Macroeconomic Uncertainty and Fear Measures Extracted From
Index Options*

Alexander David
Haskayne School of Business
University of Calgary

Pietro Veronesi
University of Chicago
NBER and CEPR

December 2008

*We thank Gurdip Bakshi, Michael Brennan, John Cochrane, George Constantinides, Gabrielle Demange, Darrell Duffie, Bernard Dumas, Rob Engle, Rick Green, Hayne Leland, Massimo Massa, Joe Ostroy, Rene Garcia, Ken Singleton, Allan Timmermann, Robert Whitelaw, Hong Yan for their comments and discussions. Parts of this paper have been presented at Carnegie Mellon, University of Chicago, CIRANO, Federal Reserve Board, INSEAD, London Business School, London School of Economics, McGill, Princeton, Stanford, Stockholm School of Economics, UCLA Economics, UCLA Finance, UCSD Economics, Washington University, Annual Derivatives Securities Conference at Boston University, European Summer Symposium in Financial Markets at Gerzensee, the Rodos Economic Theory Conference, the Western Finance Association meetings, and the World Congress of the Bachelier Finance Society in Paris for their comments and discussion. Address (David): Address: 2500 University Drive NW, Calgary, Alberta T2N 1N4, Canada. Phone: (403) 220-6987. Fax: (403) 210-3327. E. Mail: adavid@ucalgary.ca. Address (Veronesi): Graduate School of Business, University of Chicago, 5807 South Woodlawn Avenue, Chicago, IL, 60637. Phone: (773) 702-6348. E-mail: pietro.veronesi@gsb.uchicago.edu. David thanks the SSHRC for a research grant. Veronesi gratefully acknowledges financial support from the James S. Kemper Faculty Research Fund at the Graduate School of Business, University of Chicago.
Macroeconomic Uncertainty and Fear Measures Extracted From Index Options

Abstract

We demonstrate that about half the time series variation in two popular fear indices extracted from S&P 500 index options prices – the at-the-money implied volatility, and the ratio of implied volatilities of out-of-the-money puts and calls – can be explained by investors’ learning of the state of fundamentals through business cycles. The drifts of earnings growth at S&P 500 firms and inflation jointly jump, with greater instability during periods of higher inflation. Fear of the economy going into stagflation causes volatility during recessions, while the belief of a new economy growth state in the late 1990s led to an atypical bout of volatility in this period that accompanied strong fundamentals. Put-call ratios are higher in periods of stronger earnings growth when investors are more fearful of market corrections, and among these states, are highest during periods of medium inflation, which are states with the lowest upside potential. As further support for the learning mechanism we demonstrate the ability of our model to explain (i) the positive relation between volatility and the volatility of volatility, (ii) the puzzling change in sign of the relation between P/E and put-call ratios, and (iii) a significant amount of variation in the implied volatility premium.
Since the classic work of Breeden and Litzenberger (1978) it has been clear that option prices contain valuable information on investors’ forward looking state price density function (expected future stock price times marginal utility of the aggregate investor of the economy). In particular, two key statistics based on index option prices, the implied volatility of at-the-money (ATMIV) options and the ratio of implied volatilities of out-of-the-money (OTM) put to call options (P/C) are both popularly quoted as ‘fear’ indexes. One variant of the former trades under the ticket VIX on the Chicago Board Options Exchange (CBOE) which the exchange describes as:

One of the most interesting features of VIX, and the reason it has been called the “investor fear gauge,” is that, historically, VIX hits its highest levels during times of financial turmoil and investor fear. [CBOE Bulletin on VIX, 2003].

The latter, which is a direct market assessment of downside relative to upside risk has also been studied extensively since the work of Bates (1991) to imply, in particular, fears that investors have about a market crash. Reasonable alternatives that cause market wide fears include pessimism about macroeconomic prospects, increased risk aversion, or investor sentiment. However, the option pricing literature has generally failed to provide guidance on identifying these alternative channels that impact the fear indexes. In particular this literature has failed to extract macroeconomic information from options prices since most of the work assumes exogenous stock prices and their volatilities. In contrast, there is now a successful and burgeoning literature that extracts macroeconomic information from the term structure of interest rates [see, e.g. Estrella and Mishkin (1998), Ang and Piazzesi (2003) and Ang, Piazzesi, and Wei (2006)]. In this paper we integrate the two literatures by specifying a pricing model that uses information in fundamental data, stock prices, the term structure of interest rates, and in addition, options prices, which we use to extract information on the macroeconomy and explain the variation in the two fear indexes. Thus we provide a purely rational and non-behavioral explanation of fear in the

---

1Since the classic work of Black and Scholes (1977) the major innovations have been the addition of stochastic volatility [see, e.g., Hull and White (1987) and Heston (1993)], jumps in prices [see e.g. Bates (1996) and Bates (2000), and Pan (2002)], and jumps in volatility [see, e.g. Eraker, Johannes, and Polson (2003)]. A tremendous amount of empirical work has been done on these extensions of the BS formula that has enriched our understanding of stock price dynamics, and of options returns. Bakshi, Cao, and Chen (1997) provides a specification analysis of some of these models. Among more recent innovations, Christoffersen, Jacobs, Ornthanalai, and Wang (2008) build multi-factor stochastic volatility models, and somewhat related to our paper, Polson, Johannes, and Stroud (2008) price options when exogenously specified volatility follows an unobserved process that investors learn about. Finally, Constantinides, Jackwerth, and Perrakis (2008) find that several exogenously specified volatility models, such as GARCH, can be rejected as possible data generating processes for S&P 500 index options.
index options markets. We provide further support for the Bayesian learning mechanism that drives our model by studying the ability of our model to match three other stylized facts: (i) the volatility of volatility, (ii) the time variation relation between price-earning (P/E) ratios and P/Cs, (iii) the implied volatility premium. Neither of these variables is explicitly fitted by our empirical procedure, and therefore, these are cross-sectional out-of-sample predictions of our model.

Our model has a parsimonious joint specification of inflation and earnings growth, the two macroeconomic fundamentals that we focus on. The drift rates of these processes jump erratically through the business cycle. Earnings growth is more stable during periods of low inflation, and the worst outcome for stock prices is stagflation — a period of high inflation and low earnings growth. We assume that investors do not observe the drifts of the fundamental processes but learn about them from observations of fundamentals and additional signals (possibly generated by the analyst community). In periods when investors are less confident of their current estimates of the drift process, they revise their beliefs more rapidly, which causes greater volatility [see related work by David (1997) and Veronesi (2000)]. Therefore, prices and volatilities are endogenous in the model distinguishing it from classical option pricing models. The term structure is also endogenous and reflects investors assessments of the state of fundamentals. The model is similar to those by David (2006) and David and Veronesi (2008) that analyze credit spreads and one year ahead volatility forecasts, respectively, but do not study the implications for options prices. We fit the parameters of our structural model with an overidentified Simulated Method of Moments (SMM) procedure, which uses all the variables listed in the opening paragraph. It is important to note that when analyzing pricing relationships we consider a single process of uncertainty over time to gauge the success of our model in explaining the time-series variation in fundamentals and asset prices. This distinguishes our work from related work on options with learning that resets the uncertainty in each period and focuses on conditional price reactions.²

The model’s calibrated ATMIV explains about 50 percent of the variation in the data ATMIV over a 20-year sample from mid-1986 through mid-2006, a period in which we had two macroeconomic

²For example Buraschi and Jiltsov (2006) study option prices and volume in a model with learning in which agents have different priors. Dubinsky and Johannes (2006) study the reaction of options prices on individual stocks to news about earnings. Benzoni, Collin-Dufresne, and Goldstein (2005) show that the increase in investors’ perception about the average jump size of stock prices led to steepening of the implied volatility smirk after the stock market crash of 1987, but do not study its time variation in subsequent years.
recessions and several periods of massive market turbulence. It is important to note that we report all our results without lags of the endogenous variables, so that our empirical results are clearly interpretable. We find that the model ATMIV spikes in two joint states of earnings and inflation. One of these is the stagflation state, which never occurs in the 20-year sample, but nevertheless affects asset prices since investors’ expectations of transitioning to this state varies over our sample period. In particular, our model implies that volatility spikes when recessions fears increase due to the increased risk of stagflation. The other joint state in which volatility increases is during periods of low inflation and unusually high earnings growth, which occurred in our sample the late 1990s. Because investors in the model were never fully convinced that the economy had entered such a state, the high earnings growth in these periods led to increased earnings uncertainty and a high model ATMIV. This result is new to the volatility literature since volatility is generally associated with bad news. We further show that our model contains all the information in standard explanatory variables in the volatility literature such as interest rates, lagged stock returns, and the NBER-dated recession indicator.

The model’s calibrated P/C explains almost half the variation in the data P/C as well over the 20-year options sample. The model P/C is high in periods in which investors fear a rapid price drop in the following quarter. Generally, P/C ratios are larger than one in high growth states, as has been the case for almost this entire sample. Once again, fears of recessions and stagflation lower the P/C ratio in sub-periods. The P/C ratio is the highest in periods of medium inflation and high earnings growth, when earnings growth is quite unstable, but there is little upside risk since investors’ assess that it is impossible to transition to the new economy growth rate from this state directly in a quarter. This chance of transitioning to the new economy state is small, but higher in the low inflation and high growth state, and thus this state sustains a lower P/C. Our calibration shows that the run-up in the P/C ratio from 1990-1995 was largely explained by the increase in investors’ assessed likelihood of a period of low inflation and high earnings growth, which is stable. The decline in P/Cs during the second half of the 1990s is largely due to the continued “conquest of inflation” [see also Sargent, Williams, and Zha (2006)] that lowered investors’ beliefs of medium inflation and high growth which shifted towards the new economy state, which supports lower P/Cs. Finally, these two trends were somewhat reversed in the current decade that led to a slower increase in P/Cs. Therefore, macroeconomic factors drive

---

3Its important to note that we use fundamental data as well as data on stock and bond prices over a longer sample that starts in 1960.
a significant amount of variation in fears of price corrections in the market. To verify if this fear is relevant over and above other sentiment measures that might be proxies for fear or risk-aversion, we use as control variables two sentiment measures that affect the options smirk in Han (2008). The first is a measure of “trader sentiment” from trades of speculators in S&P 500 index futures. The second is a measure of “institutional investor sentiment” obtained from surveys of institutional money managers. Both these variables explain a small amount of variation in the P/C ratio (8 and 12 percent, respectively) and are insignificant once we include our model P/C.

We address three other stylized features of option prices that lend further support to the view that option price dynamics can be explained by investors learning about the state of fundamentals through business cycles. The first, as noted by Jones (2003), is that during periods of high ATMIV, such as in 1987, 1998, and 2001, ATMIV also fluctuated rapidly, that is, there is a positive relationship between volatility and the volatility of volatility. Jones (2003) points out that this stylized fact is incompatible with the Heston (1993) stochastic volatility model, and instead requires an explosive volatility process (one which violates certain regularity conditions). Our model permits a closed-form expression for the volatility of stock variance, and we are thus able to study its calibrated properties. Indeed we find that during periods of high fundamental uncertainty, investors revise their beliefs very rapidly causing not only high volatility, but also high volatility of volatility. The correlations between these variables for our model and in the data are of similar magnitude. Moreover, our model volatility process satisfies the regularity conditions that ensure volatility is stable.

Second, our model is able to shed light on the puzzling and compelling time varying relationship between P/E and P/C ratios. The overall relationship is negative in the data (correlation of -0.24), but its positive in the first 8 years of our sample, and strongly negative thereafter. Therefore, for the overall sample in periods of optimism, investors also price the risk of being disappointed and bid up the prices of puts relative to calls. Our model shows that this relationship reversed in the first part of our sample when beliefs of investors essentially shifted between the recession (medium inflation and low earnings growth) state and boom state (low inflation and high earnings). Both ratios are lower in the former state (see earlier discussion of P/Cs) and thus the series are positively correlated. In the second half of the sample, the major belief revisions are between the medium inflation and high growth and the new economy states. P/Es are higher in the latter while P/Cs are higher in the former state, thus delivering
the negative relation. Besides providing an understanding of the time variation our analysis suggests a different class of information that is relevant for options prices than previously understood. While it is widely accepted that information in analyst forecasts of earnings are relevant for pricing stocks, there is little literature on their relevance for options prices. Our analysis suggests that once condition to the state of fundamentals, analyst views can be used for better pricing of options as well.

