higher volatility reduces prices for any level of the elasticity of intertemporal substitution.

We specify news arrivals with a Markov-switching multifractal (MSM), a stochastic volatility model characterized by a small number of parameters but an arbitrarily large number of frequencies (Calvet and Fisher, 2001, 2002, 2004; Calvet, Fisher, and Thompson, 2004). Under this specification, news volatility is hit by exogenous shocks with highly heterogeneous durations, which range from one day to more than a decade in empirical applications. Earlier work shows that MSM captures the outliers, volatility persistence and power variation⁶ of financial series, while permitting maximum likelihood estimation and analytical multi-step forecasting. MSM compares favorably with standard volatility models such as GARCH(1, 1) both in- and out-of-sample (Calvet and Fisher, 2004). It is now natural to embed it into an equilibrium framework.

The multifrequency equilibrium model inherits the appealing properties of MSM. It is tightly parameterized and permits structural estimation by maximum likelihood. We estimate our specification on an index⁷ of US equities over the period 1926-2004. Versions of the model with six to eight volatility frequencies provide significant improvements in likelihood relative to lower dimensional specifications. The model also improves on earlier specifications of single frequency news arrivals (Campbell and Hentschel, 1992, hereafter “CH”), even though our approach uses fewer parameters.

Our model generates volatility feedback, the property that upward revisions to anticipated future volatility tend to decrease current returns. Consistent with a multifrequency perspective, previous researchers have studied this topic at a range of different

⁵Following Hamilton (1989, 1990), researchers have used regime-switching to help explain financial phenomena including stock market volatility, return predictability, the relation between conditional risk and return, the term structure of interest rates, and the recent growth of the stock market. Contributions include Abel (1994, 1999), Bansal and Zhou (2002), Cecchetti, Lam and Mark (1990), David (1997), Kandel and Stambaugh (1990), Lettau, Ludvigson and Wachter (2004), Turner, Startz and Nelson (1989), Veronesi (1999, 2000, 2004), Wachter (2004), and Whitelaw (2000).

⁶Power variation relates to the behavior at small time scales of sums of powers of absolute values of returns. See Calvet and Fisher (2002), Barndorff-Nielsen and Shephard (2003) and Andersen, Bollerslev, and Diebold (2003).

⁷We splice the Schwert (1990a) and value-weighted CRSP indices, as in CH.

horizons. For example, French, Schwert, and Stambaugh (1987, hereafter “FSS”), CH, and Wu (2001) assess feedback effects in daily, weekly, and monthly data, while Pindyck (1984), Poterba and Summers (1985), Bansal and Yaron (2004), and Lettau, Ludvigson and Wachter (2004) emphasize volatility movements at the business cycle range and beyond.⁸ A multifrequency approach might therefore prove useful in this context. Intuition suggests that high-frequency volatility shocks help to capture the dynamics of typical day-to-day variations, while lower-frequency movements generate the strong feedback required to fit the most extreme daily returns. Volatility feedback models are thus a natural setting where the interaction of various frequencies seems intuitively important. The paper can be viewed in this sense as a first step towards bringing together branches of the lower-frequency macro-finance and higher-frequency financial econometrics literature.

The multifrequency equilibrium generates substantially larger feedback effects than previous research. For instance, CH find that feedback amplifies the volatility of dividend news by only about 2%; they attribute this result to the property of GARCH-type specifications that the volatility of volatility can only be large if volatility itself is high. In our stochastic volatility MSM specification for dividend news, feedback rises with the number of components and the likelihood function, increasing to between 12% and 24% for the preferred specifications with six to eight components. The multifrequency equilibrium model thus generates an unconditional feedback that is 6 to 12 times larger than in previous literature.⁹

A substantial level of endogenous skewness is difficult to obtain in our full-information equilibrium with symmetric dividends. Earlier volatility feedback studies attempt to address this by introducing predictive asymmetry (CH) or skewness (Wu, 2001) directly into the econometric specification of dividends. Our work instead investigates whether higher return moments can be modeled through the endogenous equilibrium implications of imperfect investor information and learning. We thus generalize our setup to allow that the investor observes noisy signals of the volatility components and then makes Bayesian inferences about the latent volatility state. The separation of dividends from consumption implies that the price:dividend ratio is linear in investor beliefs, making the model tractable.

Our learning model generates two main sets of results. First, signal precision has little effect on the unconditional mean and variance of stock returns. To explain this,

⁸Investigation of volatility feedback in a general equilibrium setting was pioneered by Barsky (1989) in a two-period setting and Abel (1988) in the dynamic case. French, Schwert and Stambaugh (1987) and Campbell and Hentschel (1992) use GARCH-type processes to show that ex-post returns are negatively affected by positive innovations in volatility. Bekaert and Wu (2000) provide further support for this hypothesis.

⁹Based on the parameter estimates presented in Wu (2001), unconditional volatility feedback is 3.5% for his model estimated on monthly data, and is negative for his model estimated on weekly data.

we show that the price:dividend ratio (P:D) in the learning model is the conditional expectation of its full-information counterpart. This implies the same mean and lower variance, which is the analogue in our setup of the variance bounds discussed by LeRoy and Porter (1981) and Shiller (1981). In our model, however, the reduction in variance is negligible because large movements of the P:D ratio are induced by shifts in the most persistent volatility components. Learning about these rare changes is therefore a transitory phenomenon that has limited impact on stock return variance, and we verify this logic numerically. The sizeable volatility feedback implied by our full-information specification is thus robust to changes in information quality.

Our second learning result is that changes in the most persistent components have strong effects on the higher moments of returns. In particular, varying the precision of the volatility signal generates a sizeable tradeoff between endogenous skewness and kurtosis. When investors have perfect information, volatility shocks are incorporated fully and immediately into price, regardless of the direction of change. By contrast, when the volatility component signals are poor, investors rely on dividend news to make inferences about the volatility state. They still learn quickly about volatility increases, because a single extreme fluctuation is highly improbable with low volatility. Learning about reduced volatility must be slow, however, because dividend news observations near the mean are a relatively likely outcome regardless of the true volatility state. Thus, bad news about volatility incorporates quickly into price, while good news trickles out slowly.¹⁰ This asymmetry creates the observed tradeoff between endogenous skewness and kurtosis as information quality changes.

The linearity of the P:D ratio distinguishes our framework from previous research that builds on the convexity of the price function with respect to beliefs (e.g., Brennan and Xia, 2001; David, 1997; Lettau, Ludvigson and Wachter, 2004; Veronesi, 1999, 2000). Convexity leads to skewness because price is more sensitive to bad news in good times than it is to good news in bad times. The origin of convexity in these models is that signals about the latent state jointly influence investor beliefs about cash flows and investor beliefs about the stochastic discount factor. This comovement creates an “uncertainty” price discount, which is most pronounced when beliefs are intermediate and therefore most variable. By contrast, investors in our framework learn only about cash flows. The endogenous skewness mechanism identified in our paper thus complements the “uncertainty” channel considered in previous learning studies.

Our work also complements earlier research by Veronesi (2000) on how information quality affects stock returns. Whereas Veronesi considers a latent dividend drift, we investigate learning about volatility and show that signal precision then has powerful

¹⁰A statistical interpretation is that odds ratios tend to be larger in the tails of competing distributions than in the bells. This implies that learning about volatility increases should be faster than learning about volatility decreases.

effects in a multifrequency environment. Further, our study directly demonstrates that structural models of investor learning are empirically relevant to higher-frequency daily stock returns, which is an advance to this literature.¹¹

Finally, we emphasize the considerable challenge of using a five parameter economic equilibrium to explain more than 20,000 observations of daily stock returns over a time span exceeding eight decades. The model must account for high-frequency regularities such as thick tails, volatility persistence, skewness, and feedback. It must also be flexible enough to account for market conditions that change considerably over such a long time span (Schwert, 1989). Moreover, our approach is structural in the sense that economic theory places strong restrictions on what would otherwise be degrees of freedom in a purely statistical setup. Our exogenous dividend news process is conditionally Gaussian, has constant mean, and is both conditionally and unconditionally symmetric. Thus, endogenous equilibrium effects must play a significant role in explaining stock returns.

Section 2 presents the asset pricing model and the equilibrium solution for a general 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. All proofs are in the Appendix.

2. An Asset Pricing Model with Regime-Switching Dividends

This section develops a discrete-time equilibrium model with regime-shifts in the mean and volatility of dividend growth. The model resolves a well-known difficulty in the volatility feedback literature. In a Lucas (1978) tree economy, an increase in consumption volatility affects the pricing kernel and thus reduces the aggregate market price only under special choices of relative risk aversion (e.g., Abel, 1989; Whitelaw, 2000) or the elasticity of intertemporal substitution (Bansal and Yaron, 2004; Lettau, Ludvigson, and Wachter, 2004). In our model, dividend news volatility has no impact on the pricing kernel and thus generates a negative relation between volatility and prices for all preference parameters.

¹¹Contributions to the literature on learning in financial markets include Detemple (1986), Dothan and Feldman (1986), Genotte (1986), Timmermann (1993, 1996), David (1997), Veronesi (1999, 2002), Brennan and Xia (2001), and Lettau, Ludvigson and Wachter (2004). Empirical implementation tends to focus on calibration at lower frequencies. For example, Veronesi (2004) 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. David (1997) and Brennan and Xia (2001) calibrate at a monthly frequency. Guidolin and Timmermann (2003) develop estimation and forecasting for a model of learning about the drift on a binomial lattice, and apply this to pricing options at a weekly frequency. At a monthly frequency, Turner, Startz, and Nelson (1989) and Kim, Morley, and Nelson (2004), consider learning about volatility in a two-state specification with feedback effects, where the signals that drive investor learning are not specified.

2.1. Preferences, Consumption and Dividends

We consider an exchange economy defined on the regular grid $t = 0, 1, 2, \dots, \infty$. 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 $\theta = (1 - \alpha)/(1 - \psi^{-1})$. When $\alpha = \psi^{-1}$, the specification reduces to expected utility. While the Epstein-Zin-Weil preferences facilitate comparison with earlier studies, most of our results hold with standard expected utility.

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:

$$c_t - c_{t-1} = g_c + \sigma_c \varepsilon_{c,t}, \quad (2.1)$$

where the shocks $\{\varepsilon_{c,t}\}$ are IID $\mathcal{N}(0, 1)$. This standard specification is consistent with the empirical evidence that consumption growth is approximately IID in postwar US consumption data (e.g., Campbell, 2003).¹²

The volatility feedback literature suggests that aggregate stock prices decrease with the volatility of dividend news. When the stock market is a claim on aggregate consumption, this negative relation arises in equilibrium only for specific values of the preference parameters. For instance, under expected utility $\alpha = 1/\psi$, the price:dividend ratio declines with volatility only if risk aversion is less than one (Barsky, 1989; Abel, 1988).¹³ For arbitrary Epstein-Zin preferences, volatility reduces prices only if the EIS is strictly larger than 1 and relative risk aversion differs from unity: $\psi > 1$ and $\alpha \neq 1$ (Bansal and Yaron, 2004; Lettau, Ludvigson and Wachter, 2004).¹⁴ The empirical validity of the

¹²Bansal and Yaron (2004) argue that consumption growth contains a small but highly persistent component. This is reasonable but in the raw consumption data difficult to distinguish from the iid case, which is still the predominant assumption in the asset pricing literature. Future work might consider incorporating into our framework the type of consumption growth suggested by Bansal and Yaron.

¹³The price:dividend ratio is then $Q_t \equiv P_t/C_t = \mathbb{E}_t \sum_{n=1}^{+\infty} \delta^n (C_{t+n}/C_t)^{1-\alpha}$, or equivalently

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

When future consumption becomes riskier, the ratio is affected by two opposite effects. First, the covariances become more negative and *reduce* the price:dividend ratio Q_t , as desired. Second, the precautionary motive increases the expected marginal utility of future consumption $\mathbb{E}_t [(C_{t+n}/C_t)^{-\alpha}]$, which lowers interest rates and tends to *increase* Q_t . We eliminate the second effect by disentangling volatility shocks to the stock market from aggregate consumption.

¹⁴The Euler equation is then $Q_t^\theta = \delta^\theta \mathbb{E}_t [(C_{t+1}/C_t)^{1-\alpha} (1 + Q_{t+1})^\theta]$. When consumption growth is

EIS restriction has not been resolved. Attanasio and Weber (1993), Vissing-Jørgensen (2002), and Bansal and Yaron (2004) report estimates of ψ larger than 1, while Campbell and Mankiw (1989), Campbell (2003), and Yogo (2004b) find ψ to be small and in many cases statistically indistinguishable from zero.

We resolve these difficulties by: 1) separating dividends from consumption, and 2) permitting that shocks to dividend volatility do not simultaneously impact consumption. Specifically, the log-dividend $d_t = \ln D_t$ follows a random walk with state-dependent drift and volatility:

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

where $\varepsilon_{d,t}$ is IID $\mathcal{N}(0,1)$ and correlated with $\varepsilon_{c,t}$, the drift μ_d and volatility σ_d are deterministic functions of the state variable M_t , and the Itô term $\sigma_d^2(M_t)/2$ guarantees that expected dividend growth $\mathbb{E}[D_t/D_{t-1}|M_t] = e^{\mu_d(M_t)}$ is controlled only by $\mu_d(M_t)$. We assume that the state variable M_t is a first-order Markov process, and leave the exact specification of the drift and volatility fully general in the rest of Section 2.

The model thus separates stock returns from aggregate consumption growth and the stochastic discount factor. This assumption, which is common in finance, is consistent with a variety of empirical facts.¹⁵ 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 returns substantially fluctuates through time. 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. Consumption and dividend shocks should thus be correlated, but not identical.

2.2. Asset Pricing under Complete Information

The information available to the investor is a key variable of the model. To develop intuition, we begin by considering that the agent directly observes 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 investor information set $I_t = \{(C_s, D_s, M_s); s \leq t\}$ thus characterizes *full information* economies.

IID, the price dividend ratio is constant and satisfies $Q/(1+Q) = \delta \{\mathbb{E}[(C_{t+1}/C_t)^{1-\alpha}]\}^{1/\theta}$. It decreases with volatility if $(1-\alpha)/\theta > 0$ or equivalently $\psi > 1$ and $\alpha \neq 1$.