Finally, the learning mechanism of our model has implications for the time series variation in the implied volatility premium (IVP), which is the difference between the ATMIV and a forecast of future volatility under the objective measure to the maturity of the options. We note that in a pure Gaussian model, the IVP would be comprised mainly of the forecast volatility risk premium (FVRP), which is the difference between the volatility forecasts under the risk-neutral and objective measures. For our calibrated model we find the average FVRP is small, of the order of 0.6 percent, while the IVP is much larger at 3.3 percent. This is a general property of pure diffusion models in which the FVRP can only arise from the difference in drifts under the two measures, and the drift effect is small for short horizons such as a quarter. To compare with the IVP in the data, we use the data ATMIV and a forecast of future volatility using standard regressors for volatility such as lagged volatility and stock returns. The data ATMIV is similar magnitude of around 3.7 percent, and our model IVP explains about 30 percent of its variation over the 20 year sample. We further investigate as to what aspect of our model causes its ATMIV and find that the calibrated volatility of volatility explains about a third of its variation in a simple linear regression. The latter is a measure of the fourth moment, which is known to increase the prices (implied volatilities) of options. Therefore, the non-normality of returns in our pure diffusion model arises endogenously from the dynamic updating of the conditional mean and volatility of returns and the resulting mixture of distributions. Finally, the model’s volatility of volatility is higher in periods of higher fundamental uncertainties as discussed above.

The plan for the remainder of the paper is as follows. In section 1, we provide the structure of the model and derive some key pricing results. In section 2 we estimate the parameters of our model using an overidentified simulated method of moments procedure. In section 3, we study the ability of our model to explain the two fear indices, and in section 4, we study its ability to understand the three additional stylized facts. Section 5 concludes. Two technical appendices provides proofs of technical results and the estimation methodology, respectively.
1 Structure of the Model

In this section we describe the structural model, which is fully described by Assumptions 1 through 6. The model has several of the features in David (2006), who introduces a model to price defaultable bonds.

**Assumption 1.** The price of the single homogeneous good in the economy, $Q_t$, evolves according to the process

$$\frac{dQ_t}{Q_t} = \beta_t \, dt + \sigma_Q \, dW_t,$$

(1)

where $W_t = (W_{1t}, W_{2t}, W_{3t}, W_{4t})'$ is a four-dimensional vector of independent Weiner processes, the $1 \times 4$ constant vector $\sigma_Q = (\sigma_{Q,1}, \sigma_{Q,2}, 0, 0)$ is assumed to be known by investors, and the process followed by $\beta_t$ is described below.

**Assumption 2.** Real earnings, $E_t$, evolves according to the process

$$\frac{dE_t}{E_t} = \theta_t \, dt + \sigma_E \, dW_t,$$

(2)

where the $1 \times 4$ constant vector $\sigma_E = (0, \sigma_{E,2}, 0, 0)$ is assumed known by investors and the process for $\theta_t$ is described below.

**Assumption 3.** Investors observe unbiased signals, $S_t$, of the drift rate of earnings that follow the process

$$\frac{dS_t}{S_t} = \theta_t \, dt + \sigma_S \, dW_t,$$

(3)

where the $1 \times 4$ constant vector $\sigma_S = (0, \sigma_{S,2}, \sigma_{S,3}, \sigma_{S,4})$ is assumed known by investors. The signals are not observed by the econometrician.

We use these fundamentals and the signal process to price all assets in our model. In order to do so, we need to specify a stochastic discount factor to be used to discount real payoffs. We make the following assumption:

**Assumption 4.** There exists a real pricing kernel $M_t$ taking the form

$$\frac{dM_t}{M_t} = -k_t \, dt - \sigma_M \, dW_t,$$

(4)
where \( \sigma_M = (\sigma_{M,1}, \sigma_{M,2}, \sigma_{M,3}, \sigma_{M,4}) \) is a \( 1 \times 4 \) constant vector of the market prices of risk, and \( k_t = \alpha_0 + \alpha \theta_t + \alpha \beta_t \) is the real short rate of interest conditional on the hidden state variables \( \theta_t \) and \( \beta_t \). The kernel \( M \) is used to price real claims and determine the expected real returns of all securities. We restrict the real rate of interest to be a linear function of the two (hidden) state variables of the model.

The kernel is observed by investors but not by the econometrician. Using it, agents price a generic real payout flow of \( \{D_t\} \) as

\[
M_t P_t = E \left[ \int_0^\infty M_s D_s ds | \mathcal{F}_t \right],
\]

where \( \mathcal{F}_t \) is investors’ filtration to be described in Assumption 5. As in several recent papers [for example, Berk, Green, and Naik (1999) and Brennan and Xia (2002)], we specify an exogenous pricing kernel process to formulate equilibrium relationships among endogenous financial variables. The linear dependence of the real rate on the drifts of fundamentals has a theoretical basis in general equilibrium models. For example, in a Lucas (1978) economy where investors have power utility \( U(C_t, t) = e^{-\frac{\phi C_t}{1-\gamma}} \), we would have \( C_t = E_t, M_t = U'(E_t) \), and hence \( k_t = \phi + \gamma \theta_t + \frac{\gamma}{2}(1-\gamma) \sigma_E \sigma'_E \) and \( \sigma_M = \gamma \sigma_E \). In this case, the real rate is not affected by the inflation drift, that is, \( \alpha \beta = 0 \). Our specification generalizes the real rate process to economies where expected inflation affects the real cost of borrowing. This specification of the real rate can be supported in an equilibrium if we allow a storage technology that can be bought or sold, is in perfectly elastic supply, and offers a safe instantaneous return of \( k_t \). We provide further motivation for its dependence on expected inflation next.

The literature on the effect of expected inflation on the real rate of return is extensive, and it is beyond the scope of this paper to build in all the effects without significantly complicating our analysis. David (2006) provides further description of these effects, some of which we describe below. At the micro level there is a fairly large literature on various tax and accounting channels through which expected inflation affects the real return on capital [Feldstein (1980a), Feldstein (1980b), Auerbach (1979), and Cohen and Hassett (1999)]. Among the several monetary channels that lead to the non-neutrality of money, of particular relevance are cash-in-advance models in which expected inflation is a tax on money balances and raises the real cost of transactions, thus affecting the real interest rate, capital accumulation, and business cycle variation [see Cooley and Hansen (1989) for an empirical analysis of the size of this effect]. Given the multiplicity of channels that affect the relationship between expected
inflation and the real short rate, we treat the sign and size of this relationship as an empirical question that can be estimated from the joint time-series variation of fundamentals and asset prices. Our chosen functional form of the real rate can be seen as a reduced form for the above noted effects.

For notational convenience, we stack the fundamental processes (1) and (2). Let \( Y_t = (Q_t, E_t)' \), so that

\[
\frac{dY_t}{Y_t} = \varrho_t \, dt + \Sigma_2 \, dW_t,
\]

(6)

where \( \frac{dY_t}{Y_t} \) is to be interpreted as “element-by-element” division, \( \varrho_t = (\beta_t, \theta_t)' \), and \( \Sigma_2 = (\sigma'_E, \sigma'_Q)' \). Similarly, we find it useful to add the signal and kernel processes to the stacked vector and write \( X_t = (Q_t, E_t, M_t, S_t)' \), which has the drift vector \( \nu_t = (\beta_t, \theta_t, -k_t, \theta_t)' \), and volatility matrix \( \Sigma = (\sigma'_Q, \sigma'_E, -\sigma'_M, \sigma'_S)' \).

**Assumption 5.** The drift vector \( \nu_t \) follows an \( N \)-state, continuous-time finite state Markov chain with generator matrix \( \Lambda \), that is, over the infinitesimal time interval of length \( dt \)

\[
\lambda_{ij} \, dt = \text{prob}(\nu_{t+dt} = \nu_j | \nu_t = \nu_i), \quad \text{for } i \neq j, \quad \lambda_{ii} = -\sum_{j \neq i} \lambda_{ij}.
\]

The transition matrix over states in a finite interval of time, \( s \), is \( \exp(\Lambda s) \) [see, for example, Karlin and Taylor (1982)].

**Assumption 6.** Investors do not observe the realizations of \( \nu_t \) but know all the parameters of the model.

Since investors do not observe \( \nu_t \), they need to infer it from the observations of past earnings, inflation, the signal, and the pricing kernel. This learning process will generate a distribution on the possible states \( \nu_1, ..., \nu_N \) that in turn generates changes in “uncertainty” as they learn about the current state. At time \( t \), investors’ distribution over hidden states is summarized by the posterior probabilities

\[
\pi_{it} = \text{prob}(\nu_t = \nu_i | \mathcal{F}_t),
\]

where \( \mathcal{F}_t \) is the filtration generated by observing the entire path \( (X_s)_{0 \leq s \leq t} \). Let \( \pi_t = (\pi_{1t}, ..., \pi_{Nt}) \) be the vector of beliefs.
Lemma 1. Given an initial condition $\pi_0 = \hat{\pi}$ with $\sum_{i=1}^N \hat{\pi}_i = 1$ and $0 \leq \hat{\pi}_i \leq 1$ for all $i$, the probabilities $\pi_{it}$ satisfy the $N$-dimensional system of stochastic differential equations:

$$d\pi_{it} = \mu_i(\pi_t) dt + \sigma_i(\pi_t) d\tilde{W}_t,$$

in which $\mu_i(\pi_t) = E_t[\pi_t \Delta i]$, $\sigma_i(\pi_t) = \pi_{it} [\nu_i - \overline{\nu}(\pi_t)]' \Sigma^{-1}$, and $\overline{\nu}(\pi_t) = \sum_{i=1}^N \pi_{it} \nu_i = E_t \left( \frac{dX_t}{X_t} | \mathcal{F}_t \right)$.

$$d\tilde{W}_t = \Sigma^{-1} \left[ \frac{dX_t}{X_t} - E \left( \frac{dX_t}{X_t} | \mathcal{F}_t \right) \right] = \Sigma^{-1} (\nu_t - \overline{\nu}(\pi_t)) dt + dW_t.$$

Moreover, for every $t > 0$, $\sum_{i=1}^N \pi_t = 1$.

Proof. See Wonham (1964) or David (1993).

The filtering theorem for jumps in the underlying drift was first derived (to the best of our knowledge) in Wonham (1964). David (1993) provides a proof using the limit of Bayes’ rule in discrete time. The first application of this theorem in financial economics, as well as several properties of the filtering process, are derived in David (1997). In a parallel development in mathematical finance, Elliott, Aggoun, and Moore (1995) provide results for this filtering process under an alternative equivalent measure that simplifies the filtering process for some purposes.

We make the following summary comments about the updating process (7): (i) The elements of the generator matrix, $\Lambda$, completely capture the transition probabilities between states. Absent any new information, beliefs tend to mean-revert to the unconditional stationary probabilities that are completely determined by $\Lambda$. For example, if there are only two states with $\lambda_{12} = 2$ and $\lambda_{21} = 1$, then the belief that the economy is in state 1 mean reverts to $1/3$. (ii) The diffusion term, $(\nu_i - \overline{\nu}(\pi_t)) (\Sigma')^{-1}$, describes the magnitude of the change in beliefs due to uncertainty in each of the components: inflation, earnings, and real rates. For example, investors’ probability of being in state $i$ will react more to inflation shocks when the inflation rate in that state is further away from investors’ conditional mean of inflation at that moment. In such periods investors revise their beliefs more rapidly. Their beliefs are also more volatile in economies with lower noise in fundamentals and higher precision of signals as captured in the $(\Sigma')^{-1}$ term. (iii) Agents update their beliefs about underlying states by observing not only the path of fundamentals and signals, but also their pricing kernel. This is analogous to the Lucas
economy mentioned above in which the agent would learn about the drift rate of the marginal utility of consumption.

It is useful to note that investors’ beliefs change with their inferred shocks, \(d\tilde{W}\), in equation (9) as opposed to the true shocks, \(dW\), that affect fundamentals. Similarly, they infer that the process \(\{X_t\}\) follows: \(dX_t/X_t = \nu(\pi_t)dt + \Sigma d\tilde{W}_t\). At each point in time, substituting the definition of \(d\tilde{W}_t\), we have \(dX_t/X_t = \nu(\pi_t)dt + \Sigma^{-1} [\nu_t - \Sigma(\pi_t)]dt + dW_t\) = \(\nu_t dt + \Sigma dW\), which is the same process as in Assumptions 1 through 3. In the filtering literature, \(d\tilde{W}_t\) is an “innovations” process under the investors’ filtration and under the separation principle it can be used for dynamic optimization. See David (1997) for a discussion. As a special case we write

\[
\frac{dY_t}{Y_t} = \tilde{\vartheta}(\pi_t)dt + \Sigma_2 d\tilde{W}_t = \vartheta(\pi_t)dt + \Sigma_2 dW_t, \tag{10}
\]

and the kernel under investors’ filtration as \(dM_t/M_t = -\kappa(\pi_t)dt - \sigma_M d\tilde{W}_t\), where the real rate in the economy, \(\kappa(\pi_t)\), is its expected value conditional on investors’ filtration.

1.1 Stock Prices and the Term Structure of Interest Rates

To evaluate nominal claims and nominal risk premiums we will also use the nominal pricing kernel, \(N_t = M_t/Q_t\), which follows

\[
\frac{dN_t}{N_t} = -r_t dt - \sigma_N dW_t, \tag{11}
\]

where \(r_t = k_t + \beta_t - \sigma_N \sigma'_Q\) and \(\sigma_N = \sigma_M + \sigma_Q\). The nominal rate differs from the real rate by the sum of expected inflation and the inflation risk premium, which is the covariance between inflation and the nominal pricing kernel. With unobserved states, the projected (observable) nominal interest rate at time \(t\) is \(\bar{r}(\pi) = \sum_{i=1}^N r_i \pi_{it}\), where \(r_i = k_i + \beta_i - \sigma_N \sigma'_Q\) is the nominal rate that would obtain in the \(i^{th}\) state, were the states observable. The following proposition provides expressions for the price-earnings (henceforth P/E) ratio and the nominal bond price:

**Proposition 1.**
(a) The P/E ratio at time $t$ is

$$\frac{P_t}{E_t}(\pi_t) = \sum_{j=1}^{N} C_j \pi_{jt} \equiv C \cdot \pi_t,$$

(12)

where the vector $C = (C_1, \ldots, C_N)$ satisfies $C = A^{-1} \cdot 1_N$.

$$A = \text{Diag}(k_1 - \theta_1 + \sigma_M \sigma'_E, k_2 - \theta_2 + \sigma_M \sigma'_E, \ldots, k_N - \theta_N + \sigma_M \sigma'_E) - \Lambda.$$  

(13)

(b) The price of a nominal zero-coupon bond at time $t$ with maturity $\tau$ is

$$B_t(\pi_t, \tau) = \sum_{i=1}^{N} \pi_{it} B_i(\tau),$$

(14)

$$B_i(\tau) = E \left( \frac{M_{t+\tau}}{M_t} \cdot \frac{Q_t}{Q_{t+\tau}} | \nu_t = \nu_i \right) = \sum_{i=1}^{N} \Omega_i e^{\omega_i \tau} \cdot (\Omega^{-1} 1_N),$$

(15)

where $\Omega_i$ and $\omega_i$, $i = 1, \ldots, N$, are the $i^{th}$ eigenvector and $i^{th}$ eigenvalue, respectively, of the matrix $\hat{\Lambda} = \Lambda - \text{Diag}(r_1, r_2, \ldots, r_N)$.

The proof follows from a simple extension of the proofs for stock and bond prices in Veronesi (2000) and Veronesi and Yared (1999), respectively.

In (a), each constant $C_i$ represents investors’ P/E valuation of stocks conditional on the state being $\nu_i$ today. As in the classic Gordon model, the valuation depends on the expected growth of earnings, the real rate, and the equity premium. Since investors do not observe the state $\nu_i$, they weight each $C_i$ by its conditional probability $\pi_i$ thereby obtaining (12). Notice in particular that the form of the constant vector $C$ suggests that: (i) if $\alpha_\theta < 1$, a higher growth rate of earnings implies a higher P/E, (ii) if $\alpha_\beta > 0$, a higher inflation state implies a lower P/E, which is the real rate effect of inflation, and (iii) in addition, the P/E ratio in a given state of growth depends on the future sustainability of the growth rate and is determined by the transition probabilities $\lambda_{ij}$ as shown in the solution to the $N$-equations in (a).

Similarly, the bond price is a weighted average of the nominal bond prices that would prevail in each state $\nu_i$. Since investors do not actually observe the current state, they price the bond as a weighted average. Both higher inflation and higher growth rate of earnings lead to lower long term bond prices.
when $\alpha_\beta > -1$ and $\alpha_\theta > 0$. It is useful to notice that all asset prices follow continuous paths even though the drift rates for earnings and inflation jump between a discrete set of states. This results from the continuous updating process.

Let $P_t^N = P_t \cdot Q_t$ be the nominal value of stock, where $P_t$ is the real value of stocks in Proposition 1. Using the dynamics of the inflation and earnings processes under the observed filtration, we now formulate the nominal return processes for stocks and bonds.

**Proposition 2.**

(a) The nominal stock return process under the investor’s filtration is given by

$$\frac{dP_t^N}{P_t^N}(\pi_t) = (\mu^N(\pi_t) - \delta(\pi_t)) \, dt + \sigma^N(\pi_t) \, d\tilde{W}_t,$$

where $\delta(\pi)$ is the dividend yield, and the nominal stock price volatility is

$$\sigma^N(\pi_t) = \sigma_E + \sigma_Q + \sum_{i=1}^{N} C_{it} \pi_{it} \left( \nu_t - \overline{\nu}(\pi_t) \right)'(\Sigma')^{-1} \sum_{i=1}^{N} C_{it} \pi_{it}.$$

(16) **Proof:** Follows from a simple adaptation of the proof in Veronesi (2000).

Stock price volatility has an exogenous component due to noise in the fundamental process and a learning-based endogenous component, which depends on the volatility of each state probability $\pi_i$. The volatility of each state probability depends on the fundamental uncertainties as discussed in comment (ii) below Lemma 1. However, the volatility of stock prices depends additionally on the valuation of stocks in each state as measured by the P/E vector, $C$. For a given news content, states in which valuations are higher contribute more to the total volatility. Since the valuations vary with both expected earnings growth and discount rates, stock price volatility in our model will endogenously load differently on news about these components in different stages of the business cycle as found in Boyd, Hu, and Jagannathan (2005). If, for example, in state 5 the earnings drift is far from its value in other states, so that the volatility of the probability of this state is most reactive to earnings news, then stock volatility will give a larger (smaller) weight to this earnings news if the P/E ratio is high (low) in this state. We will discuss this point further in Section 2 for our calibrated model.
1.2 Return Volatility and its Covariance with Stock Returns

In this subsection we characterize the volatility of equilibrium stock returns as well as its covariance with the return process itself. As is well known, these two quantities are extremely important in determining the behavior of option prices.

We start by introducing the following notation. Let

\[\pi_i = \frac{\pi_i C_i}{\sum_{j=1}^{N} \pi_j C_j}\]  \hspace{1cm} (17)

As in Veronesi (2000), we call \(\pi^o = (\pi^o_1, \ldots, \pi^o_N)\) the *value-weighted probabilities* (notice that \(\pi^o_i \geq 0\) for each \(i\) and \(\sum_{i=1}^{N} \pi^o_i = 1\). From now on, a “\(o\)” denotes a quantity computed with respect to the distribution \(\pi^o\). For example, \(\overline{\theta}^o\) denotes the mean of the drift vector \(\theta\) computed using the distribution \(\pi^o\),

\[\sigma_{\theta \beta} = \sum_{i=1}^{N} \pi_i (\theta_i - \overline{\theta})(\beta_i - \overline{\beta}); \hspace{0.5cm} \sigma^o_{\theta \beta} = \sum_{i=1}^{N} \pi^o_i (\theta_i - \overline{\theta}^o)(\beta_i - \overline{\beta})\]  \hspace{1cm} (18)

are the covariances of the drift vectors \(\theta\) and \(\beta\) computed using \(\pi\) and \(\pi^o\), respectively. In addition we denote \(\sigma_{\theta \nu}\) and \(\sigma^o_{\theta \nu}\) to be the vectors of covariances of \(\theta\) with each element of the vector \(\nu\) using the two sets of probabilities respectively. We then have:

**Proposition 3**

(a) Stock return variance is given by

\[V = \sigma^N (\pi_i) \sigma^N (\pi_i) = (\sigma_E + \sigma_Q)(\sigma_E + \sigma_Q) + (\nu^o - \bar{\nu})(\Sigma \Sigma')^{-1} (\nu^o - \bar{\nu}) + 2 \left[ (\overline{\theta}^o - \overline{\theta}^o)(\overline{\beta}^o - \overline{\beta}) \right] \]  \hspace{1cm} (19)

(b) Return variance \(V\) follows the process \(dV = \mu_V dt + \sigma_V dW\), where \(\sigma_V =

\[2 \left[ \sum_i \left( [\pi^o_i (\nu_i - \nu^o) - \pi_i \nu_i'] (\Sigma \Sigma')^{-1} (\nu^o - \bar{\nu})(\nu_i - \bar{\nu})' \right) + (\sigma^o_{\theta \nu} - \sigma_{\theta \nu})' + (\sigma^o_{\beta \nu} - \sigma_{\beta \nu})' \right] \Sigma'^{-1}\]  \hspace{1cm} (20)

*Proof.* In Appendix 1.
The proposition implies that return variance is stochastic and the covariance between return and variance is stochastic given by
\[
\text{Cov} \left( dV, \frac{dS}{S} \right) = \sigma_V \sigma'_S, \tag{21}
\]
where stock volatility is in (16) and and the volatility of variance in (20). We will see below in Section 3 that for our calibrated model this covariance can change sign and magnitude leading to changes in the slope of the implied volatility curve for options prices.

### 1.3 Option Prices

By standard no-arbitrage pricing [e.g., Chapter 5.G in Duffie (1996)], every contingent claim \( f(t, z_t, \pi) \) must satisfy the fundamental partial differential equation (PDE):
\[
r(\pi) f = \frac{\partial f}{\partial t} + \frac{\partial f}{\partial z}(r(\pi) - \delta(\pi)) - \frac{1}{2} \sigma^N(\pi) \sigma'^N(\pi) + \sum_{j=1}^{N} \frac{\partial f}{\partial \pi_j} (\mu_j(\pi) - \rho_j(\pi))
\]
\[
+ \frac{1}{2} \frac{\partial^2 f}{\partial z^2} \sigma^N(\pi) \sigma'^N(\pi) + \sum_{j=1}^{N} \frac{\partial^2 f}{\partial z \partial \pi_j} \sigma^N(\pi) \sigma'_j(\pi) + \frac{1}{2} \sum_{k=1}^{N} \sum_{j=1}^{N} \frac{\partial^2 f}{\partial \pi_k \partial \pi_j} \sigma_k(\pi) \sigma'_j(\pi),
\]
where \( z = \log(P^N) \), \( \mu_i(\pi) \) and \( \sigma_j(\pi) \) are given in Equation (8), \( \sigma_N(\pi_l) \) is in Equation (16), and \( \rho_i(\pi) \), which is the market price of risk of the state variable \( \pi_i \), is the covariance of \( \pi_i \) with the nominal pricing kernel of the economy. David (2006) shows that
\[
\rho(\pi) = \pi_i \cdot (\beta_i - \bar{\beta}(\pi)) - (k_i - \bar{k}(\pi)).
\]

In addition, natural boundary conditions at 0 and 1 for each belief process and terminal conditions determined by the final payoff of the contingent claim are imposed. David and Veronesi (2002) solves this PDE using Fourier Transforms in the stock price dimension and series solutions in the beliefs dimensions when \( N = 2 \). David (2006) solves this PDE with \( N = 4 \) using projection methods in the beliefs dimensions. The number of complete polynomials each of length 7 needed to span this functional operator is 220. Increasing \( N \) to 5 (which is the number of drift states that we require in this paper), increases the length of the basis to 715 and makes the method too time consuming. Instead, in
this paper we use Monte-Carlo simulation using quasi-random sequences and several variance reduction techniques. The advantage of the simulation methods is that they do not suffer from the curse of dimensionality as the projection methods above. Details of the simulation procedure are provided in Appendix 2.

2 Estimation

In this section we provide a brief description of the estimation methodology and the parameter estimates of our model.