¹⁵See for instance Campbell (1996) or Campbell and Cochrane (1999).

The stochastic discount factor satisfies

$$SDF_{t+1} = \delta' \left(\frac{C_{t+1}}{C_t} \right)^{-\alpha}, \quad (2.2)$$

where $\delta' = \delta \{ \mathbb{E}[(C_{t+1}/C_t)^{1-\alpha}] \}^{\frac{1}{\theta}-1}$. This expression is proportional to the stochastic discount factor obtained under expected utility ($\theta = 1$), suggesting that the elasticity of intertemporal substitution affects the interest rate but not the price of risk.

We now turn to equilibrium pricing. The interest rate is constant through time and satisfies the familiar relationship: $r_f = -\ln \mathbb{E}_t(SDF_{t+1}) = -\ln \delta' + \alpha g_c - (\alpha \sigma_c)^2/2$. The equilibrium stock price is proportional to the current dividend: $P_t = Q(M_t)D_t$. The Markov state thus controls the P:D ratio. The gross return on the stock

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

satisfies the pricing condition $\mathbb{E}[(1 + R_{t+1})SDF_{t+1}|I_t] = 1$, or equivalently

$$\delta' \mathbb{E} \left[\left(\frac{C_{t+1}}{C_t} \right)^{-\alpha} \frac{D_{t+1}}{D_t} \frac{1 + Q(M_{t+1})}{Q(M_t)} \middle| I_t \right] = 1.$$

The price:dividend ratio therefore solves the fixed-point equation

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

where $\rho_{c,d} = Cov(\varepsilon_{c,t}, \varepsilon_{d,t}) > 0$ denotes the constant correlation between the Gaussian noises in consumption and dividends.

When the volatility 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 $Q(M_t) = \mathbb{E}_t \sum_{n=1}^{+\infty} [\prod_{h=1}^n e^{\mu_d(M_{t+h}) - r_f - \alpha \rho_{c,d} \sigma_c \sigma_d(M_{t+h})}]$. High volatility thus feeds into low asset prices for any choices of the relative risk aversion α and the EIS ψ .

By (2.3), the log excess return $r_{t+1} \equiv \ln(1 + R_{t+1}) - r_f$ is determined by the price:dividend ratio and the realization of the dividend growth:

$$r_{t+1} = \ln \frac{1 + Q(M_{t+1})}{Q(M_t)} + \mu_d(M_{t+1}) - r_f - \frac{\sigma_d^2(M_{t+1})}{2} + \sigma_d(M_{t+1})\varepsilon_{d,t+1}. \quad (2.5)$$

Equations (2.4) and (2.5) are the building blocks of the empirical work.

In empirical applications, we conveniently assume that the Markov state M_t takes a finite number of values $\{m^1, \dots, m^d\}$. Fixed-point condition (2.4) implies that the equilibrium P:D ratio can be easily computed numerically for every possible state $Q(m^1), \dots, Q(m^d)$. Econometric inference is also straightforward. While the investor observes the true volatility state M_t , the econometrician has a smaller information set $I_t^0 \subseteq I_t$. Specifically, we assume that the econometrician only observes excess returns: $I_t^0 \equiv \{r_s\}_{s=1}^t$. We show in the Appendix that the likelihood function $L(r_1, \dots, r_T)$ is then available in closed-form.

3. Volatility Feedback with Multifrequency Shocks

A large body of research documents that corporate profits and dividends are hit by shocks with heterogeneous persistence. Classic asset pricing papers (e.g., CH) nonetheless assume an autoregressive specification in which volatility shocks decline at a single frequency. In standard setups, multifrequency shocks would of course require a large number of parameters. Recent econometrics literature (Calvet and Fisher, 2001, 2002, 2004) develops Markov-switching multifractal (MSM) processes as a tractable solution to this problem. We now adopt this specification for dividend news.

3.1. A Multifrequency Specification for Dividend News

We assume that the volatility state M_t follows an MSM process, as is now explained. The Markov state is a row vector

$$M_t = (M_{1,t}; M_{2,t}; \dots; M_{\bar{k},t}) \in \mathbb{R}_+^{\bar{k}},$$

where the scalars $M_{1,t}, M_{2,t}, \dots, M_{\bar{k},t}$ are volatility components with heterogeneous durations. Persistence is highest for the first component, and progressively diminishes with the component index k . MSM thus contains a finite number of volatility components, each of which corresponds to a different frequency.

Dividends have a constant growth rate

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

and a stochastic volatility equal to the renormalized product

$$\sigma_d(M_t) \equiv \bar{\sigma}_d (M_{1,t} M_{2,t} \dots M_{\bar{k},t})^{1/2}. \quad (3.1)$$

We assume that $\bar{\sigma}_d$ is a positive constant and that each random *multiplier* $M_{k,t}$ has an unconditional mean equal to unity: $\mathbb{E}M_{k,t} = 1$.

The volatility state vector is easily constructed through time.¹⁶ Let M_t denote the Markov state at date t . For each $k \in \{1, \dots, \bar{k}\}$, the next period multiplier $M_{k,t+1}$ is drawn from a fixed distribution M with probability γ_k , and is otherwise equal to its current value: $M_{k,t+1} = M_{k,t}$. The construction can be summarized as:

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

The switching events and new draws from M are independent across k and t . The volatility components $M_{k,t}$ differ in their transition probabilities but not in their marginal

¹⁶This 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.

distribution. Each component $M_{k,t}$ therefore follows a Markov process that is identical except for time scale. These features greatly contribute to the parsimony of the model.

MSM can accommodate any distribution M with positive support and unit mean. In this paper, we choose for parsimony a binomial taking values $m_0 \in [1; 2]$ and $2 - m_0 \in [0; 1]$ with equal probability. We tightly parameterize the transition probabilities by assuming

$$\gamma_k = 1 - (1 - \gamma_{\bar{k}})^{(b^k - \bar{k})}. \quad (3.2)$$

Calvet and Fisher (2001) introduce this specification through the discretization of a Poisson arrival process, and subsequent work demonstrates its empirical validity (Calvet and Fisher, 2004; Calvet, Fisher, and Thompson, 2004). Condition (3.2) implies that the transition probabilities grow approximately geometrically: $\gamma_k \sim \gamma_{\bar{k}} b^{k - \bar{k}}$. The parameter $\gamma_{\bar{k}}$ thus controls the persistence of the highest frequency component, while b determines the spacing between component frequencies.

The specification (3.1) is appealing to model the high variability and strong persistence of financial volatility. Low-frequency multipliers deliver persistent and discrete switches, consistent with evidence of apparent non-stationarity in return variance (e.g., Schwert, 1989; Pagan and Schwert, 1990).¹⁷ Persistent changes also have a strong impact on the P:D ratio and thus generate large feedback effects in stock returns. High-frequency multipliers give additional outliers through their direct effect on the tails of the dividend news process. Further, multiplicative interaction implies that subperiods of low volatility can be observed in highly volatile periods. Conversely, total volatility can quickly switch from an extreme to a normal level, as has been observed in equity data (e.g., Schwert, 1990b). We expect that these features of MSM will help to fit US stock returns over a long time span as well as to generate substantial volatility feedback.

The equilibrium excess return

$$r_{t+1} = \ln \frac{1 + Q(M_{t+1})}{Q(M_t)} + g_d - r_f - \frac{\sigma_d^2(M_{t+1})}{2} + \sigma_d(M_{t+1})\varepsilon_{d,t+1} \quad (3.3)$$

is fully specified by the MSM volatility dynamics and fixed-point condition (2.4). Volatility feedback manifest itself in the return equation through the term $\ln \frac{1 + Q(M_{t+1})}{Q(M_t)}$. Let $\sigma_{c,d} = \sigma_c \bar{\sigma}_d \rho_{c,d}$. By Euler condition (2.4), the price-dividend ratio is determined by six coefficients: $g_d - r_f$, $\alpha \sigma_{c,d}$ and the MSM volatility parameters $(m_0, \gamma_{\bar{k}}, b, \bar{\sigma}_d)$. As is standard in the literature (e.g. Campbell and Shiller, 1988), we calibrate the mean price:dividend ratio to its empirical value. This constraint guarantees that volatility feedback estimates do not arise from a counterfactually high share of dividends in stock

¹⁷Although our model is strictly stationary, even very long-samples would be difficult to distinguish from a non-stationary process due to low-frequency switches. We view this as a convenient framework to model the low-frequency uncertainty that is present in financial data.

returns. Specifically, consider

$$\ln \rho \equiv \mathbb{E} \ln \frac{Q(M_t)}{1 + Q(M_t)} = \mathbb{E} \ln \frac{P_t}{P_t + D_t}. \quad (3.4)$$

A unique $\alpha\sigma_{c,d}$ ensures that the price-dividend ratio matches the empirical value of ρ for any choice of the other five parameters.¹⁸

The return dynamics are thus specified by

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

where m_0 denotes the high level of a volatility component, $\gamma_{\bar{k}}$ is the transition probability of the least persistent component, b quantifies the growth rates of frequencies, and $g_d - r_f$ and $\bar{\sigma}_d$ quantify the mean and standard deviations of dividend news.

3.2. Loglinear Approximation

We now develop intuition for the multifrequency equilibrium by loglinearizing the pricing equation. Specifically, assume that the price-dividend ratio is loglinear in the volatility components:

$$\ln Q(M_t) \approx \bar{q} - \sum_{k=1}^{\bar{k}} q_k (M_{k,t} - 1). \quad (3.5)$$

Equilibrium fixed-point condition (2.4) implies that¹⁹

$$q_k = \frac{\alpha\sigma_{c,d}}{2} \frac{1 - \gamma_k}{1 - (1 - \gamma_k)\rho}, \quad (3.6)$$

$$\ln \left(\frac{e^{\bar{q}}}{1 + e^{\bar{q}}} \right) \equiv \ln \rho = g_d - r_f - \alpha\sigma_{c,d}. \quad (3.7)$$

We observe that each coefficient q_k increases with the probability $1 - \gamma_k$ of the corresponding component remaining in the current state. The price:dividend ratio is thus a persistence-weighted sum of the volatility components. High frequency components have negligible effects on the P:D ratio: $q_k \rightarrow 0$ when $\gamma_k \rightarrow 1$. On the other hand for very persistent components, the coefficient q_k is large since ρ is empirically close to one at the usual frequencies.

By (3.3) and (3.7), the unconditional expected return satisfies $\mathbb{E}r_t = \alpha\sigma_{c,d}$, as in the consumption CAPM. Volatility innovations thus have no impact the unconditional

¹⁸The price:dividend ratio $Q(M_t) = \mathbb{E}_t \sum_{n=1}^{+\infty} e^{n(g_d - r_f) - \alpha\sigma_{c,d}[\sqrt{g(M_{t+1})} + \dots + \sqrt{g(M_{t+n})}]}$ monotonically decreases from $+\infty$ to 0 as $\alpha\sigma_{c,d}$ increases from $-\infty$ to $+\infty$. Thus for every $(m_0, \gamma_{\bar{k}}, b, \bar{\sigma}_d, g_d - r_f)$ and $\rho < 1$, equation (3.4) has a unique solution $\alpha\sigma_{c,d}$.

¹⁹We show in the Appendix that the approximate solution holds for all choices of $\gamma_1, \dots, \gamma_{\bar{k}}$, and thus does not hinge on restrictions (3.2).

equity premium. Realized returns are of course affected by multifrequency shocks and satisfy

$$r_{t+1} \approx \alpha \sigma_{c,d} + \sum_{k=1}^{\bar{k}} q_k [(M_{k,t} - 1) - \rho(M_{k,t+1} - 1)] + \sigma_d(M_{t+1})\varepsilon_{d,t+1}. \quad (3.8)$$

The regimes generate large clustered outliers, as in equity data. We now investigate how they affect the conditional and unexpected components of returns.

Given investor information $I_t = \{(C_s, D_s, M_s); s \leq t\}$, the *conditional* return is the persistence-weighted sum of volatility components:

$$\mathbb{E}_t r_{t+1} \approx \alpha \sigma_{c,d} \left[1 + \frac{1}{2} \sum_{k=1}^{\bar{k}} (1 - \gamma_k)(M_{k,t} - 1) \right]. \quad (3.9)$$

Multipliers with high durations command a high expected return. We note that the formula contrasts with the relationships obtained in traditional volatility models, where the conditional return is typically a function of current volatility (e.g. Merton, 1980; CH). As in Cecchetti, Lam, and Mark (1990), the volatility components generate mean reversion in returns: $\mathbb{E}_t r_{t+n} \approx \alpha \sigma_{c,d} [1 + \sum_{k=1}^{\bar{k}} (1 - \gamma_k)^n (M_{k,t} - 1)/2] \rightarrow \alpha \sigma_{c,d}$ as $n \rightarrow \infty$. Note, however, that convergence 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 return *innovation* $u_{t+1} = r_{t+1} - \mathbb{E}_t r_{t+1}$ satisfies

$$u_{t+1} \approx -\rho \sum_{k=1}^{\bar{k}} q_k (M_{k,t+1} - \mathbb{E}_t M_{k,t+1}) + \sigma_d(M_{t+1})\varepsilon_{d,t+1}. \quad (3.10)$$

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 implies that volatility and returns are negatively correlated, which 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: $-\rho q_k (M_{k,t+1} - \mathbb{E}_t M_{k,t+1}) > 0$. On the other hand if volatility is initially high, no arrival is bad news for stock returns. Thus, 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. The stock market amplification of exogenous shocks is quantified by the *unconditional volatility feedback*

$$\frac{Var(r_{t+1})}{Var(d_{t+1} - d_t)} \approx 1 + \frac{Var(M)}{\bar{\sigma}_d^2} \sum_{k=1}^{\bar{k}} q_k^2 [2\rho\gamma_k + (1 - \rho)^2] > 1, \quad (3.11)$$

which increases with the duration and size of the volatility components. The relation between feedback and volatility persistence is thus consistent with previous literature (e.g., Poterba and Summers, 1985; Bansal and Yaron, 2004), and extends to a multifrequency setting.