2.1 Estimation Methodology

We estimate the model by using information in fundamentals, stock and bond prices, and options prices. Fundamentals are included since investors’ information sets clearly contain the history of all fundamental data. However, since they likely have additional information on the hidden states of macroeconomic fundamentals, we attempt to extract it from asset prices. We estimate the time series of investors’ beliefs over fundamental states, as well as the underlying parameters using a Simulated Method of Moments (SMM) method as described below.

Let \( \Psi \) denote the set of structural parameters in the fundamental processes of inflation, earnings, earnings signals, and the pricing kernel in equations (1), (2), (3), and (4), respectively. Let the likelihood function for the fundamentals data observed at discrete points of time (quarterly) be \( L \). To extract information about investors’ beliefs from the prices of stocks, bonds, and stock options, we use the pricing formulae for the P/E ratio and Treasury bond prices in Proposition 1 to generate model-determined moments. In addition we use the simulation based options prices as shown in Appendix 2. Let \( \{e(t)\} \) denote the errors of the asset prices, and define \( \epsilon(t) = (e(t)', \frac{\partial L}{\partial \Psi} (t)')' \), where the second term is the score of the likelihood function of fundamentals with respect to \( \Psi \). We now form the SMM objective:

\[
\begin{align*}
    c &= \left( \frac{1}{T} \sum_{t=1}^{T} \epsilon_t \right)' \cdot \Omega^{-1} \cdot \left( \frac{1}{T} \sum_{t=1}^{T} \epsilon_t \right) \\
    (23)
\end{align*}
\]

The details of the procedure are in the Appendix.
It is worth emphasizing three key points about our choice of the SMM method of estimation. First, a simulation-based approach is necessary in our case since the likelihood function for the fundamental data observed at discrete points in time is not available in closed-form. In addition, series-based approximations of the likelihood function as in Ait-Sahalia (2002) are cumbersome given the high dimensionality of our state place (five dimensions for the beliefs and two for fundamentals). Second, our SMM approach allows for the fact that the econometrician observes data only on fundamentals while investors in addition observe their pricing kernel and signals about earnings, and hence update their beliefs about fundamental drifts based on a finer information filtration. Finally, the procedure combines information in asset price and volatility moments with the information in fundamental data so that the extracted investors’ beliefs are potentially quite different from estimation methods that rely only on fundamental information [see, for example, Hamilton (1989)].

2.2 Estimation Results for the Regime Switching Model

In this subsection, we briefly describe the results of the estimation of our model. We start with the description of the data series used. Our data sample runs from 1960 to 2006. Aggregate earnings for the economy are approximated as the operating earnings of S&P 500 firms, and these data are obtained from Standard and Poor’s. Similarly, the aggregate P/E ratio is estimated as the equity value of these firms divided by their operating earnings. Dividends for these firms, also obtained from Standard and Poor’s, are used with the prices to compute returns. We use the Consumer Price Index (CPI), obtained from the Federal Reserve Bank of St. Louis, as our inflation series, which is used both for discounting nominal earnings, as well as for forecasting future real earnings using our joint regime switching model of fundamentals. The time series of zero-coupon yields and returns on bonds of different maturities are obtained from the Fama-Bliss data set available at the University of Chicago. We obtain transactions data on S&P 500 index options from 1986:Q2 to 1996:Q1 from the CBOE. These data are no longer available from 1996:Q1, and therefore, we use data on these same options from Option Metrics from 1996:Q2 to 2006:Q2. We construct quarterly time series by using the option prices around the 15th day of the first month of each quarter. In our empirical exercises, we use the fundamental data observations available at the end of the previous quarter to match with these options data. The 15 day lag in using the option data helps to ensure that the fundamental data are available to traders in the options markets,
which may not be available in real time. It is also important to note that Option Metrics provide the average of bid and ask prices at the end of each trading day, and not prices based on actual transactions.

We estimate a model with three regimes each for inflation characterized by the states $\beta_1 < \beta_2 < \beta_3$ and earnings growth $\theta_1 < \theta_2 < \theta_3$, which lead to nine composite states overall. In our estimation, we find that the unconstrained estimates of the transition matrix led to several zero elements, leading to a more parsimonious five state model with the following states: $\{(\beta_1, \theta_2), (\beta_2, \theta_1), (\beta_2, \theta_2), (\beta_3, \theta_1), (\beta_1, \theta_3)\}$. We find that the remaining four states have close to zero probability of occurring in the sample. Overall, the five-state and nine-state models lead to almost the same value for the SMM objective function. Gray (1996) and Bansal and Zhou (2002) use a similar criterion for the choice among alternative regime specifications.

The top three panels of Table 1 reports estimates of the drifts and volatilities of fundamentals and signals. The fourth panel reports the transition probability (generator) matrix, while the fifth panel reports the estimates of the parameters of investors’ pricing kernel. We estimate that inflation averages 2.1%, 4.1%, and 8.4% in the three states, which we shall refer to as low (LI), medium (MI), and high (HI) states of inflation. Earnings growth averages -3.5%, 2.5%, and 5.4% which we classify as regular low (LG) and high (HG) growth rate states, and the “new economy” (NG) growth rate state, respectively. We provide further motivation for the names given to these states by looking at the implied quarterly and five-year transition probability matrices in the top two panels of Table 2. These matrices are derived solely from the elements of the generator elements in Table 1. We provide some summary descriptions next.

Notice first that the regular high growth rate of earnings, $\theta_2$, is far more persistent in the low inflation state: from the (LI-HG) state, there is a 99% chance of returning to this state and a 1% chance of transitioning to the (MI-HG) state in a quarter, and thus there is almost a zero chance of growth slowing in a quarter in which there is also low inflation. From the (MI-HG) state, on the other hand, there is an 2.2% chance of a transition to the (MI-LG) state in the following quarter. This is the signaling role of inflation – it provides an early warning signal of an unsustainable high growth rate of fundamentals. Second, we see from the bottom panel of Table 2 that, by our model estimates, regular high growth rates (states 1 and 3) would occur about 70% of the time and regular low growth states would occur about 27.5% of the time. In fact, in a model with a similar structure, David (2006) shows that these four
states provide a good fit for fundamentals and credit spreads. However, including short-term (quarterly) stock volatility as one of the overidentifying conditions leads to a rejection of the four-state model. We will see shortly that this failure is largely due to the spectacular rise of stock volatility in the late 1990s and its equally spectacular fall in the first half of the current decade. These massive swings seem to be unrelated to the transitions of the economy between the four basic macro states as seen in the time series of implied investors’ state probabilities that are shown in Figure 1. Overall, as seen, investors have held fairly strong views that the U.S. economy has remained in the LI-HG states. For most of the decade of the 1960s this probability hovered around 85%, while in the 1990s it was lower at around 70%. There were several switches within the medium inflation states in the 1970s, with two bouts of high inflation and low growth in the mid-1970s and early 1980s. The mid-to-late 1980s were characterized by high (nearly 40%) probabilities of the MI-HG state, which steadily declined through the end of the 1990s, although the probabilities of this state rose substantially to about 35% before the recession in 1990. This trend decline in expected inflation has also been noted by other authors [see, for example Sargent, Williams, and Zha (2006)]. In the current decade, the probability of the MI-HG state has slowly trended up and has averaged about 5% from 2002 to 2005. Unlike all previous recessions, the biggest reassessment in investors’ expectations in the 2001 recession related to earnings growth rather than inflation, as investors sharply revised their beliefs that the economy was in the NG (fifth) state. We discuss the implications of this revision further below.

The fifth state, (LI-NG), has a stationary (long-run) probability of only about 2.5%, but its inclusion in our model helps to a large extent to explain the high P/Es and high stock market volatility in the late 1990s. During this period, earnings growth was far above its historical average, and led investors to conjecture that due to productivity increases, there was a “new economy growth rate.” As we see from investors filtered probabilities in Figure 1, investors were very uncertain that the economy was in this state, which led to high stock market volatility. In this sense, the experience of the late 1990s was extraordinary: while earnings grew rapidly in the 1960s as well, the probability of the NG state only briefly touched 5%, while in 1999 it reached 25%. This channel of expectations of high growth rates causing the high P/Es and high volatility in the 1990s was also used by Pastor and Veronesi (2006). As in their model, we see in Proposition 1 that P/Es are convex in expected growth rates of earnings, which implies that in periods of higher uncertainty, ceteris paribus, P/Es are higher. However, their paper does
not analyze the longer term relationship between uncertainty, P/Es, and volatility. Examining the time series of uncertainty however, we see that in all other periods of high uncertainty in our sample, P/Es were low, because in such periods investors feared a recession and had low expectations of earnings growth. Finally, it is also useful to note that by our model estimates, the new economy growth only occurs in low inflation states, so that a pickup in inflation lowers investors’ assessed probability that the new economy growth will persist. As seen in Figure 1, the slight trend increase in the probability of the MI-HG state in the current decade has been accompanied by a steady decline in the probability of the LI-NG state.

The low probability of occurrence of the high drift rate of earnings lends some similarity of our model-estimated parameters to the recent work on “long-run risks” starting with the work of Bansal and Yaron (2004). This channel is used to generate a large equity premium in their paper. In our model, investors are confronted with the possibility that a low probability switch of fundamentals occurred in the late 1990s. As can be seen from the middle panel of Table 2, even at the five-year horizon there is only about a 1% chance of the economy entering the NG state from any other state. This low probability explains why investors’ filtered probability of this state reached a maximum of only around 25%, raising P/Es, but due to their high uncertainty, also raising stock market volatility. This role of the small probability of entering the LI-NG state distinguishes it from the LI-HG state, even though the persistence of both states, as measured by the probabilities of returning to the respective states in the following quarter, are very similar. Indeed, the volatility of stocks remains contained when investors perceive moderately high growth and mainly increase their probabilities of the LI-HG state. The increases in volatility following strong growth in earnings growth in the early 1960s and late 1990s explains why the leverage effect is fairly weak in our sample. This effect documents that volatility increases mainly following bad news. We will discuss this more completely in the option pricing results in Section 3. These episodes also help explain why the power of interest rates in forecasting volatility for our full sample has been quite limited, since they were not preceded by rising rates. Glosten, Jagannathan, and Runkle (1993) point out that interest rates are useful forecasters of stock market volatility in alternative subsamples of stock market data.

We next turn to the kernel parameter estimates, the fifth set of parameters in Table 1. We notice immediately that \( \alpha_0 \) is very close to zero, which implies that the real rate does not depend on the state
of real fundamentals, and that $\alpha_\beta$ is significantly positive, which means that the real rate is higher in states of higher inflation. Ceteris paribus, these estimates imply lower P/E ratios and higher Treasury yields during periods with higher expected inflation.\(^4\) These observations are confirmed in the bottom panel of Table 2, which reports the implicit parameters for the P/E and bond yields across the states using the pricing formulas for stocks and bonds in equations (12) and (14), respectively.

Using the time series of investors’ state probabilities in Figure 1 and the estimated parameters we generate time series of model-implied expected fundamental growth and prices in Figure 2. Our SMM procedure also uses the moments of assets volatilities. Table 3 shows the fits of our model to expected fundamental growth and the overidentifying price and volatility variables. The first two lines of the table show that the regression of historical on expected growth give an $R^2$ of 46.7% and 9.5% for inflation and earnings growth, respectively. We note that our SMM procedure, which maximizes the likelihood of investors observing the historical fundamental processes, does not have an explicit prediction on the fitted actual fundamentals in each period, but instead characterizes expected fundamental growth. Therefore, these fits reflect not simply the accuracy or our model, but in addition, investors’ estimates on the fraction of variation in fundamental growth that is related to shifts in trend growth rates as opposed to purely idiosyncratic variation. Also note that the $\beta$ coefficients in both expectations regressions are in excess of two, so that actual fundamentals are more than twice as volatile as their expectations.

In contrast to fundamental growth rates, the overidentifying price and volatility moments used in the estimation lead to explicit predictions of model-implied prices and volatilities in each period and are functions of investors beliefs of the states of the fundamentals. Indeed, our model fits the prices fairly accurately: the fit for the P/E ratio is about 72.7%, although notably, the model fails to fully fit the high valuations of the late 1990s, as investors were not fully convinced of the new economy growth state. In particular, the model correctly fits a P/E in the high teens for most of the 1960s and low teens for much of the 1970s and early 1980s. The fits for historical yield series have $R^2$s of between 61% and 67%, with the better fits for the longer maturities. In strong support of our model, the $\alpha$ coefficient in each of the regressions for the financial variables is not significantly different from zero, and the $\beta$ coefficients of each financial variable are significant at the 1% level and close to one. The model fits

\(^4\)Some authors have called this inverse relationship between P/Es and Treasury yields the “Fed Model,” a term that has been attributed to Prudential Securities strategist Ed Yardeni. The relationship seems to hold in several countries across varying time samples [see, for example Aubert and Giot (2007)].
are not perfect though: the largest errors occur for the shorter maturity bonds following recessions. As seen in the figure, after each of the past two recessions, the Federal Reserve effectively lowered short-term rates dramatically to levels that cannot be justified by our purely fundamental-based model. The pricing errors decrease in the maturity of the bonds, as long bond yields did not decline as much in these periods.

The final components of our SMM error term are the moments based on option prices that we discuss separately in Section 3. Using the scores of the likelihood function and the errors of the price and volatility variables, we evaluate the SMM objective function, which serves as an omnibus test statistic [see for example Gray (1996) and Bansal and Zhou (2002)]. The overall SMM objective function value, which has a chi-squared distribution with 7 degrees of freedom, is 11.34, implying a $p$-value larger than 12%, so we fail to reject our model.

3 Time Variation in the Option Implied Fear Indices Explained by the Model

In this section we examine the ability of our model to explain the time variation in the two fear index indices. We also consider several control variables that have been shown in the literature to be successful explanatory variables for these indices.

3.1 Explaining Time-Variation in ATM Implied Volatility

In this subsection we study the ability of our model to explain the time-series variation in the ATM implied volatility (ATMIV). Our basic ATMIV measure is constructed by interpolating the implied volatilities of call and put prices that are closest to being at-the-money. The bottom panel of Table 2 shows the ATMIV when investors are 90 percent sure of being in each of the five states respectively. Volatility is high (at about 33 percent) when investors assess a high (90%) chance of being in the MI-LG and LI-NG states. In the former state the high volatility results from the large uncertainty of the economy slipping into the stagflation (HI-LG) state, when P/Es decline further. This is the signaling channel of inflation that investors take heed to during most periods of rising inflation. ATMIV is also high in the LI-NG state due to the high earnings uncertainty as explained in the previous section. In
the stagflation state, volatility is at 22 percent, which is a little lower that in the MI-LG state, because investors are looking ahead to better times. Finally, volatility is low (averages 12 percent) in the two HG states, when earnings growth is the most stable. Somewhat surprisingly it is lowest in the MI-HG state since its impossible to transition to the new economy state directly from this state. It must be noted that the stock volatility in our model given in (16) is a concave function of these beliefs, and increases from these above numbers when there is greater uncertainty about the states. Therefore, the ATMIV is also ceteris paribus, higher in periods of higher uncertainty.

The historical and model-fitted series are shown in the top panel of Figure 4 and some regressions examining the fits are in Table 4. Our historical time series spans a long period of 20 years that covers the recessions of 1991 and 2001 as well unusual events such as the stock market crash of 1987, the collapse of LTCM in 1998, and the bursting of the technology bubble in 2000. As seen in the figure, during each of these events implied volatility increased above 35 percent, while its average over the sample is 19.2 percent. Our model shows that macroeconomic uncertainty could account for more than 75 percent of the volatility during each of these episodes with the exception of the third quarter of 1991 when upon the announcement of the Iraq invasion, implied volatility hit 30 percent while the model implied volatility is only 22 percent. The model captures well the decline in volatility from 1991 to 1996 when earnings growth resumed at a strong pace and investors’ belief of the normal recession (MI-LG) state declined while that of the normal boom (LI-HG) state trended upwards (Figure 1). Throughout the decade, until about 1998, the investors’ assessed likelihood of the new economy growth state slowly increased, and from about 1996; the increasing earnings uncertainty from this state offset the declining volatility from recession fears, so that uncertainty and ATM volatility started rising quite sharply. The high volatility was sustained throughout the late 1990s, but it increased quite sharply even further to its highest level in our sample in 2001 when there was a simultaneous increase in investors’ probability of the normal recession state from nearly 0 in the first quarter of 2000 to 5 percent, and a decrease in the probability of the new economy growth rate state from about 30 to 20 percent, both of which led to an increase in earnings uncertainty. The model is also very successful in explaining the subsequent decline in implied volatility until 2006. As seen in Figure 1 the major change in investors’ beliefs was a simultaneous decline in the probability of a normal expansion (LI-HG) state and an increase in the (MI-HG) state, which is the state with lowest implied volatility. However,
the model does not match the extremely low volatility of 12 percent in the first half of 2006, instead implying a volatility of about 17 percent in this period.

We next look in Table 4 at some formal statistics of the fit our model for the data ATMIV series. Line 1 shows that the $R^2$ of the simple regression of the data ATMIV on the model ATMIV for the full sample is 50 percent, and the beta coefficient is 0.989, which is very close to 1. Excluding the fourth quarter of 1987 from the regression increases the $R^2$ by another 5 percentage points (not reported). Line 2 reports the regression of the data ATMIV series on the model’s fitted stock volatility series. At time $t$ the stock volatility is given by (16) and evaluated using investors’ beliefs in Figure 1. The $R^2$ for the regression is nearly 47 percent, which only a little lower than our model ATMIV. The model ATMIV fits better since it additionally incorporates the impact of higher moments on implied volatility. In addition, we note that the intercept in the regression of 5 percentage points is highly statistically significant, and the $\beta$ coefficient is smaller than 1. The intercept is related to the model’s volatility premium that we will discuss separately in Section 4.3.

We next report the results of a set of control variables that also help explain the data ATMIV. Line 3 reports the fit from the S&P 500 price-earnings ratio. As seen the beta coefficient is positive and the $R^2$ is 22 percent. This P/E effect is particularly strong in our 20-year sample, in which one of the largest episodes of stock market in the late 1990s coincided with high P/Es. As explained in the previous section, due to the small probability of the new economy state, increases in earnings in the 1990s had a positive impact on P/Es and volatility, while in most other periods of high uncertainty in our full sample, high volatility accompanies low P/Es. Line 4 reports that the short rate does not explain ATMIV in our sample. Glosten, Jagannathan, and Runkle (1993) point out that interest rates are useful forecasters of stock market volatility in alternative subsamples of stock market data. The power of interest rates to explain the ATMIV in our sample is limited due to the unusually large episode of volatility associated with the new economy growth that was not preceded by rising rates. Line 5 shows that the NBER expansion indicator has a negative sign but explains only 4 percent of the variation in the ATMIV. In longer samples, Schwert (1989) reports that the expansion indicator explains a large part of contemporaneous volatility. Finally, line 6 shows that the leverage effect, which is the lagged S&P 500 return in periods when it is negative, is highly significant, and explains almost 33 percent
of the variation in the ATMIV. These above control variables encompass almost all existing work on volatility.

We next consider the effects of lags, and the effects of combining the controls with our model ATMIV. We note that the practice of including lags in volatility equations has a long traditions starting with the seminal work of Engle (1982) and extended with various specifications in the huge GARCH volatility literature. We investigate further the economic determinants of the significance of the lag and the extent to which our model implied volatility can capture all the economic information. Line 7 shows that the lag of ATMIV is highly significant and explains about 50 percent of the variation in ATMIV, about the same amount as our model ATMIV. The regression tells us that there are persistent economic forces that explain ATMIV. It however, does not identify these economic forces. Our modeling in contrast explicitly shows the effects of macroeconomic uncertainty on the ATMIV. In line 8, we combine the lag and our model ATMIV, and find that both remain highly significant, although the combined explanatory power increases only by a small amount, to 57 percent. This tells us that the lag and our model variable share a large amount of common explanatory power, which is persistent macroeconomic uncertainty. In addition, to macroeconomic uncertainty, the lag likely picks up effects of trading and its disruptions, which are likely sources of volatility [see e.g. Roll (1984)]. Finally, line 8 combines the lag, our model ATMIV, and all our controls, and find very little increase in the explanatory power, with the lag and our model retaining their significance, while the other controls all turn out to be insignificant. Therefore, for our sample, the effects of all the above control variables are picked up by our model. David and Veronesi (2008) look at the performance of our model and these controls over a longer sample of nearly 50 years and longer forecasting horizons.

3.2 Explaining Time-Variation in the Put-Call Ratio

In this section we study the model’s ability to explain the put-call ratio (P/C), which we define to be the ratio of implied volatilities of 5 percent out-of-the money put and call options. We obtain the implied volatilities of exact strikes by interpolating the implied volatilities of options with nearby strikes. The P/C is a natural “fear index” since it provides a market-based measurement of investors’

---

5To ensure intertemporal consistency of put-call ratios we follow Bates (1991) and set $K_{\text{put}} = S_te^{(r-\delta)t}/1.05$ and $K_{\text{call}} = S_te^{(r-\delta)t} \times 1.05$.  

24
assessed likelihood of prices falling as opposed to rising to the maturity of the option. It is useful to keep in mind that the Black-Scholes model implies that the P/C is identically 1.

We first look at Figure 3 that shows the densities of stock returns when investors are 90 percent certain of being in each of the five states (the remaining states each have equal probability). The top panel shows the densities for the states 1, 3, and 5, while the bottom panel shows the densities for states 2 and 4. Recall that states 1 and 3 have normal high growth of earnings, while state 5 has the new economy growth. In each these states the density is negatively skewed, while for the low growth states, the densities are positively skewed. These shapes naturally imply that the P/C is larger than 1 for the former states and smaller than 1 for the latter ones. In each of the higher growth states, investors fear that there is greater downside risk of slowing growth rate growth rates through the business cycle that lower P/E ratios and lead to negative returns, hence raising the P/C. The opposite holds true for the lower growth states. Apart from these sign differences, the magnitudes of the P/C are also quite different in the two sets of states, and our comments below also allude to the transition probabilities in the top panel of Table 2. The top panel the figure shows that the greatest negative skewness is in the MI-HG state, from which there is a significant (2.3 percent) chance of moving to the MI-LG state in a quarter, with no significant upside. From the regular expansion state of LI-HG, there is a smaller probability of transition directly to a lower P/E state, but also a small probability of transitioning to the new economy growth rate state with an upside. Therefore, the P/C is highest in the MI-HG state at 3.46, while its still high at 2.14 in the LI-HG state. In the new economy growth rate state, there is no upside at all but since the state is very stable there is smaller downside too than the other states, so that the P/C is lower at 1.48. Notice from the filtered probabilities, that investors were never 90 percent certain of any of the states in our sample, so that the model P/Cs were never as high as as in these states. However, the model P/C is roughly linear in these probabilities so that they provide us intuition for the model based series.

The bottom panel of Figure 4 shows the data and model P/Cs. As can be seen both series are almost always larger than 1 (the former falls below 1 only twice). In short the data P/C trended up to the recession of 1991, fell dramatically, and then rose sharply upwards to 1996, from where it began a steep decline until the 2001 recession. Finally, it trended upwards from 2001 to the end of our sample in 2006. The fall in the P/C during the two recessions also occurs in our model since the probabilities
of the two low growth states increase in these periods, and as we pointed out above, this tends to lower model P/Cs. The model P/C rises from around 1.2 in the 1991 recession to nearly 1.7 in 1996 caused largely by the decline in investors’ recession fears (probability of being in state 2). From 1996 to 1998, investors probability of being in the MI-HG state diminished rapidly as inflation was held in check, and rapid growth of earnings increased their beliefs of being in the new economy state, thus leading to a decline in the model P/C as discussed in the previous paragraph. From 1998 to 2001, the probability of the new economy state declined but the reemerging fear of the recession (MI-LG) state thus lowered the model P/C further. From 2001 to 2006, we see that both data and model P/Cs slowly trended upwards, and Figure 1 shows that the major trend during this period is the steady increase in the probability of the MI-HG state and the decrease in the probability of the LI-HG state, which as discussed above raises the model P/C. In summary, the model P/C tracks fairly well the ups and downs in the data P/C over the 20 year sample, and we present some regression results to confirm this finding next.