CH attribute their low estimates of volatility feedback to the fact that in GARCH-type processes, the volatility of volatility increases very rapidly (as a fourth power) of the volatility level. By contrast, our model with $\bar{k} = 1$ implies

$$Var_t(\sigma_{t+1}^2) = \bar{\sigma}_d^4 \gamma_1 [Var(M) + (1 - \gamma_1)(M_{1,t} - 1)^2].$$

The volatility of volatility is therefore a non-monotonic, U-shaped function of the volatility level. When $\bar{k} > 1$, it is straightforward to approximate $Var_t(\sigma_{t+1}^2)$ by a sum of U-shaped functions of the multipliers. This weaker relation between the level and conditional variance of volatility suggests that our model may yield larger estimates of the feedback effect.

4. Empirical Results with Fully Informed Investors

We now investigate the performance of the multifrequency equilibrium model on a long sample of excess US equity returns. The test is challenging for several reasons. To begin with, we use daily data spanning a long period of nearly eight decades. Thus, the model must be able to account for regularities that are important at high frequencies, such as thick tails, volatility persistence, skewness, and feedback effects. Additionally, a good model should be flexible enough to account for market conditions that vary widely across different periods in such a long time span, but also avoid overfitting. Our parsimonious multifrequency specification seems like a good candidate for this task.

Explaining the data is even more challenging because we use an equilibrium model rather than a purely econometric specification. Our approach is thus structural in the sense that economic theory places strong restrictions on what would otherwise be degrees of freedom in a purely statistical setup. Finally, our exogenous dividend news process is conditionally Gaussian, has constant mean, and is both conditionally and unconditionally symmetric.²⁰ Thus, endogenous equilibrium effects must play a significant

²⁰Our discussion of symmetry neglects the Itô adjustment term, which is negligible for all practical purposes. In empirical specifications, the contribution of this term to daily returns never exceeds a few basis points, while daily standard deviation is of the order of 1%.

role in explaining daily stock returns.

4.1. Excess Return Data

We estimate the multifrequency equilibrium model on daily excess returns of a US equity index from January 1926 to December 2003. As in CH, the index is constructed by combining the Schwert (1990a) daily index from 1926-1963 with CRSP value-weighted returns from 1963 onwards, and subtracting a daily risk-free rate imputed from 30-day Treasury bills. The entire 1926-2003 period contains 20,765 observations (“Full Sample”). We also report results for the period beginning in 1952, as is common in previous literature (e.g., Campbell, 1991; CH) because it corresponds to a change in interest rate regime with the Fed-Treasury Accord. This sample contains 13,109 observations (“Postwar Sample”).²¹

Figure 1 shows the data, demonstrating the thick tails, low-frequency cycles, and negative skewness that are widely recognized characteristics of aggregate stock returns. To further indicate how conditions change across different periods in the long span of the data, Table 1 shows moments of the excess return series for four evenly spaced subsamples of each sample. These vary substantially, consistent with the findings of Schwert (1989) and Pagan and Schwert (1990). The data thus contain complex high-frequency variations as well as substantial movements at low-frequencies, presenting a significant challenge to a parsimonious, stationary, and tractable equilibrium model of returns.

4.2. Maximum Likelihood Estimation and Volatility Feedback

We use maximum likelihood (ML) to estimate the full-information equilibrium model. This is an achievement in itself. Typically, once one ventures outside the class of GARCH models to a stochastic volatility setting, exact maximum likelihood estimation becomes difficult. We preserve this convenient and asymptotically efficient estimation method because stochastic volatility in our setting is generated by pure regime-switching. Our multifrequency approach allows us to use this technique effectively at daily intervals, whereas extensive previous work with regime-switching over the last fifteen years has been limited to longer horizons. Preserving ML estimation also deserves notice because our returns process is the endogenous outcome of an economic equilibrium. Observed returns are thus a highly complex, nonlinear, and dynamic function of the underlying exogenous state variables.

Table 2 presents our estimation results with the number of volatility components \bar{k}

²¹An earlier version of the paper carried out the same empirical procedures on the CRSP data only, and found results very similar to the postwar results reported here.

varying from 1 to 8 across the columns.²² The first column ($\bar{k} = 1$) thus corresponds to a standard regime-switching model with only two possible volatility states. The two-state case is a mainstay of the existing regime-switching literature, and the focus of a recent volatility feedback study by Kim, Morley, and Nelson (2004).

Examining the value of the likelihood function as \bar{k} increases, we see clearly the benefits of a multifrequency specification. In going from one to two volatility components, the log likelihood increases by over 750 points in the full sample, and over 330 points in the post-war sample. Since this requires adding only one additional parameter (from four to five), the increase in likelihood is large by any standard model selection criterion.²³ Increasing the number of frequencies from two to three raises the log likelihood by an additional 245 points in the full sample, and 167 points in the postwar sample, but does not increase the number of parameters since components are identical except for a tightly-parameterized vector of time-scales. Substantial increases continue, without adding additional parameters, until the likelihood flattens in the range of $\bar{k} = 6$ to $\bar{k} = 8$ volatility components for both samples.

The parameter estimates in Table 2 also deliver sensible intuition. The multiplier m_0 decreases monotonically with the number of components. As \bar{k} grows, each multiplier needs to do less work in explaining aggregate volatility fluctuations, and reducing m_0 is the way the model achieves this. The unconditional volatility of dividends varies with no apparent pattern across \bar{k} , consistent with the fact that long-run volatility is not easy to identify even in large samples. Dividend growth $g_d - r$ is positive and approximately constant at about 5 basis points per day. We explore this result further in later discussions. Finally, the switching probability $\gamma_{\bar{k}}$ of the highest frequency component is fairly stable across specifications, and b , which controls frequency spacing, drops initially and stabilizes at a value of about 2 for specification with $\bar{k} = 6$ and larger. These results imply that the highest frequency volatility shocks have durations of approximately 15 to 30 days. Further, adding volatility components at first tightens intrafrequency spacing when the number of components is small, but eventually serves more to extend the low frequency range of variations.

Table 2 also reports statistics of the first four moments for each specification. Unconditional volatility is variable but in most cases comparable to the data, and excess kurtosis is captured well when \bar{k} is large. However, the model appears to overestimate the equity premium and does not sufficiently capture negative skewness. To fully understand these difficulties, which we later overcome by introducing investor learning, we

²²The computation of the likelihood function uses the exact closed-form expressions of the price:dividend ratios $Q(m^1) \dots Q(m^d)$. See the Appendix for further details.

²³For example, using the Akaike Information Criterion (AIC) or the Bayesian Information Criterion (BIC) the necessary increase in likelihood to justify one additional parameter would be less than five points for either sample size.

will compare our equilibrium model with the seminal work of Campbell and Hentschel (1992).

Before beginning this process, we examine the unconditional volatility feedback

$$\frac{Var(r_{t+1})}{Var(d_{t+1} - d_t)},$$

which is presented in the final line of Table 2. Feedback increases as components are added, and for the best performing models with $\bar{k} \geq 6$, it contributes between 5 – 17% of total variance in the full sample, and 23 – 24% in postwar data. These numbers are 6 to 12 times larger than reported by CH for each period, and we now confirm that this large difference continues to hold in our longer samples.

4.3. Comparison with CH

The Campbell and Hentschel (1992) specification, which we describe in detail in the Appendix, provides a good comparison for our approach. First, CH also use endogenous feedback to generate restrictions on excess stock returns. They assume QGARCH dividend news and a linear pricing rule for volatility, which can be approximately reconciled with an equilibrium setup comparable to ours. Second, CH similarly address feedback effects in daily data over a very long sample, which is ambitious since departures from normality are most pronounced at shorter intervals, and changing economic conditions are more important over longer time spans (Schwert, 1989; Pagan and Schwert, 1990). More recent studies (e.g., Wu, 2001) often focus on lower frequency data over shorter time spans. Third, like CH we restrict dividend news to be symmetric and conditionally normal, which requires endogenous feedback to play a critical role in matching higher moments. By contrast, Wu (2001) allows correlation between dividend news growth and volatility, permitting exogenously skewed dividend news. Fourth, both CH and our model are relatively parsimonious, requiring only seven and five parameters respectively. Finally, both our model and CH permit convenient ML estimation, which further facilitates comparison.

Table 3 reports ML estimation results for the CH model on both samples. Panel A gives parameter estimates, which are comparable to those found in the original CH study. Using the formula given in the Appendix, we can calculate the magnitude of volatility feedback. As in the original study, feedback contributes between 1-2% of unconditional variance, and is thus small relative to our multifrequency equilibrium model.

Panel B compares in-sample fit of the CH model to our multifrequency specification with $\bar{k} = 8$ volatility components. Although our equilibrium model has two fewer parameters, its likelihood is over one hundred points larger. We adjust for the number of parameters by calculating the Bayesian Information Criterion (BIC) statistic for each specification, and assess significance using the Vuong (1989) test and a HAC-adjusted

version proposed in Calvet and Fisher (2004). These show that the difference in likelihood is highly significant in both samples. Moreover, all multifrequency equilibrium models with three or more volatility components have higher likelihood than the CH specification.

This confirms that the full-information multifrequency equilibrium generates large feedback effects and performs well in-sample relative to an important benchmark. We now investigate the multifrequency model in greater detail.

4.4. Conditional Inference

In the full-information framework, investors directly observe the volatility state M_t , but as empiricists we must make inferences based only on excess returns $I_t^0 \equiv \{r_s; s \leq t\}$. The Appendix discusses how to calculate *filtered* beliefs

$$\hat{\Pi}_t^j \equiv \mathbb{P}(M_t = m^j | I_t^0)$$

as well as the *smoothed* probabilities

$$\hat{\Psi}_t^j \equiv \mathbb{P}(M_t = m^j | I_T^0),$$

for $j \in \{1, \dots, 2^{\bar{k}}\}$. The filtered beliefs are most useful for an empiricist wanting to use the model for forecasting, while the smoothed beliefs allow the most informative *ex post* analysis of the data.

Figure 2 displays marginal probabilities for components of both types of beliefs when $\bar{k} = 8$. Specifically, for each $k \in \{1, \dots, 8\}$ the amounts $\hat{\Pi}_t^{M(k)} \equiv \mathbb{P}(M_{k,t} = m_0 | I_t^0)$ and $\hat{\Psi}_t^{M(k)} \equiv \mathbb{P}(M_{k,t} = m_0 | I_T^0)$ give marginal probabilities that volatility component k is in a high state.²⁴

Filtered beliefs on the left side of the figure show sensible patterns. For the lowest frequency $k = 1$, beliefs drop over time from 0.5 to a value close to zero until the 1987 crash, and then jump immediately to almost 1.0. The most likely explanation of the very large price drop is an increase in low-frequency volatility. By contrast, when a smaller but still substantial drop in price of about -8% occurs just after 1955, the lowest frequency $k = 1$ moves little, but the second lowest frequency $k = 2$ jumps upwards substantially. When a similar size price drop occurs in the early 1960's, beliefs about the second volatility component are already high, so it cannot absorb the shock. Filtered beliefs about components $k = 3$ and higher thus increase. In general, the cycles in marginal beliefs have shorter durations as k increases, consistent with intuition. 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

²⁴A complete depiction is not possible since the full belief vectors each contain 256 elements.

up and down rapidly, but rarely reach their boundaries. Conditioning on all returns as in the smoothed beliefs on the right hand of the figure leads to more refined inference. The smoothed marginal probabilities move less frequently but in larger increments, and spend more time near the boundaries of zero and one.

In the first two panels of Figure 3, we use the filtered beliefs to compute the one-step-ahead conditional mean and variance of returns. As implied by our feedback specification, these are positively correlated, showing small peaks in the early 1970's with higher levels in 1987 and around 2000. The asset-pricing literature emphasizes that the market discount rate exhibits small and persistent variations through time. Feedback models focus on cyclical variations in dividend news volatility as a possible source of these fluctuations.²⁵ While our multifrequency specification of volatility is appealing because it permits multiple sources of volatility fluctuations in accord with economic intuition, one might worry that this would lead to a conditional mean that is “too variable” or “too jumpy.” Figure 3 shows that this is in fact not the case. The conditional discount rate moves slowly because it is dominated by the most persistent volatility component, as predicted by the loglinear solution in Section 3.

The last panel of Figure 3 shows the *conditional feedback* $Var_t(r_{t+1}) - Var_t(d_{t+1})$, which moves positively with dividend volatility, but not to the same degree as in CH. In fact, conditional feedback is relatively stable, and the ratio of its maximum to minimum value is less than two. Regressing the log of conditional feedback on the log of conditional news volatility $Var_t(d_{t+1})$, we find a coefficient of 0.223 with a standard error of 0.001 and an R^2 of 73.6%. This contrasts with CH where the regression coefficient would be equal to two and the R^2 equal to one, and confirms that feedback in our model is a dampened rather than magnified version of conditional variance.

4.5. Return Decomposition

We now decompose U.S. equity returns into a conditional expectation, feedback innovation, and dividend news. This provides a convenient *ex post* quantification of the impacts of volatility feedback in the sample.

The fully informed investor observes in period $t + 1$ the excess return r_{t+1} given by (3.3). At time $t + 1$ or later, the investor can thus implement the decomposition

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

This separates the realized return into: 1) its expected value at time t ; 2) the innovation due to the volatility feedback; and 3) the multifrequency dividend news.

²⁵Other explanations include investor heterogeneity (e.g., Constantinides and Duffie, 1996; Dumas, 1989), habit-formation (Abel, 1990; Constantinides, 1990; Campbell and Cochrane, 1999), prospect theory (Benartzi and Thaler, 1995), or irrational expectations (Barberis, Shleifer and Vishny, 1998). See Campbell (2003) for a recent review of this literature.

Even after the entire sample is observed, the empiricist has a smaller information set $I_T^0 \subset I_T$ than the investor, and thus derives an analogous but less precise decomposition. Specifically, the relation $r_{t+1} = \mathbb{E}(r_{t+1}|I_T^0)$ implies

$$r_{t+1} = \mathbb{E}_{\hat{\Psi}(t)} r_{t+1} + \left(\mathbb{E}_{\hat{\Psi}(t+1)} - \mathbb{E}_{\hat{\Psi}(t)} \right) r_{t+1} + \hat{e}_{d,t+1}, \quad (4.1)$$

where

$$\hat{e}_{d,t+1} \equiv \mathbb{E}[\sigma_d(M_{t+1})\varepsilon_{d,t+1}|I_T^0] \quad (4.2)$$

is the *ex post* estimate of realized dividend news. By the law of iterated expectations, $\hat{e}_{d,t+1}$ has mean zero.

We implement the *ex post* decomposition in Figure 4. The top panel (4A) illustrates the excess return series $\{r_t\}$, and the remaining panels show consecutively the three smoothed terms of (4.1): conditional return, volatility feedback, and dividend news. We examine these successively.

The smoothed conditional return in Panel 4B shows small persistent variations, very much like the *ex ante* conditional return in Figure 3. Thus, allowing for more precise inference about the time t volatility state does not greatly influence expected returns.

By contrast, the smoothed feedback in Panel 4C differs sharply from the modest-sized and fairly stable *ex ante* feedback in Figure 3. In particular, smoothed feedback appears in strong, intermittent bursts. On most days it is small, but its occurrences coincide with the most substantial variations in the series, and on these days it contributes a large portion of realized returns. These features are consistent with the intuition of the model that low-frequency volatility changes are infrequent but have a large price impact when they occur.

To understand the large difference between *ex ante* and *ex post* feedback estimates, recall that changes in the individual volatility states are unpredictable. Thus, even given precise information about the state M_t , the empiricist has limited ability to 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.

Finally, in Panel 4D, the residual $\hat{e}_{d,t+1}$ is the filtered version of a symmetric MSM process. We calculate its sample moments, and find a variance of 0.693, skewness coefficient -0.121 , and kurtosis 8.39.²⁶ Relative to the actual return data, the residual variance is 13% smaller, skewness is 89% smaller, and leptokurtosis is 78% smaller. The

²⁶These are not necessarily unbiased estimates. For example, applying Jensen's inequality to (4.2) shows that the variance of $\hat{e}_{d,t+1}$ is a lower bound for the variance of dividend news.

filtered residual thus has a much lower skewness and kurtosis than the equity returns. These findings suggest that endogenous volatility feedback plays an important role in explaining the higher moments of returns in our sample.

4.6. Overidentifying Tests of the Full-Information Equilibrium

The previous section shows that, given observed stock returns, feedback plays an important role in matching higher sample moments. A different, but useful question is, how likely is the model to generate moments similar to the data? Since we have already used the exact likelihood to estimate the model, we can design overidentifying tests around this question.

For the sake of comparison, we conduct the tests for both the multifrequency and CH specifications. For each process, we simulate 1,000 paths of the same length as the data and calculate statistics of the first four moments for each path. We also report the fraction of paths for which the simulated statistic exceeds the corresponding empirical moment. Note that this is a challenging test, since both models rely on endogenous economic mechanisms to fit higher moments of the data.

Table 4 shows the results. Both models accurately capture the second moment, which falls well within the five percent simulated confidence bands for each sample period. Both models also overestimate the mean equity premium. Although CH is somewhat closer to the data, the mean return is significant at the five percent level in all cases. Likewise, the models have trouble capturing the negative skewness of the data, but CH is somewhat closer and the data falls within its five percent confidence band in the full sample. Finally, the multifrequency model captures well the high kurtosis of stock returns, while CH does not.

To understand these results, we first note the link between difficulty matching the first and third moments that has been established in the literature. FSS and CH discuss that failing to sufficiently capture negative skewness in a conditionally heteroskedastic sample tends to produce overestimates of the mean.²⁷ Accurately capturing the skewness of the data thus appears important.

In generating endogenous skewness, an important difference between CH and the multifrequency equilibrium arises from predictive asymmetry in the dividend news process. Predictive asymmetry is the property that negative innovations generate higher future volatility than positive innovations of same magnitude. When exogenous dividend

²⁷Consider a very negative observed return. If the innovation process is negatively skewed, then the observation may be accounted for without inferring a high volatility. On the other hand, if negative skewness is underestimated, then one must infer a correspondingly larger volatility. In estimating the mean of a heteroskedastic sample, less weight is given to observations with high volatility. Thus, when negative skewness is underestimated, a calculation of the population mean gives less weight to negative observations, more weight to positive observations, and the inferred population mean tends to be larger.

news has this feature, volatility feedback is immediately asymmetric, giving stronger negative skewness in returns. Early drafts of CH (e.g., Campbell and Hentschel, 1991) show that standard GARCH processes provide only a modest level of endogenous skewness. Incorporating predictive asymmetry in dividend news through the QGARCH specification helps to give stronger negative skewness, explaining the results in Table 4.

Our MSM specification for dividend growth has no predictive asymmetry. One possibility would be to modify the dividend growth process to accommodate this feature, following the path of CH. While reasonable, this would be more an econometric solution than the outcome of an endogenous economic mechanism. In the next Section, we leave our symmetric dividend process unchanged, and instead show that learning about stochastic volatility can be a powerful and economically appealing method of generating endogenous skewness.²⁸

5. Learning About Volatility and Endogenous Skewness

This section shows that learning about stochastic volatility provides a substantial source of endogenous skewness not previously identified in the literature. We assume that Bayesian investors receive imperfect signals about the state of the economy, which is a reasonable reduced-form if fundamental research is costly. The quality of the signal controls a tradeoff between endogenous skewness and kurtosis: as information quality deteriorates, kurtosis falls and returns become more negatively skewed. Although ML estimation is no longer feasible, we find through simulation that (i) the size of the volatility feedback effect is not highly sensitive to the learning environment, and (ii) intermediate information levels best capture higher moment of stock returns.

5.1. Investor Information and Stock Returns

Investors observe every period consumption, dividends, and noisy observations of the volatility components:

$$\delta_t = M_t + \sigma_\delta z_t, \tag{5.1}$$

where $\sigma_\delta \geq 0$ is a scalar, and $z_t \in \mathbb{R}^{\bar{k}}$ is an IID vector of independent standard normals. This specification nests the full information case ($\sigma_\delta = 0$). The information set $I_t = \{(C_{t'}, D_{t'}, \delta_{t'}); t' \leq t\}$ generates a conditional probability distribution Π_t over the state space $\{m^1, \dots, m^d\}$. We show in the Appendix that beliefs can be computed recursively.

The price:dividend ratio is a function of the investor beliefs: $P_t/D_t = Q(\Pi_t)$. The stochastic discount factor depends only on consumption and is thus the same as in the

²⁸Note that it would be difficult to incorporate learning in the CH model since volatility is a deterministic function of past observations in GARCH-type settings. Learning is, however, a natural path to pursue in our multifrequency latent state variable environment.

full information economy. In equilibrium, the stock price satisfies

$$\delta' \mathbb{E} \left[\left(\frac{C_{t+1}}{C_t} \right)^{-\alpha} \frac{D_{t+1}}{D_t} \frac{1 + Q(\Pi_{t+1})}{Q(\Pi_t)} \middle| I_t \right] = 1,$$

and by forward iteration

$$Q(\Pi_t) = \mathbb{E} \left[\sum_{i=1}^{\infty} \delta^i \left(\frac{C_{t+i}}{C_t} \right)^{-\alpha} \frac{D_{t+i}}{D_t} \middle| I_t \right]. \quad (5.2)$$

The P:D ratio is the conditional expectation of exogenous variables driven by the first-order Markov state M_t . We infer that P:D is linear in the current belief²⁹

$$Q(\Pi_t) = \mathbb{E} [Q(M_t) | I_t] = \sum_{j=1}^d Q(m^j) \Pi_t^j, \quad (5.3)$$

where $Q(m^j)$ is the price:dividend ratio computed under full information. The setup is thus highly tractable, as it implies that prices are a (linear) belief-weighted average of state-prices from the full information model.

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

$$r_{t+1} = \ln \frac{1 + Q(\Pi_{t+1})}{Q(\Pi_t)} + \mu_d(M_{t+1}) - r_f - \frac{\sigma_d^2(M_{t+1})}{2} + \sigma_d(M_{t+1}) \varepsilon_{d,t+1}. \quad (5.4)$$

When a new state occurs, investor learning implies that the market adjusts more gradually and thus generates less extreme returns than in the full information economy. Simulating the return process with learning is now straightforward, as discussed in the Appendix.

The equilibrium impact of signal variability σ_δ is conveniently analyzed from (5.3) for fixed values of the other parameters $(m_0, \gamma_{\bar{k}}, b, \bar{\sigma}_d, g_d - r_f)$. The P:D ratio is the filtered version of its full information counterpart, which implies equality of the means: $\mathbb{E}Q(\Pi_t) = \mathbb{E}Q(M_t)$. This result suggests that information quality has essentially no effect on the equity premium. The variance satisfies the orthogonality condition: $Var[Q(M_t)] = Var[Q(\Pi_t)] + \mathbb{E}\{[Q(\Pi_t) - Q(M_t)]^2\}$. Note that this equation is the analogue in our setup of the variance bounds considered by Leroy and Porter (1981) and Shiller (1981). In our framework, we expect the difference in variances to be small: The loglinear approximation of Section 3.2 shows that the variance of P:D is dominated by changes in the most persistent components. Since learning about these changes is a rare and transitory phenomenon, the difference $Q(\Pi_t) - Q(M_t)$ is likely to be modest most of the time. This

²⁹In a representative agent economy with Epstein-Zin-Weil utility, the P:D ratio is linear in beliefs if: (1) dividend growth is driven by a Markov state; and (2) consumption growth is a separate IID process.

suggests that the variances of P:D and returns are relatively insensitive to information quality, and we confirm this logic numerically in the next Section.

The linearity property (5.3) distinguishes our work from earlier research in which the stock price is a convex function of beliefs (e.g. Brennan and Xia, 2001; David, 1997; Lettau, Ludvigson and Wachter, 2004; Veronesi, 1999, 2000). Convexity induces skewness because price is more sensitive to bad news in good times than it is to good news in bad times. This “uncertainty” channel originates from signals that are informative about both future dividend news and future marginal rates of substitution. We will show that even in the absence of such a dual role, learning about stochastic volatility can be a powerful source of endogenous skewness.

The equilibrium effect of information quality is investigated by Veronesi (2000), who examines learning about mean dividend growth. His setup, like ours, involves a stationary equilibrium with discrete latent states and rational Bayesian investors. Our agenda of investigating learning about volatility is clearly complementary to his.³⁰ Other differences include: 1) Veronesi uses an analytically convenient continuous-time setup with a single persistence level, while we develop a tractable multifrequency specification in discrete-time, and 2) Veronesi uses a Lucas tree economy with power utility, while we separate dividends from consumption to obtain linear pricing with Epstein-Zin preferences. The tree structure in Veronesi (2000) implies that many of his results have analogues in the volatility feedback literature. For instance, an increase in the expected growth rate, like a decrease in volatility, can lead to a decline in prices. Our focus on cash flow news allows us to abstract away from these issues and develop a multifrequency learning model that can be applied to daily stock returns.

5.2. Learning Model Results

Despite the simplicity of the pricing and updating rules, econometric inference is computationally expensive in our imperfect information equilibrium. To see this, note that the state of a learning economy consists of the volatility vector M_{t+1} and the investor belief Π_{t+1} . Since the econometrician observes only excess returns, evaluating the likelihood of the data would require integrating over the conditional distribution of the state (Π_t, M_t) . When $\bar{k} = 8$, this would entail estimating a distribution defined on $\mathbb{R}^{256} \times \{m^1, \dots, m^{256}\}$ and the curse of dimensionality would set in.

We instead use a simulation-based approach to investigate the learning model. We focus on the specification with $\bar{k} = 8$ frequencies, and a base set of parameters

³⁰In concluding his study, Veronesi anticipates an extension along the lines we pursue: “There are other types of information that are certainly relevant and that are also worth investigating. These may include information about future volatilities [...] for example. The effect of information quality on these variables may have different implications on stock returns than the one discussed here.” We verify this prediction.

$(m_0, \sigma_d, g_d - r, \gamma_{\bar{k}}, b)$ that appear reasonable given the empirical results in Section 4. In particular, we set $m_0 = 1.34$ and $\sigma_d = 0.78$ close to their values estimated from the post-war data in the full information model. We reduce $g_d - r$ to the value 0.015, or about 3.75% annually. This approximately matches the empirically observed growth rate of dividends, and also gives simulated returns that better fit the mean equity premium. Finally, we set the frequency parameters to $\gamma_{\bar{k}} = 0.05$ and $b = 2.5$. These are comparable to the values from the full information estimation, but permit somewhat more contribution from low frequencies. Intuitively, when investors gradually infer changes to the volatility state, the model can accommodate lower frequency shocks without implying implausibly large price changes. This choice of specification permits a lowest frequency shock of about once every forty years, which is about half the length of the sample.

To evaluate the impact of information quality, we consider a set of signal volatilities $\sigma_\delta \in \{0, .1, \dots, 1, 1.25, \dots, 2, 3, 4, 5, 10, 15, 20\}$. For each value, we simulate a single long sample of excess returns, and calculate the first four moments of returns as well as the feedback, using the same set of random draws. We discuss all results, and report a subset in Table 5. The mean return is equal to 2.54 basis points per day across all simulations, which is close to the mean equity premium in both our sample periods. The fact that mean returns appear approximately constant with the learning environment is not surprising, since the average price:dividend ratio is independent of the signal precision. The standard deviation is likewise nearly invariant to information quality, and takes a value of about 0.82 for each simulation, close to the empirical value in the post-war period. Because the second moment is not sensitive to signal precision, feedback is also nearly constant across the different simulations, and takes a value of about 12.5%. Thus, the degree of volatility feedback is robust across different learning environments.

We do, however, find large and systematic differences in the degree of skewness and kurtosis across learning environments. Skewness is close to zero at about -0.02 when $\sigma_\delta = 0$, falling to -0.26 when $\sigma_\delta = 0.5$, to -0.71 when $\sigma_\delta = 2$, and to -1.14 when $\sigma_\delta = 20$. Returns thus become more negatively skewed as investor information becomes less precise. Kurtosis takes its highest value of about 47 when investor information is perfect. With a value of $\sigma_\delta = 0.5$ kurtosis drops to 13.8 and when $\sigma_\delta = 2$ kurtosis falls to 8.23. We thus infer a tradeoff between skewness and kurtosis. With perfect information kurtosis is large but skewness is close to zero. As the quality of investor information deteriorates, returns become more negatively skewed and kurtosis falls as well. Values in the range $\sigma_\delta \in [0.4, 1.0]$ seem most consistent with the high kurtosis and substantial negative skewness observed in the data. The tradeoff between skewness and kurtosis across all values of σ_δ is depicted in Figure 5.