Line 1 of Table 5 provides the simple regressions of the data P/C on the model P/C. As seen, the fit is very solid with a statistically insignificant alpha coefficient, a beta of 0.846 (not different statistically from 1) and an $R^2$ of 47.4 percent. As for the ATMIV regression we report the regression with the lagged P/C for comparison. The lag does pick up a similar $R^2$ of 45 percent. We compare the two jointly below with controls. We next see if our model P/C can be explained by two sentiment measures used by Han (2008). The first, which we refer to as “trader sentiment” is the net long position of large speculators on S&P 500 index futures obtained from the Commodity Futures Trading Commission’s Commitment of Traders Report. The second, “investor sentiment” is the bull-bear spread (proportion of traders bullish less bearish) in Investor’s Intelligence’s survey of investment newsletter writers. These measures are alternative measures of fear in the market and are thus compelling control variables for our measures of downside risk obtained from asset prices. While Han (2008) suggests that the significance of these measures supports a behavioral view of asset prices, we note that they could be consistent with a rational model of heterogeneous learning about the states of fundamentals such as in David (2008).

If these sentiment measures reflect investors fear of negative returns on the S&P 500 index, then in the regression coefficient of the data P/C on these measures should give a negative beta coefficient. Line 2 shows that the trader sentiment measure does have a beta coefficient that is negative, but it
is only weakly significant at the 10 percent level. Its explanatory power for the P/C is also low at only 8 percent. Han (2008) provides a regression that includes the lagged P/C and finds a statistically significant coefficient on a shorter sample from 1988 to 1997:Q2. On our full sample, line 3 shows that once the lag is included, the trader sentiment measure is statistically insignificant. We do obtain a significant regression on the shorter sample. Line 5 includes the lagged P/C, our model P/C and the trader sentiment, and we find that both our model and the lag remain significant, and the explanatory power increased to 56.5 percent. Therefore the lag does seem to have some incremental information than captured by our model, but its not the trader sentiment. We run analogous regressions with the investor sentiment, and find similar results. The investor sentiment on its own is more highly significant, and explains about 13 percent of the P/C, but is not significant in the presence of either the lagged P/C or our model. Overall, the variables suggested by Han (2008) are significant for explaining the P/C explanatory power is quite low and they do not provide additional information on the P/C ratio than our fundamental learning model. We note, that the exercise of fitting the parameters of a fundamental model with heterogeneous beliefs to match options prices will undoubtedly do better than our model, but that exercise is not performed in this paper.

4 Additional Properties of Option Prices

In this section we discuss features of observed option prices that are not directly fitted by our empirical methodology. The ability of our model to replicate these additional features provides further support for the economic mechanism that determines option prices in our model.

4.1 The Volatility of ATM Volatility

In the previous section we saw that our model ATMIV was able to explain about 50 percent of the variation in the data ATMIV. In the model, the implied volatility is to a large part determined by the endogenous volatility of stock prices, which increases during periods of inflation and earnings uncertainty. In addition, looking again at the top panel of Figure 4, we see that during episodes of high volatility such as in 1987, 1998, and 2001, ATMIV also fluctuated by its largest amounts in our sample. The positive relation between volatility and the volatility of volatility is noted in Jones (2003) who further notes that it cannot arise in the Heston (1993) stochastic volatility model, which has
been the workhorse of the option pricing literature. To obtain the level dependence of volatility, Jones (2003) proposes a generalization of the Constant Elasticity of Variance (CEV) model of Chan, Karolyi, Longstaff, and Sanders (1992). One drawback of the volatility processes in such models as is that they do not satisfy global growth and Lipschitz conditions, which are commonly used sufficient statistics for a number of important results. In this subsection we investigate if our model is able to shed light on the positive association between volatility and changes in volatility. Indeed as can be seen in (7), the Bayesian learning mechanism that drives volatility in our model implies that investors revise their beliefs faster during periods of high uncertainty as they have low confidence in their estimates of the current state of the fundamentals. David (1997) establishes that the belief processes do satisfy global growth and Lipschitz processes and it is straightforward to extend the result to stock volatility in this model.

The top panels of Figure 5 show the scatter plots of implied volatility and absolute changes in implied volatility for the data and model series. Both show a positive association of similar magnitude between these variables with correlations of 42 percent and 30 percent, respectively. We next check if the absolute changes in implied volatility are related to the volatility of stock variance in our model. To do this we construct a time series of the model’s implied volatility of variance using (20) and evaluate it at each date using the filtered beliefs in Figure 1. The scatter plot of absolute changes in implied volatility (data and model) with this series are shown in the middle panel. Note that the model series measures the ex-ante volatility of variance at each date and is compared to the ex-post realized changes in ATMIV, and is thus an expectation of the latter series. As seen, the model series is highly correlated with both the data and model absolute changes with correlations of 43 and 37 percent respectively.

Finally, to firm up our intuition about the learning mechanism we check if the absolute changes in ATMIV are related to the two fundamental uncertainties (again ex-ante concepts) in our model. We define investors’ conditional uncertainty about earnings growth (and analogously for inflation) at any time as

\[
RMSE_E(t) = \left( \sum_{i=1}^{N} \pi_i(t) \left( \theta_i - \bar{\theta}(t) \right)^2 \right)^{1/2}.
\]

(24)
The scatter plots are shown in the bottom panels of Figure 5, and the correlations with earnings and inflation uncertainty are 43 percent and 28 percent, respectively. Thus investors’ fundamental uncertainties are both significant drivers of volatility changes. Relatedly, Beber and Brandt (2008) show that volatility in stock and bond markets declines faster following periods of high macroeconomic uncertainty extracted from the economic derivative markets over a shorter sample from 2002-2005.

4.2 The Relationship Between Price-Earnings and Put-Call Ratios

Should we expect a relationship between P/E ratios and P/Cs? Suppose, P/E ratios tend to mean revert. Then, in periods of a high P/E, investors should expect a decline in future P/Es and keeping earnings fixed, should anticipate downside risk in future stock returns that would raise P/Cs. However, since earnings are not fixed, the impact of future returns is not clear since the high P/E may forecast future earnings and raise the prospect of upside risk in returns, which would lower the current P/C. Even holding earnings fixed, the tendency of P/Es to mean-revert over the next quarter (the maturity of the options) may depend on the state of macroeconomic fundamentals so that the impact of current valuations on P/Cs can be state dependent, making the relationship more complex. In this subsection we see that the sign of the relationship between P/Es and P/Cs has varied over time, and our model is able to shed light on this time variation.

Figure 6 shows scatter plots of realized P/Es and P/Cs over the 20 year sample. The left panels show the relationship in the data and the right from the model generated ratios. The correlation of the two variables in the data is -0.3, while in the model it is -0.24. However, a closer look at the two plots shows that there are points of positive correlation between the variables as well. In the middle panel, we make the same scatter plots only over the initial part of the sample from 1986:Q2 to 1993:Q4, and in these panels we find the relationship is positive. The correlation between the two variables is 0.38 in the data, and a stronger 0.76 for our model. Moving to the bottom panels, which cover the period from 1994:Q1 to 2006:Q2, we find the relationships are strongly negative, with correlations of -0.6 and -0.84 in the data and our model respectively. Therefore, our model does well in picking up the turning point in the sign of the relationship between P/Es and P/Cs.

We shed some light on the sign reversal by looking again at the filtered beliefs of investors in Figure 1 and the calibrated P/Es and P/Cs in the five states in the bottom panel of Table 2. In the first part
of the sample, the dominant movement in beliefs was a the shifts between the two medium inflation (states 2 and 3) and the two low inflation states. (states 1 and 5). Coming into the 1990 recession there was an increase in the probability of the second state from near zero to 15 percent, and of the third state from around 30 to 40 percent. At the same time there were dips in the probabilities of the two low inflation states. These changes reversed at the end of the recession. The two medium inflation states both sustain low P/E ratios due to the instability of earnings in these states. As discussed in the previous section, the P/C is highest is at its lowest point in the MI-LG state, while its at its highest in the MI-HG states. Since the rise in the probability of the second state is higher, the net effect is to lower the P/C in these periods, so that P/Es and P/Cs have a positive comovement. In the second subsample, which includes the 2001 recession, the dominant movements were between the third (MI-HG) and (LI-NG) states. Throughout the second half of the 1990s, the probability of the MI-HG state declined and that of the new economy state increased as inflation was controlled and the boom in technology companies boosted P/Es. In the first half of the current decade, these trends reversed although not as strongly as we had saw them take hold in the 1990s, and this milder reversal helped explain the slow decline in P/Es and increase in P/Cs during this period.

4.3 The Volatility Premium

The ‘volatility premium’ is a measure of the difference in volatility forecasts under the risk-neutral (Q) and objective (P) measures. If volatility is systematically positively related to investors’ pricing kernel (marginal utility of consumption), then as a priced factor it carries a negative risk premium that leads to a higher forecast of volatility under the Q measure or a positive volatility premium.

The empirical finance literature now has more than one operational definition of this quantity. The first, which we call the implied volatility premium (IVP) is defined as the difference between at-the-money implied volatility and a forecast of future volatility to the maturity of the option under the objective (P) measure. The forecast under the P measure is constructed for specific volatility models. A second definition, which we call the forecast volatility risk premium (FVRP), simply takes the difference in forecasts of future volatility under the two measures using the different dynamics under the two
measures. The two different forecasts are formed with the same structural model and differ only by the difference in volatility drift, which is determined by the specification of the pricing kernel process.\footnote{In addition, Bollerslev and Zhou (2007) use a measure of an ex-post volatility premium that takes the difference between implied volatility and realized volatility to predict future stock returns.}

In this subsection we evaluate the ability of our model to shed light on the time variation of the IVP process. Thus our focus is different from the early papers on this topic that were focused around forecasting volatility using implied volatility and the implied volatility premium. Notable among these papers are Canina and Figlewski (1993) and Christensen and Prabhala (1998) who find that implied volatility is a useful albeit biased forecaster of future realized volatility. In more recent work, Drechsler and Yaron (2008) and Eraker (2008) construct equilibrium models with ‘long term risks’ to understand the size of the IVP and some of its unconditional moments. However, neither of these papers explicitly study the ability of their model to match the historical time series variation of the IVP, and thus the conditional predictions of their models are largely untested. We contribute to this literature by studying explicitly the properties of our model and data IVP historical processes.

In discussing the cause of the forecasting bias, Christensen and Prabhala (1998) note that while the Black-Scholes implied volatility can be thought of as a volatility forecast, it can also be interpreted as a measure of the option’s price. In addition to the level of volatility forecast under the $Q$ measure, the option price is determined by the higher moments of the forecasted return distribution. Therefore, the interpretation of implied volatility as a forecast of future volatility holds only under the assumption of Gaussian stock returns. This implies that a model with non-Gaussian returns may well have a small FVRP but an IVP that matches historical experience.

To construct a data based ex-ante IVP series we need forecasts of realized volatility, which we discuss first. To construct a ‘data’ forecast, we appeal to well established results in the volatility forecasting literature literature, which we discussed in Section 3.1. In particular, for our sample from 1986:Q2 to 2006:Q2, we run a regression of realized volatility on its one-quarter lag, the lagged P/E ratio, and lagged returns on the S&P 500 index in periods when they are negative. The results of this regression are:

$$
\text{Vol}(t + 1) = 0.034 + 0.357 \text{Vol}(t) + 0.321 P/E(t) - 0.363 \text{Ret}^-(t); \quad R^2 = 0.195
$$

\begin{align*}
| & 1.123 & | 3.435 & | 3.030 & | -2.004 \\
\end{align*}

\[ (25) \]
where $\text{Vol}(t + 1)$ is the volatility realized in quarter $t + 1$, which we define as the square root of the sum of squared S&P 500 index returns in the quarter, $P/E(t)$ is the S&P 500 price-to-operating earnings ratio, and $\text{Ret}^-(t)$ is the return on the S&P 500 index in periods when it is negative. T-Statistics are in parenthesis and are adjusted for heteroskedasticity and autocorrelation using the Newey and West (1987) method. The regression $R^2$ improves to 41 percent for the post-crash subsample starting in 1988:Q2.\footnote{Drechsler and Yaron (2008) use a similar regression equation using a post-crash sample (and hence report a higher $R^2$) in constructing their measure of the volatility premium. They do not use the P/E ratio and lagged returns in their forecast, but include lagged implied volatility. For our sample, their volatility forecast has a 90% correlation with our forecast and makes little difference to any of our results.} It is also important to note that we use non-overlapping data by constructing one-quarter ahead volatility forecasts at the quarterly frequency so that the t-statistics are more reliable. The realized volatility series and its fitted value from the forecast in (25) are shown in the top panel of Figure 7.

We construct our model-based volatility forecast of one-quarter ahead realized volatility under the P measure as shown in Appendix 2. A regression of realized volatility on our lagged model forecasted volatility gives the following result:

\[
\text{Vol}(t + 1) = -0.007 + 0.883 \hat{\text{Vol}}(t + 1|t); \quad R^2 = 0.338
\]

\[[-2.040] \quad [3.859],\]

where $\hat{\text{Vol}}(t + 1|t)$ is the optimal forecast of volatility from our model. The realized volatility series and our model based forecast are shown in the middle panel of Figure 7. For the full sample our model forecast explains 33.8% of the variation in realized volatility, while for the sample restricted to the post-crash period it climbs to 45%. As seen, our model forecast equation does not include lags, and the forecast is efficient in the sense that the intercept is statistically insignificantly different from 0 and the slope is insignificantly different from 1 at the 1% significance level. The realized volatility series and the optimal model forecast are shown in the middle panel of Figure 7.

The data IVP is the difference between the data ATM series and the fitted value from the regression in (25). We similarly construct a model based IVP series by taking the difference between the model implied volatility analyzed in Section 3.1 and the model forecast in (26). The data and model IVP series are shown in the bottom panel of Figure 7. The means of the data and model series are very similar at 3.78% and 3.3% respectively, however the data series has a volatility of 5.04% which is greater than
the model volatility of 3.34%. Both data and model premiums are positive in most periods, with the proportion of positive data points being 80% and 91.2% respectively. These positive implied volatility premiums are consistent with the past studies cited earlier. In addition, both series are highly autocorrelated: the first order autocorrelations of the data and model series are 0.47 and 0.60 respectively. The relationship between the data and model volatility premiums is

\[
\text{Vol Prem}_D = 0.007 + 1.081 \text{Vol Prem}_M; \quad R^2 = 0.295. \tag{27}
\]

Thus the model volatility premium explains nearly 30% of the time variation in the data premium. In addition we fail to reject that the intercept is 0 and the slope is 1 with 95% confidence. We note that our estimation procedure in Section 3 does not include the volatility premium as one of the overidentifying moments, and hence this is an out-of-sample fit of our model.

We next examine if the IVP of our model is related to its FVRP. The Q forecast of our model is constructed using the same methodology as the P forecast and is shown in (46) in Appendix 2. The FVRP is simply the difference in the two forecasts. Constructing the series conditional on beliefs for our full sample we find that the FVRP is positive but small – of around 0.6%. We note that the projected belief dynamics under the two measures differ only in their drifts and the forecast horizon of one quarter is relatively short so that the drift effect is small. Thus the IVP is largely unrelated to volatility being a priced risk factor in our model.

The small FVRP in our model is also true of the Heston (1993) stochastic volatility (SV) model. Chernov and Ghysels (2000) is one of the few papers that calibrate the parameters of the Heston model under both objective and risk-neutral measures. We use their parameter values to calculate the variance premium for our sample in this model. For this model we have exact formulas for variance forecasts.
under each measure.\footnote{Chernov and Ghysels (2000) estimate the model}

To compute the variance premium for the SV model we use the annualized realized variance on the S&P 500 index proxied as the sum of squared daily returns for each day in the quarter. We then use forecasts using (30) using the two different sets of parameters and compute the difference. For the sample between 1986:Q2 to 2006:Q2, the same set of dates for which we have options data, we find an annualized premium of -0.27%.

What explains the model’s implied volatility premium?

\[
\text{Vol Prem}^M = 0.058 + 0.915 \log(\text{Vol. of Vol.}^M); \quad \bar{R}^2 = 0.328. \tag{31}
\]

The volatility of volatility is a measure of the fourth moment of asset returns. The result tells us that implied volatility is higher during periods of a higher fourth moment. This supports the view that the volatility premium results from the non-normality of returns that makes the values of options higher than that of the Black-Scholes model with the same volatility forecast. The non-normality of returns in our pure diffusion model arises endogenously (under either measure) as investors dynamically learn about the state of fundamentals and adjust their conditional mean and volatility of future returns, which over any finite horizon is then formed as a mixture of distributions.

5 Conclusion

Option prices provide key forward looking information on investors’ expectations, and market attention is often focused on two key fear indices, the at-the-money implied volatility, and the ratio of implied

\[
\begin{align*}
\frac{dS(t)}{S(t)} &= R dt + \sqrt{V(t)} dW_S(t) \quad \tag{28} \\
\frac{dV(t)}{} &= (\theta - \kappa V(t)) dt + \sigma_V \sqrt{V(t)} dW_V + \sigma_V \rho \sqrt{V(t)} dW_S. \quad \tag{29}
\end{align*}
\]

For this volatility process the annualized $\tau$ period ahead forecast of future variance under the objective measure conditional on on observing variance $V(t)$ at time $t$ is

\[
\hat{V}(t, t + \tau) = \frac{\theta}{\kappa} + \frac{1}{\kappa \tau} (1 - e^{-\kappa \tau}) \left( V(t) - \frac{\theta}{\kappa} \right). \tag{30}
\]

Under the risk-neutral measure $\theta$ and $\kappa$ are replaced by $\theta^*$ and $\kappa^*$, respectively. Chernov and Ghysels (2000) report the following parameter values using data both on the stock price and options from 1985 to 1993 in Table 2: $\theta = 0.014$, $\theta^* = 0.007$, $\kappa = 0.931$, $\kappa^* = 0.690$. 

34
volatilities of out-of-the-money puts and calls. The former is measure of market turbulence, while the latter is a measure of downside risk. Standard option pricing models use exogenous stock prices and their volatilities that are unrelated to fundamentals, and are hence unable to identify specific economic factors that can explain these fears. We provide a model in which stock, bond, and option prices, are functions of investors’ beliefs of the state of fundamentals. Fear of the economy going into stagflation causes volatility during recessions, while the belief of a new economy growth state in the late 1990s led to an atypical bout of volatility in this period that accompanied strong fundamentals. Put-call ratios are higher in periods of stronger earnings growth, and among these, are highest during periods of medium inflation, which are states with the lowest upside potential. Our model explains about half the time series variation in the two fear measures and is able to reflect all the information in sentiment indices and variables shown to be significant for forecasting volatility in the empirical finance literature.

As further support for the learning mechanism we demonstrate the ability of our model to explain (i) the positive relation between volatility and the volatility of volatility, (ii) the puzzling change in sign of the relation between P/E and put-call ratios, (iii) a significant amount of variation in the implied volatility premium. Properties (i) and (iii) are shown to be directly related to investors’ uncertainty about fundamentals, while (ii) relies on their assessment of the stability of earnings growth. Overall, we find strong evidence for a learning based mechanism to explain investors fears embedded in options prices.
References


Appendix 1

For proving Proposition 2 we will use the algebraic result stated in the following lemma.

Lemma 2

\[ \frac{\partial \bar{\theta}}{\partial \pi_i} = C_i \left( \theta_i - \bar{\theta} \right) \left( \sum_j \pi_j C_{ij} \right). \]

Proof of Lemma 2:

\[
\frac{\partial \bar{\theta}}{\partial \pi_i} = \frac{\partial \left( \sum_j \pi_j C_{ij} \theta_j \right)}{\partial \pi_i} = \frac{C_i \theta_i \left( \sum_j \pi_j C_{ij} \right) - C_i \left( \sum_j \pi_j C_{ij} \theta_j \right)}{\left( \sum_j \pi_j C_{ij} \right)^2} = \frac{C_i \theta_i}{\left( \sum_j \pi_j C_{ij} \right)^2} - \frac{C_i \theta_i}{\left( \sum_j \pi_j C_{ij} \right)} - \frac{C_i \bar{\theta}^2}{\left( \sum_j \pi_j C_{ij} \right)} = \frac{C_i \left( \theta_i - \bar{\theta} \right)}{\left( \sum_j \pi_j C_{ij} \right)}.
\]

which completes the proof. \( \blacksquare \)

Proof of Proposition 3: Let the second term in the variance equation be \( V_2 = (\bar{\nu}^\circ - \bar{\nu})' (\Sigma \Sigma')^{-1} (\bar{\nu}^\circ - \bar{\nu}) \). Then, using Lemma 2 on each element of the drift vector \( \nu \) we have

\[
\frac{\partial V_2}{\partial \pi_i} = 2 \left[ \frac{C_i (\nu_i - \bar{\nu}^\circ)}{\sum_j \pi_j C_{ij}} - \nu_i \right]' (\Sigma \Sigma')^{-1} (\bar{\nu}^\circ - \bar{\nu}).
\]

Then, using the volatilities of the beliefs process in equation (8), we have \( dV_2 = \mu_{V,2} dt + \sigma_{V,2} \), where

\[
\sigma_{V,2} = \sum_i \frac{\partial V_2}{\partial \pi_i} \sigma_i = 2 \sum_i \pi_i \left[ \frac{C_i (\nu_i - \bar{\nu}^\circ)}{\sum_j \pi_j C_{ij}} - \nu_i \right]' (\Sigma \Sigma')^{-1} (\bar{\nu}^\circ - \bar{\nu}) (\nu_i - \bar{\nu})' (\Sigma \Sigma')^{-1}.
\]
Similarly, let the third term in the variance equation be \( V_3 = 2[(\bar{\sigma} - \bar{\theta}) + (\bar{\beta} - \bar{\beta})] \). Then we have \( dV_3 = \mu_{V,3}dt + \sigma_{V,3} \), where

\[
\sigma_{V,3} = \sum_i \frac{\partial V_3}{\partial \pi_i} \sigma_i = 2 \sum_i \frac{\partial[(\bar{\sigma} - \bar{\theta}) + (\bar{\beta} - \bar{\beta})]}{\partial \pi_i} \sigma_i
\]

\[
= 2 \sum_i \pi_i \left( \left[ \frac{C_i(\theta_i - \bar{\theta})}{\sum_j \pi_j C_j} - \theta_i \right] + \left[ \frac{C_i(\beta_i - \bar{\beta})}{\sum_j \pi_j C_j} - \beta_i \right] \right) (\nu - \bar{\nu})' \Sigma^{-1}
\]

\[
= 2 \left[ (\sigma_{\theta_j}^o - \sigma_{\theta_j})' + (\sigma_{\beta_j}^o - \sigma_{\beta_j})' \right] \Sigma^{-1}
\]

where the second equality follows from Lemma 2, the third the definition of \( \pi_i^o \), and the fourth from the fact that

\[
\sum_j \pi_j^o (\theta_j - \bar{\theta}) (\beta_j - \bar{\beta}) = \sum_j \pi_j^o \theta_j \beta_j - \bar{\theta} \bar{\beta} = \sigma_{\theta \beta}^o,
\]

and analogous terms for the other elements of \( \nu \). Now summing \( \sigma_{V,2} \) and \( \sigma_{V,3} \) provides the statement of (b).

**Appendix 2**

1. **SMM Estimation of the Regime Switching Model**

   We provide here the details of the SMM estimation procedure. Since fundamentals are stationary in growth rates, we start by defining logs of variables: \( y_t = \log(Y_t) \), \( s_t = \log(S_t) \), and \( m_t = \log(M_t) \). Using (10), (3), and (4) we can write

\[
dy_t = (\bar{\theta}(\pi_t) - \frac{1}{2} (\sigma\sigma'_{Q\sigma}) dt + \Sigma_2 d\bar{W}_t,
\]

\[
ds_t = (\bar{\theta}(\pi_t) - \frac{1}{2} \sigma\sigma'_{E\sigma}) dt + \Sigma_1 d\bar{W}_t,
\]

\[
dm_t = (-\bar{\theta}(\pi_t) - \frac{1}{2} \sigma\sigma'_{M\sigma}) dt + \Sigma_3 d\bar{W}_t.
\]

It is immediate that investors’ beliefs \( \pi_t \) completely capture the state of the system \((y_t, s_t, m_t)\) for forecasting future growth rates. The specification of the system is completed with the belief dynamics in (7).