To understand these results, consider the role played by dividend growth in the investor updating process. When information is perfect, dividend growth plays no role in

determining investor beliefs about the volatility state. Regardless of whether volatility state variables increase, decrease, or stay the same, investors find out immediately and fully incorporate into price the impact of any changes. The speed of learning is independent of the direction of the volatility change, and returns are approximately symmetric. Kurtosis is high and skewness close to zero.

At the other extreme, when σ_δ is arbitrarily large the corresponding signals are not useful, and investors rely on dividend news to infer the latent state. If volatility increases, investors may get a single extreme observation that is implausible under their existing beliefs. In this case beliefs quickly revise upward. On the other hand, a volatility decrease (good news) can only be revealed slowly. This is because investors learn about low volatility by observing dividend growth close to its mean, but this is a relatively likely outcome regardless of the volatility level. Thus, bad news about increased volatility can be incorporated into price quickly, while good news about low volatility trickles out slowly. This asymmetry explains why skewness increases and kurtosis falls as information quality about the volatility state deteriorates.

To further illustrate the effect of information quality, Figure 6 displays four simulations with length $T = 20,000$ of the learning economy with different signal precisions. Consecutively from top to bottom, $\sigma_\delta = 0$ corresponds to full information, $\sigma_\delta = 0.6$ and $\sigma_\delta = 1.0$ give two reasonable intermediate values, and $\sigma_\delta = 20$ corresponds to nearly uninformative signals. All simulations use identical sets of random draws to facilitate comparison, and we clearly see the anticipated effects. With perfect information, large and symmetric feedback gives substantial outliers of both signs. As information quality decreases, gradual learning causes feedback to be spread out across multiple days, and fewer extreme returns occur. The attenuation is noticeably stronger for positive returns, and skewness thus becomes more pronounced with σ_δ . When $\sigma_\delta = 20$, this effect is so extreme that no large positive returns occur in the simulation. The intermediate cases where $\sigma_\delta = 0.6$ and $\sigma_\delta = 1.0$ appear most consistent with daily stock returns.

6. Conclusion

This paper develops a tractable asset-pricing framework for economies with multifrequency shocks to dividend news. We focus on a news specification with constant mean, multifrequency stochastic volatility, and a conditionally Gaussian noise. The five-parameter equilibrium model accounts for skewness, thick tails, time-varying volatility and feedback in over eighty years of daily stock returns.

Two endogenous mechanisms play a critical role in matching equity data. First, volatility feedback contributes between 12 and 24% of return variance, or 6 to 12 times the amount in previous literature (e.g., Campbell and Hentschel, 1992). Feedback from persistent components helps to capture extreme returns, while higher-frequency varia-

tions match day-to-day volatility movements. Second, investor learning generates substantial endogenous skewness. Building on Veronesi (2000), we consider investor signals about the volatility state, and show that information quality creates a tradeoff between skewness and kurtosis. Intermediate information environments best match the data.

The paper illustrates that our multifrequency approach helps to connect low-frequency macro-finance and learning literature with higher frequency financial econometrics. Convergence of these areas follows from bringing multifrequency shocks into pure regime-switching economies, which traditionally offer three major advantages: 1) asset pricing is straightforward in a Markov chain setup; 2) the econometrics of regime-switching, based on a simple filtering theory, is well-understood; and 3) learning is easily incorporated by using similar filtering techniques. The multifrequency approach thus expands the practical range of equilibrium regime-switching economies from a few states to several hundred, and from lower frequencies to daily returns.

We anticipate that the multifrequency approach can be extended in a number of useful directions. First, our Epstein-Zin setup is tractable when the latent Markov regimes affect drifts as well as volatilities. This invites future investigations on the equilibrium effect of learning about drifts in a multifrequency environment. Intuition suggests that learning about the drift should strengthen both our empirical findings and the results produced by Veronesi (1999, 2000), Brennan and Xia (2001) and others in a unifrequency environment. Second, recent econometric developments permit the effective and parsimonious specification of multifrequency comovement across assets or factors (Calvet, Fisher, and Thompson, 2004). This advance permits for instance the joint modelling of consumption and dividends, as in Bansal and Yaron (2004), or more generally paves the way for novel specifications of multifrequency covariation in stochastic discount factors and asset payoffs. Extensions along these lines are left as open areas for future research.

7. Appendix A. Economies with Regime-Switching Drift and Volatility

7.1. Stochastic Discount Factor

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 tree has price $P_{c,t} = p_c C_t$, and yields the return $1 + R_{c,t+1} = (1 + 1/p_c)C_{t+1}/C_t$. 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. Numerical Pricing

The process M_t is a Markov-chain with 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 . Consider the row vector $\iota = (1, \dots, 1) \in \mathbb{R}^d$, the equilibrium column vector

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

and the matrix $B = (b_{ij})_{1 \leq i, j \leq d}$ with components $b_{ij} = e^{\mu_d(m^j) - r_f - \alpha \rho_{c,d} \sigma_c \sigma_d(m^j)} a_{i,j}$. The pricing condition (2.4) can be rewritten as $q = B(\iota' + q)$, or equivalently

$$q = (I - B)^{-1} B \iota'. \quad (7.1)$$

In empirical work, we use this expression to compute numerically the equilibrium price:dividend ratio corresponding to a given set of parameters.

7.3. Bayesian Updating and Closed-Form Likelihood

We derive the filtered beliefs $(\hat{\Pi}_t^j)_{1 \leq j \leq d}$ and the corresponding likelihood. The derivation follows Hamilton (1989, 1990), but due to feedback the conditional density of observables at $t + 1$ depends on both t and $t + 1$ latent states.

Proposition 1. *The econometrician's conditional probabilities are computed recursively using Bayes' rule:*

$$\hat{\Pi}_{t+1} = \frac{\hat{\Pi}_t [A * F(r_{t+1})]}{\hat{\Pi}_t [A * F(r_{t+1})] \iota'}, \quad (7.2)$$

where $*$ denotes element-by-element multiplication, and $F(r)$ is the matrix with elements $F_{i,j}(r) \equiv f_{r_{t+1}}(r | M_t = m^i, M_{t+1} = m^j)$. The log-likelihood of the return process satisfies

$$\ln L(r_1, \dots, r_T) = \sum_{t=1}^T \ln \left\{ \hat{\Pi}_{t-1} [A * F(r_t)] \iota' \right\}. \quad (7.3)$$

Proof. The conditional probability $\hat{\Pi}_{t+1}^j = \mathbb{P}(M_{t+1} = m^j | I_t^0, r_{t+1})$ satisfies

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

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

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

The updated probability can now be written as

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

which implies (7.2).

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

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

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

8. Appendix B: Loglinear Approximation

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

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

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

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

We infer that equations (3.7) – (3.6) hold.

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 $(1, 1, \dots, 1)$:

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

Combining this result with (3.5), we infer

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

and conclude that (3.8) holds.

We then easily derive the conditional moments of the loglinear return process. The conditional return is given by $\mathbb{E}_t r_{t+1} \approx \alpha\sigma_{c,d} + \sum_{k=1}^{\bar{k}} q_k [1 - \rho(1 - \gamma_k)](M_{k,t} - 1)$, which is equivalent to (3.9). The return innovation $r_{t+1} - \mathbb{E}_t r_{t+1}$ is then easily computed.

By (3.8), the unconditional variance of returns satisfies $Var(r_{t+1}) \approx \bar{\sigma}_d^2 + \sum_{k=1}^{\bar{k}} q_k^2 Var[(M_{k,t} - 1) - \rho(M_{k,t+1} - 1)]$. We note that $\mathbb{E}[(M_{k,t+1} - 1)(M_{k,t} - 1)] = (1 - \gamma_k)Var(M)$ and conclude that (3.11) holds.

9. Appendix C: Full-Information Volatility Feedback Economy

9.1. Smoothed Beliefs

Proposition 2. *The econometrician's smoothed belief satisfies the backward recursion*

$$\hat{\Psi}_t^i = \hat{\Pi}_t^i \sum_{j=1}^d a_{ij} \frac{\hat{\Psi}_{t+1}^j}{\hat{\Pi}_{t+1}^j} \left[\frac{f^{ij}(r_{t+1})}{f_{r_{t+1}}(r_{t+1}|I_t^0)} \right], \quad (9.1)$$

$i \in \{1, \dots, d\}$, and the final condition $\hat{\Psi}_T = \hat{\Pi}_T$.

Proof. Since M_t is first-order Markov, we know that $\hat{\Psi}_t^i = \sum_{j=1}^d \mathbb{P}(M_{t+1} = m^j | I_T^0) \mathbb{P}(M_t = m^i | M_{t+1} = m^j, I_T^0)$ can be written as

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

By Bayes' rule, $\mathbb{P}(M_t = m^i | M_{t+1} = m^j, I_{t+1}^0) = \mathbb{P}(M_t = m^i, M_{t+1} = m^j | I_{t+1}^0) / \hat{\Pi}_{t+1}^j$, and thus

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

The conditional probability $\mathbb{P}(M_t = m^i, M_{t+1} = m^j | I_{t+1}^0)$ is by Bayes' rule equal to

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

We infer that

$$\mathbb{P}(M_t = m^i, M_{t+1} = m^j | I_{t+1}^0) = \frac{\hat{\Pi}_t^i a_{ij} f^{ij}(r_{t+1})}{f_{r_{t+1}}(r_{t+1} | I_t^0)}, \quad (9.2)$$

and conclude that (9.1) holds. \blacksquare

This smoothing rule slightly differs from the one derived by Kim (1994) for traditional Hamilton models, in which the signal observed by the econometrician 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

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

The smoothed probability thus satisfies

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

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.

9.2. Ex Post Decomposition

We condition the return equation (3.3) with respect to the econometrician's information set I_T^0 :

$$r_{t+1} = g_d - r_f + \mathbb{E} \left[\ln \frac{1 + Q(M_{t+1})}{Q(M_t)} - \frac{\sigma_d(M_{t+1})^2}{2} \middle| I_T^0 \right] + \hat{\epsilon}_{d,t+1}.$$

The definition of smoothed beliefs implies

$$r_{t+1} = g_d - r_f + \mathbb{E}_{\hat{\Psi}(t+1)} (\ln[1 + Q(M_{t+1})] - \sigma_d(M_{t+1})^2/2) - \mathbb{E}_{\hat{\Psi}(t)} \ln Q(M_t) + \hat{\epsilon}_{d,t+1}$$

Since $\mathbb{E}_{\hat{\Psi}(t)} r_t = g_d - r_f + \mathbb{E}_{\hat{\Psi}(t)} (\ln[1 + Q(M_{t+1})] - \sigma_d(M_{t+1})^2/2 - \ln Q(M_t))$, we conclude that (4.1) holds.

10. Appendix D: The Campbell-Hentschel Model

The CH specification is based on a dividend news that follows a QGARCH(1,2) process (Engle, 1990; Sentana, 1995). Excess returns follow

$$r_{t+1} = \mu + \gamma \sigma_t^2 + (1 + 2\lambda b) \eta_{d,t+1} - \lambda (\eta_{d,t+1}^2 - \sigma_t^2), \quad (10.1)$$

where the dividend news $\eta_{d,t+1}$ are $\mathcal{N}(0, \sigma_t^2)$, and

$$\begin{aligned} \sigma_t^2 &= \varpi + \alpha_1 (\eta_{d,t} - b)^2 + \alpha_2 (\eta_{d,t-1} - b)^2 + \beta \sigma_{t-1}^2, \\ \lambda &= \frac{\gamma \rho (\alpha_1 + \rho \alpha_2)}{1 - \rho (\alpha_1 + \rho \alpha_2 + \beta)}. \end{aligned}$$

The parameter ρ is calibrated to the empirical P:D ratio, and the seven parameters $(\mu, \gamma, \varpi, \alpha_1, \alpha_2, b, \beta)$ are estimated by maximum likelihood.

The conditional return $\mathbb{E}_t r_{t+1} = \mu + \gamma \sigma_t^2$ increases in conditional volatility and γ , which is related to risk aversion. Feedback appears itself in (10.1) through a quadratic term in dividend news, $\lambda(2b\eta_{d,t+1} + \sigma_t^2 - \eta_{d,t+1}^2)$. After an extreme innovation $\eta_{d,t+1}$, the investor knows that volatility will increase, and price drops. The location parameter b differentiates QGARCH from traditional GARCH. When $b > 0$, negative dividend news $\eta_{d,t}$ have a higher impact on volatility than positive news of the same size. Feedback is calculated by $(1 + 2\lambda b)^2 + 2\lambda^2 \sigma_t^2$.

11. Appendix E: Learning Economies

Consider the volatility state and investor belief (M_t, Π_t) at the end of period t . The state of the economy in the following period is computed in three steps:

1. *Latent state of nature.* We draw the volatility state M_{t+1} given M_t . We also sample $\bar{k} + 2$ independent standard normals $(z_{1,t+1}; \dots; z_{\bar{k},t+1}; \varepsilon_{d,t+1}; \eta_{c,t+1})$. The Gaussian consumption noise is $\varepsilon_{c,t+1} = \rho_{c,d}\varepsilon_{d,t+1} + \sqrt{1 - \rho_{c,d}^2}\eta_{c,t+1}$. We then compute the consumption, dividend and signal in period $t + 1$. The empirical evidence mentioned in Section 2 of the paper suggests to choose a relatively low value for $\rho_{c,d}$. We use $\rho_{c,d} = 0.20$ in our simulations.
2. *Investor belief.* The investor observes $(\delta_{t+1}, c_{t+1} - c_t, d_{t+1} - d_t)$. She then computes her new probability distribution Π_{t+1} over volatility states with Bayes' rule :

$$\Pi_{t+1}^j \propto f(\delta_{t+1}, c_{t+1} - c_t, d_{t+1} - d_t | M_{t+1} = m^j) \sum_{i=1}^d a_{i,j} \Pi_t^i, \quad (11.1)$$

for all $j \in \{1, \dots, d\}$.

3. *Stock Return.* We can then compute the corresponding excess return using (5.4).

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 (1990), Asset Prices Under Habit Formation and Catching Up with the Joneses, *American Economic Review* 80, 38-42.
- [3] Abel, Andrew (1999), Risk Premia and Term Premia in General Equilibrium, *Journal of Monetary Economics* 43, 3-33.
- [4] Abel, Andrew (2003), The Effects of a Baby Boom on Stock Prices and Capital Accumulation in the Presence of Social Security, *Econometrica* 71, 551-578.
- [5] Andersen, Torben, Tim Bollerslev, and Francis Diebold (2003), Some Like it Smooth, and Some Like it Rough: Untangling Continuous and Jump Components in Measuring, Modeling, and Forecasting Asset Return Volatility, Working paper, University of Pennsylvania.
- [6] Andersen, Torben, Tim Bollerslev, Francis Diebold and Clara Vega (2004), Real-Time Price Discovery in Stock, Bond and Foreign Exchange Markets, Working paper, University of Pennsylvania.
- [7] Antweiler, Werner, and Murray Frank (2004), Is All That Talk Just Noise? The Information Content of Internet Stock Message Boards, *Journal of Finance* 59, 1259-1295.
- [8] Attanasio, Orazio, and Guglielmo Weber (1993), Consumption Growth, The Interest Rate, and Aggregation, *Review of Economic Studies*. 60, 631-649.
- [9] Bansal, Ravi, and Amir Yaron (2004), Risks for the Long Run: A Potential Resolution of Asset Pricing Puzzles, *Journal of Finance* 49, 1481-1509.
- [10] Bansal, Ravi, and Hao Zhou (2002), Term Structure of Interest Rates with Regime Shifts, *Journal of Finance* 57, 1997-2043.
- [11] Barberis, Nicholas, Shleifer, Andrei and Robert Vishny (1998), A Model of Investor Sentiment, *Journal of Financial Economics* 49, 307-343.
- [12] Barndorff-Nielsen, Ole, and Neil Shephard (2003), Realised Power Variation and Stochastic Volatility Models, *Bernoulli* 9, 1109-1111.
- [13] Barsky, Robert (1989), Why Don't the Prices of Stocks and Bonds Move Together?, *American Economic Review* 79, 1132-1145.

- [14] Bekaert, Geert, and Guojun Wu (2000), Asymmetric Volatility and Risk in Equity Markets, *Review of Financial Studies* 13:1, 1-42.
- [15] Benartzi, Shlomo and Richard Thaler (1995), Myopic Risk Aversion and the Equity Premium Puzzle, *Quarterly Journal of Economics* 110, 73-92.
- [16] Brennan, Michael, and Xia, Yihong (2001), Stock Price Volatility and the Equity Premium, *Journal of Monetary Economics* 47, 249-283.
- [17] Calvet, Laurent E., and Adlai J. Fisher (2001), Forecasting Multifractal Volatility, *Journal of Econometrics* 105, 27-58.
- [18] Calvet, Laurent E., and Adlai J. Fisher (2002), Multifractality in Asset Returns: Theory and Evidence, *Review of Economics and Statistics* 84, 381-406.
- [19] 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.
- [20] Calvet, Laurent E., Adlai J. Fisher, and Samuel B. Thompson (2004), Volatility Comovement: A Multifrequency Approach, forthcoming, *Journal of Econometrics*.
- [21] Campbell, John Y. (1991), A Variance Decomposition for Stock Returns, *Economic Journal* 101, 157-179.
- [22] Campbell (1996), Understanding Risk and Return, *Journal of Political Economy* 104, 298-345.
- [23] Campbell, John Y. (2003), Consumption-Based Asset Pricing, in G. Constantinides and M. Harris eds., *Handbook of the Economics of Finance*, Amsterdam: North-Holland.
- [24] 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.
- [25] Campbell, John Y., and Ludger Hentschel (1991), No News is Good News: An Asymmetric Model of Changing Volatility in Stock Returns, NBER Working Paper #3742.
- [26] 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.
- [27] 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.

- [28] 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.
- [29] Cecchetti, Stephen, Pok-Sang Lam and Nelson Mark (1990), Mean Reversion in Equilibrium Asset Prices, *American Economic Review* 80, 398-418.
- [30] Constantinides, George (1990), Habit Formation: A Resolution of the Equity Premium Puzzle, *Journal of Political Economy* 98, 519-543.
- [31] Constantinides, George and Darrell Duffie (1996), Asset Pricing with Heterogeneous Consumers, *Journal of Political Economy* 104, 219-240.
- [32] David, Alexander (1997), Fluctuating Confidence in Stock Markets: Implications for Returns and Volatility, *Journal of Financial and Quantitative Analysis* 32, 427-462.
- [33] Detemple, Jérôme B. (1986), Asset Pricing in a Production Economy with Incomplete Information, *Journal of Finance* 41, 383-391.
- [34] Dothan, Michael U., and David Feldman (1986), Equilibrium Interest Rates and Multi-period Bonds in a Partially Observable Economy, *Journal of Finance* 41, 369-382.
- [35] Dumas, Bernard (1989), Two-Person Dynamic Equilibrium in the Capital Market, *Review of Financial Studies* 2, 157-188.
- [36] Engle, Robert (1990), Discussion: Stock Market Volatility and the Crash of '87, *Review of Financial Studies* 3, 103-106.
- [37] 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.
- [38] French, Kenneth, G. William Schwert, and Robert F. Stambaugh (1987), Expected Stock Returns and Volatility, *Journal of Financial Economics* 19, 3-29.
- [39] Genotte, Gérard (1986), Optimal Portfolio Choice under Incomplete Information, *Journal of Finance* 41, 733-749.
- [40] Greenwood, Jeremy, and Boyan Jovanovic (1999), The Information-Technology Revolution and the Stock Market, *American Economic Review* 89, 116-122.
- [41] Guidolin, Massimo, and Timmermann, Allan (2003), Option Prices Under Bayesian Learning: Implied Volatility Dynamics and Predictive Densities, *Journal of Economic Dynamics and Control* 27, 717-769.

- [42] Hamilton, James (1989), A New Approach to the Economic Analysis of Nonstationary Time Series and the Business Cycle, *Econometrica* 57, 357-84.
- [43] Hamilton, James (1990), Analysis of Time Series Subject to Change in Regime, *Journal of Econometrics* 45, 39-70.
- [44] Jermann, Urban (1998), Asset Pricing in Production Economies, *Journal of Monetary Economics* 4, 257-275.
- [45] Kandel, Shmuel, and Robert F. Stambaugh (1990), Expectations and Volatility of Consumption and Asset Returns, *Review of Financial Studies* 3, 207-232.
- [46] Kim, Chang-Jin (1994), Dynamic Linear Models with Markov-Switching, *Journal of Econometrics* 60, 1-22.
- [47] Kim, Chang-Jin, James Morley and Charles Nelson (2004), Is There a Positive Relationship Between Stock Market Volatility and the Equity Premium?, *Journal of Money, Credit and Banking* 36, 339-360.
- [48] Lettau, Martin, and Sidney Ludvigson (2001), Consumption, Aggregate Wealth, and Expected Stock Returns, *Journal of Finance* 56, 815-849.
- [49] Lettau, Martin, Sidney Ludvigson, and Jessica Wachter (2004), The Declining Equity Premium: What Role Does Macroeconomic Risk Play?, Working Paper, New York University.
- [50] LeRoy, Stephen, and Richard Porter (1981), The Present-Value Relation: Tests Based on Implied Variance Bounds, *Econometrica* 49, 555-574.
- [51] Lochstoer, Lars (2004), Expected Returns and the Business Cycle: Heterogeneous Agents and Heterogeneous Goods, Working Paper, UC Berkeley.
- [52] Lucas, Robert (1978), Asset Prices in an Exchange Economy, *Econometrica* 46, 1429-1445.
- [53] MacKinlay, A. Craig (1997), Event Studies in Economics and Finance, *Journal of Economic Literature* 35, 13-39.
- [54] Merton, Robert C. (1980), On Estimating the Expected Return on the Market: An Explanatory Analysis, *Journal of Financial Economics* 8, 323-361.
- [55] Pagan, Adrian, and G. William Schwert (1990), Alternative Models for Conditional Stock Volatility, *Journal of Econometrics* 45, 267-290.
- [56] Pindyck, Robert (1984), Risk, Inflation, and the Stock Market, *American Economic Review* 74, 334-351.

- [57] Poterba, James, and Lawrence Summers (1985), The Persistence of Volatility and Stock Market Fluctuations, *American Economic Review* 75, 1142-1151.
- [58] Roll, Richard (1984), Orange Juice and the Weather, *American Economic Review* 74, 861-880.
- [59] Schwert, G. William (1989), Why Does Stock Market Volatility Change Over Time?, *Journal of Finance* 44, 1115-1153.
- [60] Schwert, G. William (1990a), Indexes of United States Stock Prices from 1802 to 1987, *Journal of Business* 63, 399-426.
- [61] Schwert, G. William (1990b), Stock Volatility and the Crash of '87, *Review of Financial Studies* 3, 77-102.
- [62] Sentana, Enrique (1995), Quadratic ARCH Models, *Review of Economic Studies* 62, 639-661.
- [63] Shiller, Robert (1981), Do Stock Prices Move Too Much to be Justified by Subsequent Changes in Dividends?, *American Economic Review* 71, 421-436.
- [64] Timmermann, Allan (1993), How Learning in Financial Markets Generates Excess Volatility and Predictability of Stock Returns, *Quarterly Journal of Economics* 108, 1135-1145.
- [65] Timmermann, Allan (1996), Excess Volatility and Predictability of Stock Prices in Autoregressive Dividend Models with Learning, *Review of Economic Studies* 63, 523-557.
- [66] 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.
- [67] Veronesi, Pietro (1999), Stock Market Overreaction to Bad News in Good Times: A Rational Expectations Equilibrium Model, *Review of Financial Studies* 12, 975-1007.
- [68] Veronesi, Pietro (2000), How Does Information Quality Effect Stock Returns?, *Journal of Finance* 55, 807-837.
- [69] Veronesi, Pietro (2004), The Peso Problem Hypothesis and Stock Market Returns, *Journal of Economic Dynamics and Control* 28, 707-725.
- [70] Vissing-Jørgensen, Annette (2002), Limited Asset Market Participation and the Elasticity of Intertemporal Substitution, *Journal of Political Economy* 110, 825-853.
- [71] Vuong, Quang (1989), Likelihood Ratio Tests for Model Selection and Non-Nested Hypotheses, *Econometrica* 57, 307-333.

- [72] Wachter, Jessica (2004), A Consumption-Based Model of the Term-Structure of Interest Rates, *Journal of Financial Economics*, forthcoming.
- [73] Weil, Philippe (1989), The Equity Premium Puzzle and the Risk-Free Rate Puzzle, *Journal of Monetary Economics* 24, 401-421.
- [74] Whitelaw, Robert (2000), Stock Market Risk and Return: An Equilibrium Approach, *Review of Financial Studies* 13, 521-547.
- [75] Womack, Kent (1996), Do Brokerage Analysts' Recommendations Have Investment Value? *Journal of Finance* 51, 137-167.
- [76] Wu, Guojun (2001), The Determinants of Asymmetric Volatility, *Review of Financial Studies* 14, 837-859.
- [77] Yogo, Motohiro (2004a), A Consumption-Based Explanation of Expected Stock Returns, *Journal of Finance*, forthcoming.
- [78] Yogo, Motohiro (2004b), Estimating the Elasticity of Intertemporal Substitution When Instruments Are Weak, *Review of Economics and Statistics* 86, 797-810.

TABLE 1. – EXCESS RETURN MOMENTS

	Sample	By Subperiod			
	Moment	1	2	3	4
<i>A. Full Sample: 1926-2003</i>					
mean	0.022	0.010	0.043	0.013	0.024
standard deviation	1.104	1.643	0.755	0.773	1.003
skewness	-0.30	0.09	0.96	0.07	-1.55
kurtosis	20.54	11.48	11.36	5.79	32.24
<i>B. Postwar: 1952-2003</i>					
mean	0.023	0.044	-0.004	0.021	0.030
standard deviation	0.858	0.666	0.766	0.932	1.023
skewness	-1.05	-0.70	0.13	-2.83	-0.21
kurtosis	26.51	13.34	5.76	59.84	7.18

Notes: This table reports statistics of the first four moments of daily excess returns. The statistics are reported for our 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

	$k = 1$	2	3	4	5	6	7	8
<i>A. Full Sample: 1926-2003</i>								
\hat{m}_0	1.812 (0.006)	1.656 (0.006)	1.584 (0.007)	1.509 (0.008)	1.456 (0.007)	1.412 (0.009)	1.387 (0.007)	1.381 (0.009)
$\hat{\sigma}$	1.477 (0.029)	1.402 (0.025)	1.397 (0.025)	1.412 (0.048)	1.390 (0.037)	1.350 (0.037)	1.218 (0.032)	1.042 (0.059)
$\hat{g}_d - r$	0.044 (0.004)	0.047 (0.005)	0.057 (0.004)	0.055 (0.004)	0.058 (0.004)	0.057 (0.004)	0.056 (0.004)	0.052 (0.004)
$\hat{\gamma}k$	0.030 (0.010)	0.065 (0.008)	0.083 (0.009)	0.065 (0.009)	0.062 (0.017)	0.068 (0.013)	0.037 (0.008)	0.065 (0.015)
\hat{b}	-	8.24 (1.19)	5.12 (0.416)	3.13 (0.254)	2.43 (0.248)	2.12 (0.142)	1.45 (0.074)	1.87 (0.074)
$\ln L$	69146.92	69905.72	70241.25	70363.96	70416.87	70444.32	70463.67	70451.98
$E[r_t]$	0.049	0.052	0.063	0.061	0.064	0.063	0.064	0.062
$Var[r_t]^{1/2}$	1.482	1.416	1.432	1.453	1.437	1.396	1.249	1.129
$Skew[r_t]$	-0.040	-0.054	-0.070	-0.070	-0.072	-0.069	-0.077	-0.072
$Kurt[r_t]$	4.95	6.09	7.76	8.38	8.78	8.64	7.75	18.65
feedback	1.008	1.020	1.051	1.059	1.069	1.070	1.051	1.170
<i>B. Postwar: 1952-2003</i>								
\hat{m}_0	1.711 (0.009)	1.584 (0.008)	1.519 (0.010)	1.440 (0.009)	1.434 (0.011)	1.380 (0.010)	1.372 (0.006)	1.334 (0.005)
$\hat{\sigma}$	1.050 (0.020)	1.005 (0.021)	1.040 (0.022)	0.944 (0.021)	1.022 (0.037)	0.943 (0.028)	0.801 (0.022)	0.779 (0.031)
$\hat{g}_d - r$	0.043 (0.004)	0.051 (0.004)	0.052 (0.004)	0.052 (0.006)	0.048 (0.005)	0.049 (0.006)	0.047 (0.006)	0.048 (0.006)
$\hat{\gamma}k$	0.026 (0.007)	0.057 (0.008)	0.039 (0.005)	0.038 (0.005)	0.060 (0.009)	0.059 (0.013)	0.053 (0.009)	0.077 (0.015)
\hat{b}	-	5.530 (1.06)	2.662 (0.268)	2.114 (0.133)	2.830 (0.199)	2.234 (0.153)	2.094 (0.095)	1.997 (0.091)
$\ln L$	45675.21	46008.95	46175.35	46218.80	46251.37	46263.60	46268.18	46268.07
$E[r_t]$	0.051	0.060	0.060	0.061	0.056	0.058	0.057	0.058
$Var[r_t]^{1/2}$	1.055	1.017	1.060	0.969	1.088	0.999	0.895	0.864
$Skew[r_t]$	-0.032	-0.048	-0.056	-0.055	-0.055	-0.054	-0.054	-0.051
$Kurt[r_t]$	4.49	5.37	6.17	6.26	17.80	12.49	36.62	31.35
feedback	1.010	1.024	1.041	1.056	1.133	1.122	1.247	1.230

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

A. Campbell-Hentschel Model Parameter Estimates							
	$\omega \times 10^7$	α_1	α_2	β	$b \times 10^3$	$\mu \times 10^4$	γ
Full Sample	1.87 (0.78)	0.140 (0.01)	-0.073 (0.01)	0.925 (0.004)	3.05 (0.18)	3.60 (0.53)	0.14 (0.03)
Postwar	0.53 (0.74)	0.145 (0.01)	-0.088 (0.01)	0.934 (0.005)	3.04 (0.20)	3.47 (0.57)	0.47 (0.09)

B. Likelihood Comparison							
		No. of Parameters	$\ln L$	BIC	BIC p -value vs. Multifractal		
					Vuong (1989)	HAC Adj	
Full Sample	Multifractal	5	70451.98	-6.7833			
	QGARCH	7	69920.65	-6.7311	< 0.001	< 0.001	
Postwar	Multifractal	5	46268.07	-7.0554			
	QGARCH	7	46057.26	-7.0218	< 0.001	< 0.001	

Notes: Panel A shows parameter estimates from the CH volatility feedback model. 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 CH specification would be rejected in favor of the multifrequency model.

TABLE 4. – MOMENT COMPARISON

	Standard		
Mean	Deviation	Skewness	Kurtosis
A. Full Sample			
<i>Data</i>			
0.022	1.10	-0.30	20.54
<i>Full-Information Regime-Switching</i>			
0.062	1.129	-0.072	18.65
(0.01)	(0.13)	(0.11)	(3.12)
[0.000]	[0.407]	[0.034]	[0.676]
<i>Campbell-Hentschel QGARCH</i>			
0.038	1.04	-0.19	7.78
(0.01)	(0.09)	(0.23)	(7.45)
[0.021]	[0.823]	[0.114]	[0.973]
B. Postwar			
<i>Data</i>			
0.023	0.858	-1.05	26.51
<i>Full-Information Regime-Switching</i>			
0.058	0.858	-0.060	32.73
(0.01)	(0.09)	(0.30)	(12.80)
[0.000]	[0.490]	[0.001]	[0.359]
<i>Campbell-Hentschel QGARCH</i>			
0.038	0.827	-0.253	5.64
(0.01)	(0.06)	(0.17)	(3.08)
[0.022]	[0.761]	[0.007]	[0.996]

Notes: This table shows a moment-based comparison of the full information regime-switching feedback model against the CH 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. – MOMENTS OF THE LEARNING MODEL

	Signal Standard Deviation σ_δ					
	0	0.2	0.5	1	1.5	2
$E[r_t]$	0.0254	0.0254	0.0254	0.0254	0.0254	0.0254
$Var[r_t]^{1/2}$	0.8183	0.8184	0.8170	0.8178	0.8185	0.8189
$Skew[r_t]$	-0.0250	-0.0742	-0.2573	-0.5063	-0.6301	-0.7116
$Kurt[r_t]$	47.40	29.61	13.78	9.54	8.61	8.23
feedback	1.1263	1.1266	1.1230	1.1250	1.1269	1.1281

Notes: This table shows the effect of learning on different moments of the data. The parameters $m_0 = 1.34$, $\sigma_d = 0.78$, $g_d - r = 0.15$, $\gamma_k = 0.05$, and $b = 2.5$ are held constant across all simulations. Each column of the table reports simulation results for different values of the signal standard deviation σ_d . When $\sigma_\delta = 0$, information is perfect, and as σ_d becomes larger the signal precision weakens. For each specification, we simulate a single long series of $T = 3 * 10^6$ returns using the same set of random draws, and report moments of the simulated data. Mean, variance, and feedback are nearly constant across simulations. Skewness becomes more negative and kurtosis falls as signal precision becomes weaker.

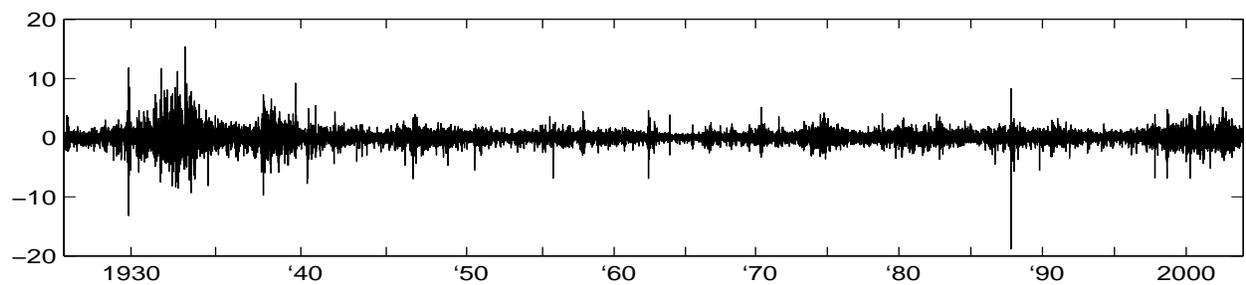


Figure 1: Daily Excess Returns. This figure shows daily excess returns from 1926 to 2003. The market return series splices the Schwert (1990a) data from 1926-1963 with the CRSP value weighted index from 1963-2003. The risk free rate is proxied by the return on 30 day U.S. Treasury bills.

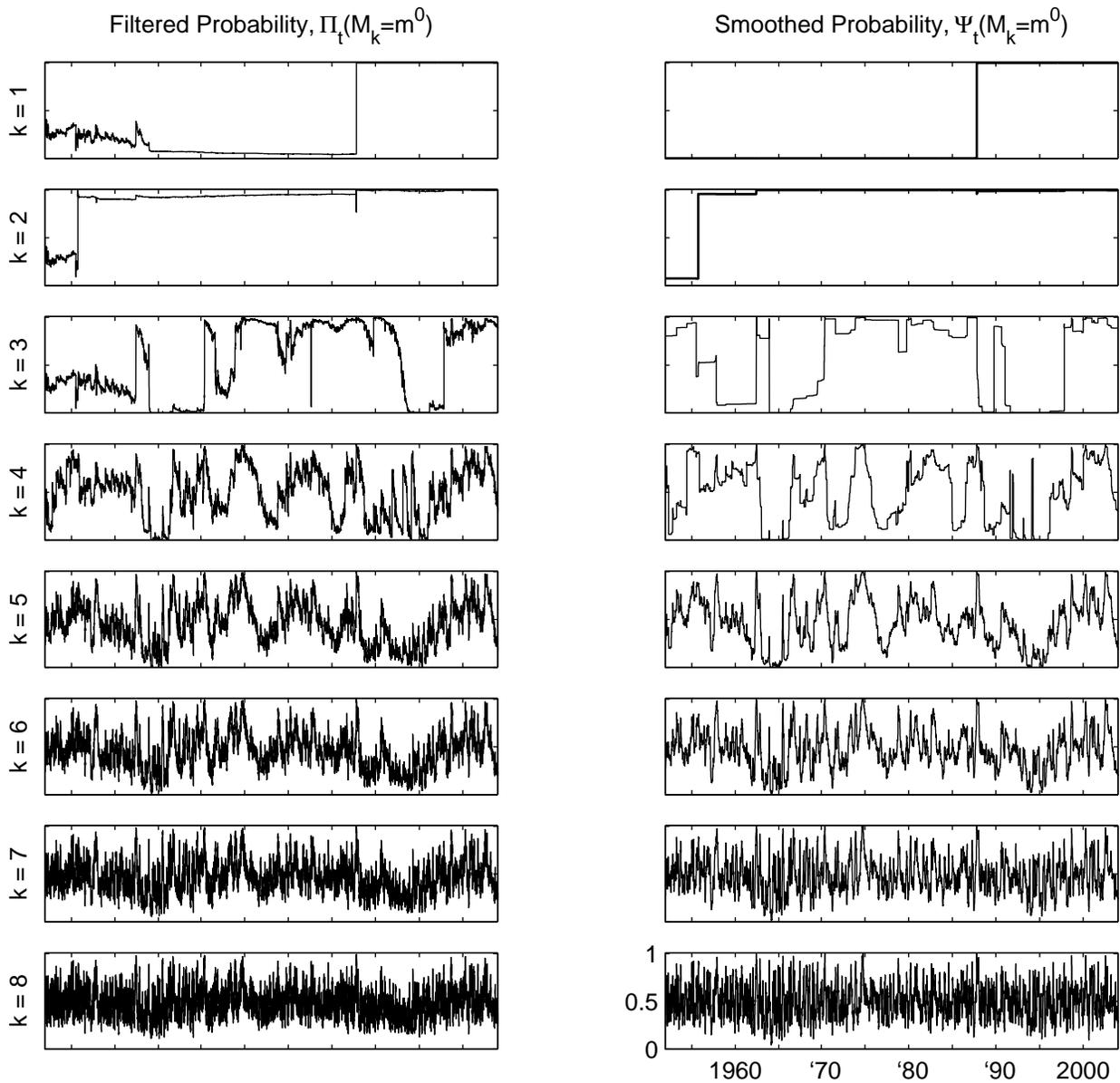


Figure 2: Volatility Component Conditional Beliefs. This figure shows *ex ante* and *ex post* conditional beliefs for the values of each volatility component in the full information regime-switching specification with $\bar{k} = 8$ components. The filtered probabilities Π_t are in the left-hand column, and the smoothed probabilities 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.

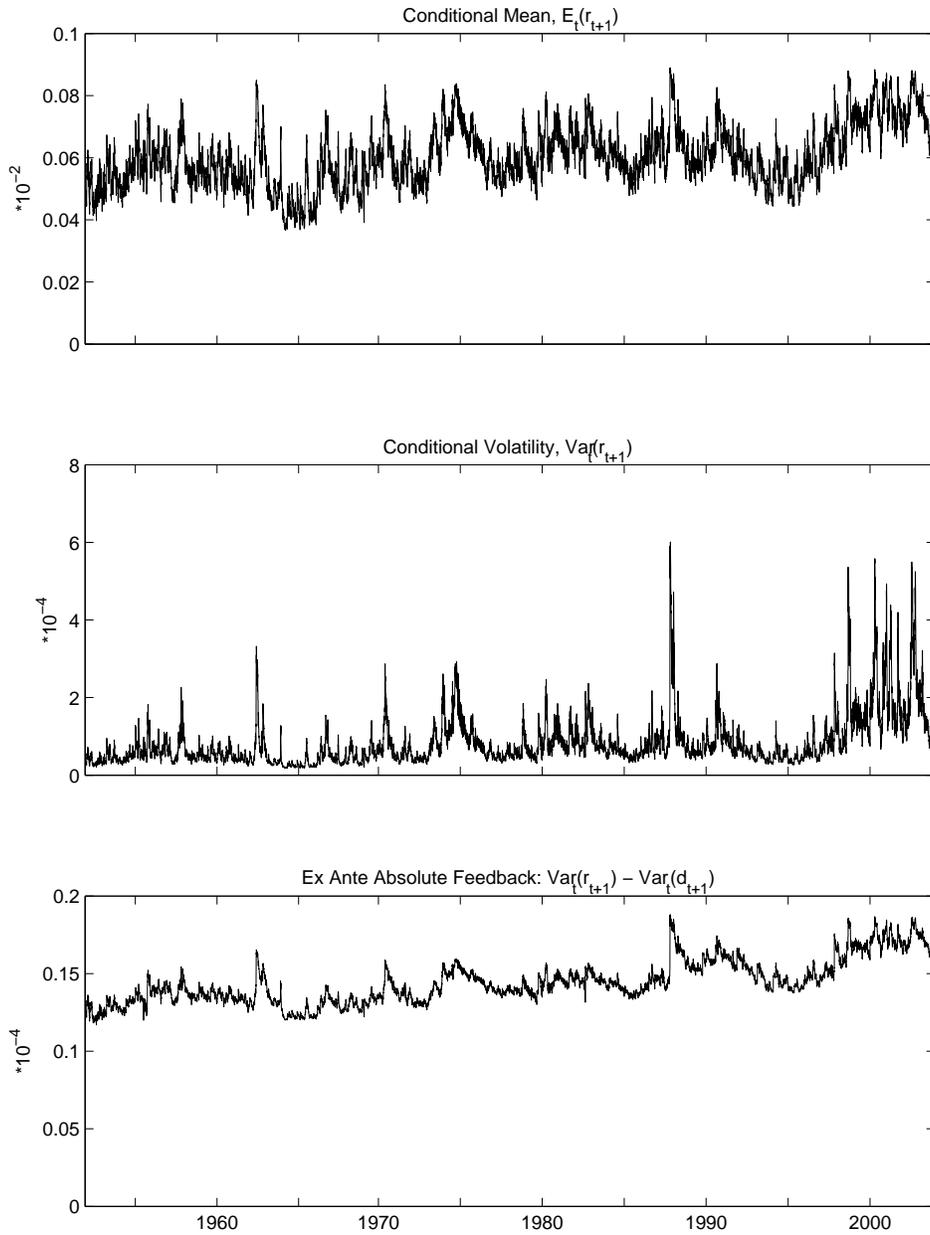


Figure 3: Ex Ante Conditional Mean, Volatility, and Feedback. This figure shows conditional moments of excess returns under the full information regime switching specification with $\bar{k} = 8$ volatility components. Conditioning information is the *ex ante* information set of returns up to and including date t . The three panels show consecutively the conditional mean and variance of returns, and the ex ante absolute feedback.

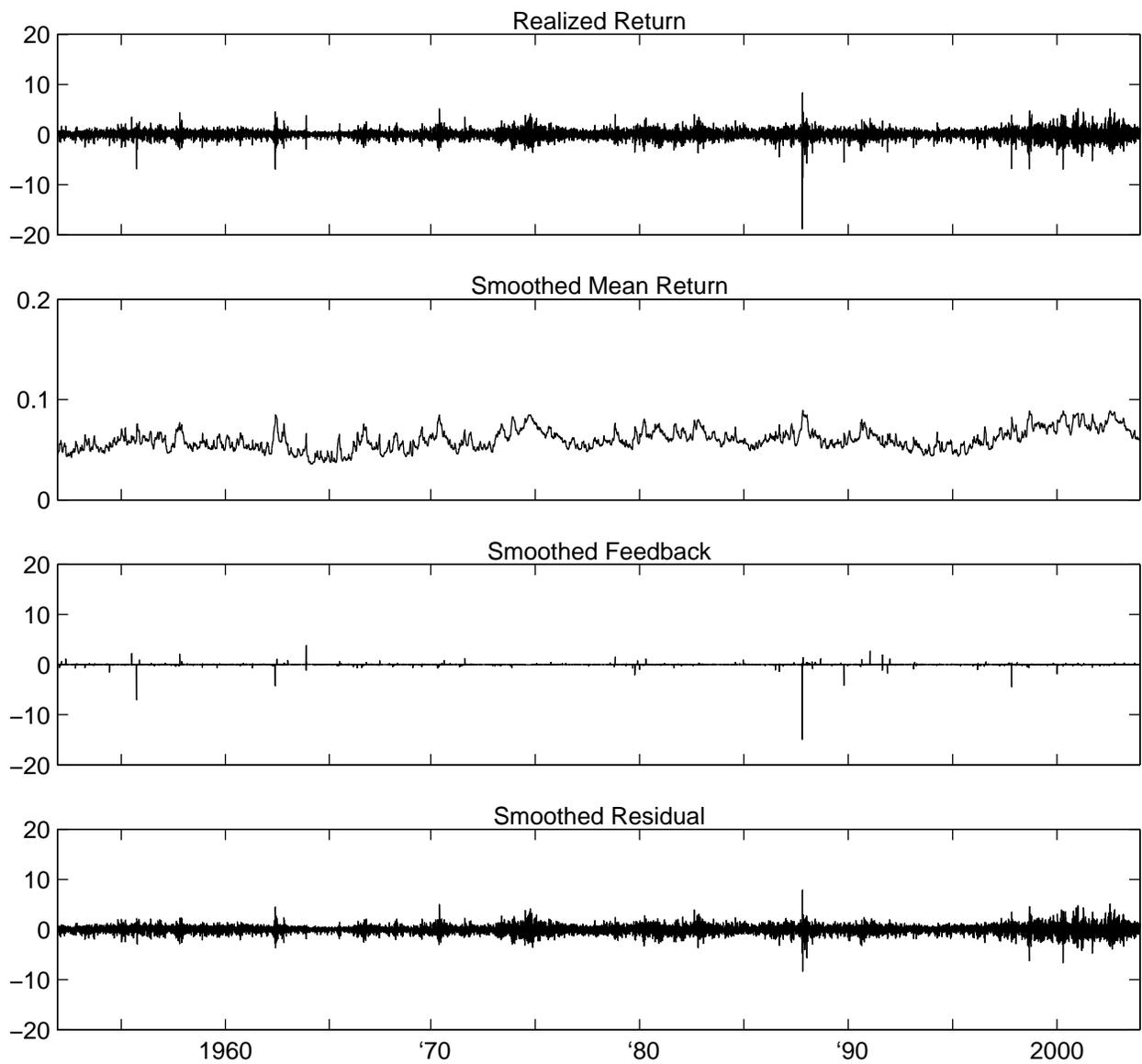


Figure 4: Ex Post Return Decomposition. This figure shows an *ex post* decomposition of realized returns using the full information regime-switching feedback model. The decomposition uses the smoothed beliefs Ψ_t obtained by using the conditioning information set of all returns. The first panel shows 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.

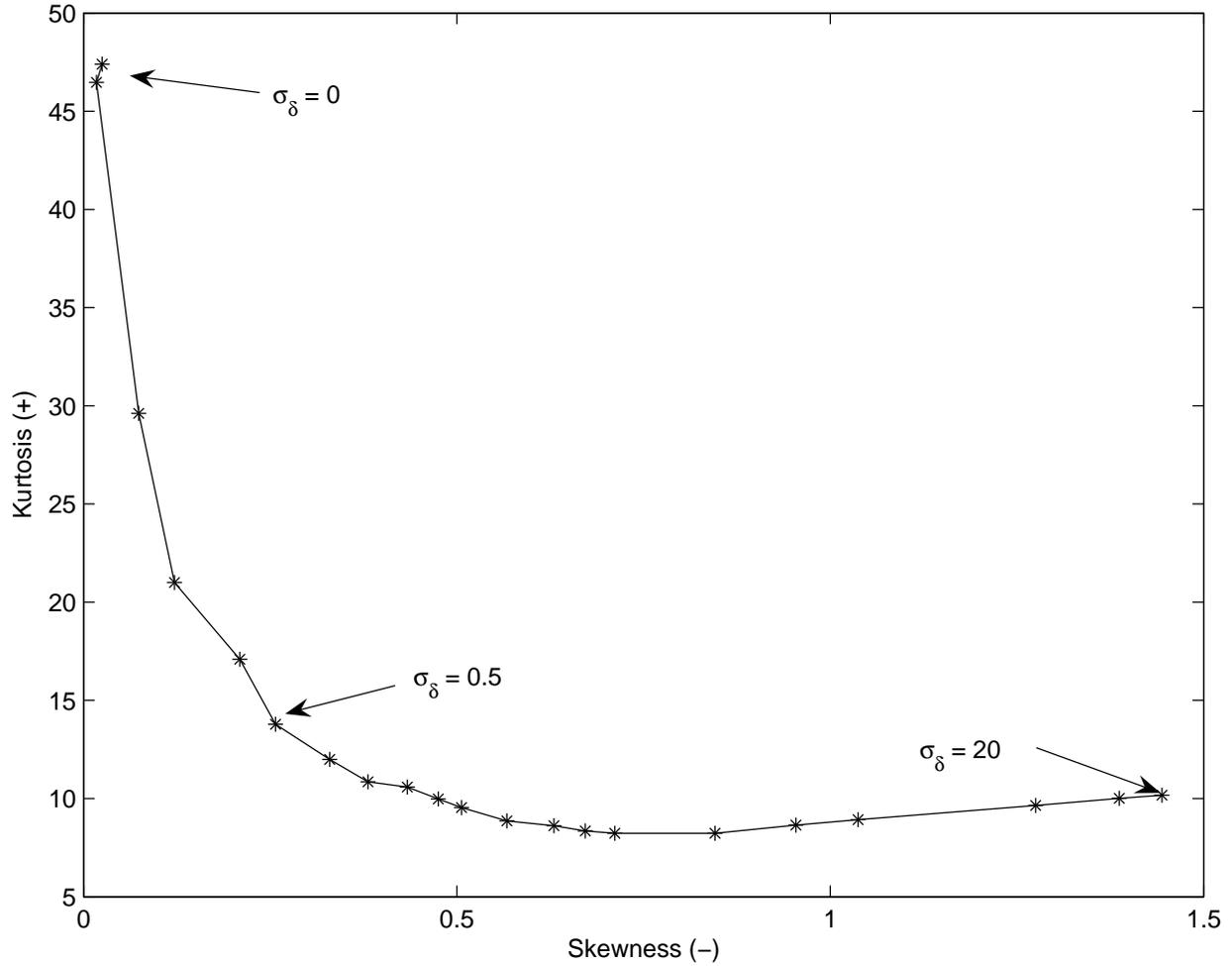


Figure 5: Learning Model Skewness and Kurtosis. This figure shows skewness and kurtosis for different information environments in the simulated learning model. The parameters $m_0 = 1.34$, $\sigma_d = 0.78$, $g_d - r_f = 0.15$, $\gamma_{\bar{k}} = 0.05$, and $b = 2.5$ are held constant across all simulations. Each simulation is then based on a different value of the signal standard deviation $\sigma_\delta \in 0, 0.1, \dots, 1, 1.25, \dots, 2, 3, 4, 5, 10, 15, 20$. When $\sigma_\delta = 0$, information is perfect, and as σ_δ becomes larger the signal precision weakens. For each specification, we simulate a single long series of $T = 3 \times 10^6$ returns using the same set of random draws. Each marked point on the plot represents a different simulation, progressing from $\sigma_\delta = 0$ in the top left to $\sigma_\delta = 20$ in the bottom right.

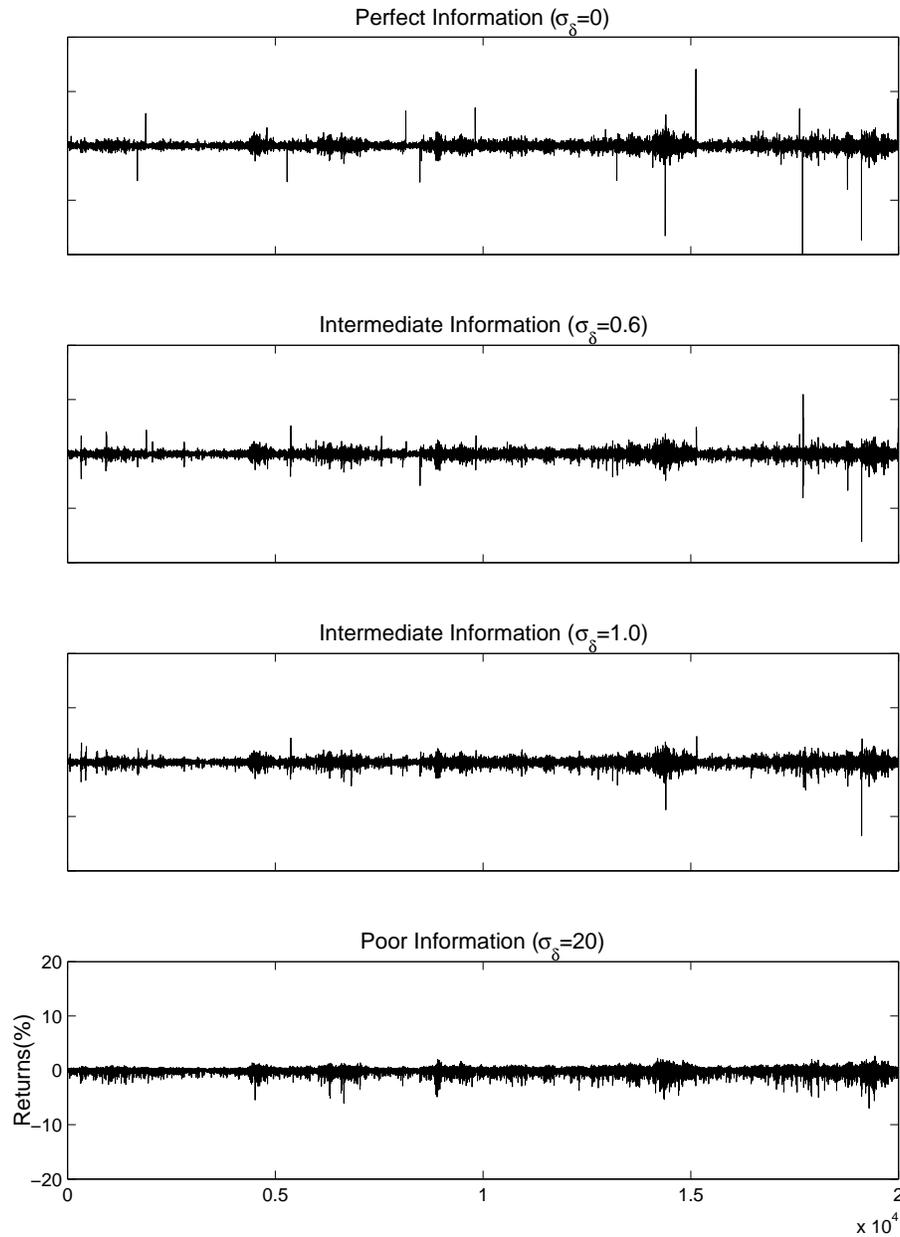


Figure 6: Simulations of the Learning Economy. This figure shows learning economy simulations with length $T = 20,000$. All simulations are based on the same set of base parameters $m_0 = 1.34$, $\sigma_d = 0.78$, $g_d - r_f = 0.15$, $\gamma_{\bar{k}} = 0.05$, and $b = 2.5$, and the same random draws for dividends, signal noises, and multipliers. Signal variability, σ_δ , is the only item that changes across simulations. The top panel corresponds to full-information. The middle two panels show intermediate learning environments, and in the final panel information quality is poor. We observe that noise in the investor signals attenuates extreme feedback realizations, but attenuation is stronger for positive than negative realizations. This generates increasing skewness as information quality falls.