The econometrician has data series \( \{y_{t1}, y_{t2}, \ldots, y_{tk}\} \). Let \( \Psi \) be the set of parameters of the model. We start by specifying the likelihood function over data on fundamentals observed discretely using the procedure in the SML methodology of Brandt and Santa-Clara (2002). See also Duffie and Singleton (1993). Adapting their notation, let

\[
\mathcal{L}(\Psi) = p(y_{t1}, \ldots, y_{tK}; \Psi) = p(y_{t0}; \Psi) \prod_{k=1}^{K} p(y_{tk+1} - y_{tk}, t_{k+1} | y_{tk}, t_k; \Psi),
\]

where \( p(y_{tk+1} - y_{tk}, t_{k+1} | y_{tk}, t_k; \Psi) \) is the marginal density of fundamentals at time \( t_{k+1} \) conditional on investors’ beliefs at time \( t_k \). Since \( \{\pi_{tk}\} \) for \( k = 1, \ldots, K \) is not observed by the econometrician, we maximize

\[
E[\mathcal{L}(\Psi)] = \int \cdots \int \mathcal{L}(\Psi) f(\pi_{t1}, \pi_{t2}, \ldots, \pi_{tK}) d\pi_{t1}, d\pi_{t2}, \ldots, d\pi_{tK},
\]

where the expectation is over all continuous sample paths for the fundamentals, \( \bar{y}_{tk} \), such that \( \bar{y}_{tk} = y_{tk}, k = 1, \ldots, K \). In general, along each path, the sequence of beliefs \( \{\pi_{tk}\} \) will be different.

As a first step, we need to calculate \( p(y_{tk+1} - y_{tk}, t_{k+1} | y_{tk}, t_k; \Psi) \). Following Brandt and Santa-Clara (2002), we simulate paths of the state variables over smaller discrete units of time using the Euler discretization scheme
Shocks are approximated at roughly a daily frequency. The pricing kernel and beliefs along the entire path of the series of fundamentals. Each path is started with an initial belief, model-implied price-earning ratios and bond yields at the discrete data points overidentifying moments to the SML method above. From Proposition 1, we can compute the time series of often report where the generator matrix The Strong Law of Large Numbers (SLLN) implies that where is a 4 × 1 vector of standard normal variables, and h = 1/M is the discretization interval. The Euler scheme implies that the density of the 2 × 1 fundamental growth vector yt over h is bivariate (since ) normal.

We approximate with the density , which obtains when the state variables are discretized over M subintervals. Since the drift and volatility coefficients of the state variables in (7), and (32) to (34) are infinitely differentiable, and is positive definite, Lemma 1 in Brandt and Santa-Clara (2002) implies that as .

The Chapman-Kolmogorov equation implies that the density over the interval with M subintervals satisfies for quarterly data, so that shocks are approximated at roughly a daily frequency. The pricing kernel and beliefs along the entire path of the time series are obtained by iterating on (38) and (39). We approximate the expected likelihood as

where is the density approximated in (41). The SLLN implies that as . We often report which is the econometrician’s expectation of investors’ belief at . To extract investors’ beliefs from data on price levels and volatilities in addition to fundamentals we add overidentifying moments to the SML method above. From Proposition 1, we can compute the time series of model-implied price-earning ratios and bond yields at the discrete data points , as

where

\begin{align*}
\hat{y}_{t+h} - \hat{y}_t &= \left( \bar{\theta}(\pi_t) - \frac{1}{2}(\sigma_Q\sigma_Q', \sigma_E\sigma_E') \right) h + \Sigma_2 \sqrt{h} \hat{\epsilon}_t; \\
\hat{s}_{t+h} - \hat{s}_t &= \left( -\bar{\theta}(\pi_t) - \frac{1}{2} \sigma_S \sigma_S' \right) h + \sigma_S \sqrt{h} \hat{\epsilon}_t; \\
\hat{m}_{t+h} - \hat{m}_t &= \left( -\bar{k}(\pi_t) - \frac{1}{2} \sigma_M \sigma_M' \right) h + \sigma_M \sqrt{h} \hat{\epsilon}_t; \\
\pi_{t+h} - \pi_t &= \mu(\pi_t) h + \sigma(\pi_t) \sqrt{h} \hat{\epsilon}_t
\end{align*}

(36) (37) (38) (39)

where is a 4 × 1 vector of standard normal variables, and h = 1/M is the discretization interval. The Euler scheme implies that the density of the 2 × 1 fundamental growth vector yt over h is bivariate (since ) normal.

We approximate with the density , which obtains when the state variables are discretized over M subintervals. Since the drift and volatility coefficients of the state variables in (7), and (32) to (34) are infinitely differentiable, and is positive definite, Lemma 1 in Brandt and Santa-Clara (2002) implies that as .

The Chapman-Kolmogorov equation implies that the density over the interval with M subintervals satisfies for quarterly data, so that shocks are approximated at roughly a daily frequency. The pricing kernel and beliefs along the entire path of the time series are obtained by iterating on (38) and (39). We approximate the expected likelihood as

\begin{align*}
\hat{p}_M \left( y_{t_{k+1}} - y_{t_k}, t_{k+1} | \pi_{t_k}, t_k ; \Psi \right) &= \frac{1}{L} \sum_{l=1}^{L} \phi \left( y_{t_{k+1}} - y_{t_k}; \varphi(\pi_l), \Sigma_2 \Sigma_2' h; \Psi \right) = \int \int \phi \left( y_{t_{k+1}} - y; \varphi(\pi), \Sigma_2 \Sigma_2' h; \Psi \right) \times p_M \left( y - y_{t_k}, \pi, m, t_k + (M - 1)h | \pi_{t_k}, t_k \right) d\pi dy,
\end{align*}

(40)

where denotes a bivariate normal density. Now can be approximated by simulating L paths of the state variables in the interval and computing the average

\begin{align*}
\hat{p}_M \left( y_{t_{k+1}} - y_{t_k}, t_{k+1} | \pi_{t_k}, t_k ; \Psi \right) &= \frac{1}{L} \sum_{l=1}^{L} \phi \left( y_{t_{k+1}} - y_{t_k}; \varphi(\pi_l), \Sigma_2 \Sigma_2' h; \Psi \right).
\end{align*}

(41)

The Strong Law of Large Numbers (SLLN) implies that as .

To compute the expectation in (35), we simulate S paths of the system (36) to (39) “through” the full time series of fundamentals. Each path is started with an initial belief, , the stationary beliefs implied by the generator matrix . In each time interval ( , ) we simulate (M-1) successive values of using the discrete scheme in (36), and set . The results in the paper use for quarterly data, so that shocks are approximated at roughly a daily frequency. The pricing kernel and beliefs along the entire path of the simulation are obtained by iterating on (38) and (39). We approximate the expected likelihood as

\begin{align*}
\hat{L}^{(S)}(\Psi) &= \frac{1}{S} \sum_{S=1}^{S} \prod_{k=0}^{K-1} \hat{p}_M \left( y_{t_{k+1}}^{(s)} - y_{t_k}^{(s)}, t_{k+1} | \pi_{t_k}^{(s)}, t_k ; \Psi \right),
\end{align*}

(42)

where is the density approximated in (41). The SLLN implies that as . We often report which is the econometrician’s expectation of investors’ belief at . To extract investors’ beliefs from data on price levels and volatilities in addition to fundamentals we add overidentifying moments to the SML method above. From Proposition 1, we can compute the time series of model-implied price-earning ratios and bond yields at the discrete data points , as

\begin{align*}
\hat{P}E_{t_k} &= C \cdot \pi_{t_k}, \\
\hat{t}_{t_k}(\tau) &= \frac{-1}{\tau} \log \left( B(\tau) \cdot \pi_{t_k} \right).
\end{align*}
We note that the constants $C_s$ and the functions $B(\tau)$ both depend on the parameters of the fundamental processes, $\Psi$. Hence, we let the pricing errors be denoted
\[
e_{tk}^P = \left( \frac{P_{t+k}}{P_{t}} - \pi_{t+k} (0.25) \right) - \hat{\pi}_{tk} (0.25), \hat{\pi}_{tk} (1) - i_{tk} (1), \hat{\pi}_{tk} (5) - i_{tk} (5) \right).
\]
Also note that since the pricing formulas are linear in beliefs, $1/S \sum_{s=1}^{S} C \cdot \pi_{tk}^{(s)} = C \cdot \tilde{\pi}_{tk}$ (and similarly for the bond yields) and no information is lost by simply evaluating the errors at the econometrician’s conditional mean of beliefs. We similarly formulate the errors from options prices as
\[
e_{tk}^O = \left( \hat{V}_{tk} - V_{tk}, (P/C)_{tk} - (P/C)_{tk} \right),
\]
where $V$ is the at-the-money implied volatility of options with maturity closest to 90 days, and $P/C$ is the ratio of implied volatilities of 5% out-of-the-money puts and calls. The model-implied options prices are calculated using Monte-Carlo simulations as described below.

To estimate $\Psi$ from data on fundamentals as well as financial variables, we form the overidentified SMM objective function as in (23). The moments used are the scores of the log likelihood function from fundamentals, the pricing errors from financial variables, and their volatilities. Since the number of scores in $\frac{\partial \log(\ell_\Psi)}{\partial \Psi}$ equals the number of parameters driving the fundamental processes in $\Psi$, the number of pricing errors is four, and the number of option price errors is two, the statistic $c$ in (23) has a chi-squared distribution with six degrees of freedom. We correct the variance covariance matrix for autocorrelation and heteroskedasticity using the Newey-West method [see, for example, Hamilton (1994) equation 14.1.19] using a lag length of $q = 24$. A long lag length is chosen since interest rates and P/E ratios used in the error terms are highly persistent processes.

2. Options Prices

As for the likelihood function we formulate options prices as expected discounted values of their terminal payoffs under the risk-neutral measure. Expectations are approximated using Monte Carlo simulation while discretizing the dynamics of the state variables of our system along the $sth$ sample path under the risk-neutral measure as:
\[
\pi_{t+h}^{* (s)} - \pi_t^{* (s)} = \left( \mu(\pi_t^{* (s)}) - \rho(\pi_t^{* (s)}) \right) h + \sigma(\pi_t^{* (s)}) \sqrt{\tilde{\epsilon}_t^{**}}
\]
(43)
\[
P_{t+h}^{* (s)} = P_t^{* (s)} \exp\left[\left( r(\pi_t^{* (s)}) - \delta(\pi_t^{* (s)}) \right) h + \sigma N(\pi_t^{* (s)}) \sqrt{\tilde{\epsilon}_t^{**}} \right]
\]
(44)
\[
B_t^{* (s)} = B_t^{* (s)} \exp\left[ -r(\pi_t^{* (s)}) h \right],
\]
(45)
where $\tilde{\epsilon}^{**}$ is a $4 \times 1$ vector of standard normal variables, and $h = 1/M$ is the discretization interval. On each sample the process for the state variables is simulated starting with $\pi_t^{* (s)} = \pi_t$, the assumed beliefs of investors at time $t$. Then the value of a European call option at time $t$ when investors have beliefs $\pi_t$ that matures at $t + T$ is given by
\[
C_{t+T}^{* (s)}(t, T, \pi_t) = \frac{1}{S} \sum_{s=1}^{S} B_{t+T}^{* (s)} \max \left[ P_{t+T}^{* (s)} - K, 0 \right].
\]
We report option prices for $M = 90$. To reduce the time of computations we use three variance reduction techniques: the first two, antithetic and control variate (with Black-Scholes prices), are well known. In addition, we use the expected martingale simulation technique of Duan et. al. The volatility forecast under the Q-measure
is approximated from the path of forecasted beliefs under this measure as

\[
\sigma^{M^*}(t, T, \pi_t) = \sqrt{\frac{1}{S} \sum_{j=1}^{M} \sigma_N(\pi_{t+j}^{*}) \sigma_N(\pi_{t+j}^{*})' h}.
\]

(46)

Volatility forecasts under the objective measure are analogously constructed using the belief process simulated under the objective measure as in (39).
The table reports SMM estimates of the following model for CPI, $Q_t$, real earnings, $E_t$, earnings signals, $S_t$, and the real pricing kernel, $M_t$:

\[
\begin{align*}
\frac{dQ_t}{Q_t} &= \beta_1 dt + \sigma_Q dW_t, \\
\frac{dE_t}{E_t} &= \theta_1 dt + \sigma_E dW_t, \\
\frac{dM_t}{M_t} &= -k_t dt - \sigma_M dW_t, \\
\frac{dS_t}{S_t} &= \theta_2 dt + \sigma_S dW_t,
\end{align*}
\]

where $\sigma_Q = (\sigma_{Q,1}, \sigma_{Q,2}, 0)$, $\sigma_E = (0, \sigma_{E,2}, 0)$, $\sigma_M = (\sigma_{M,1}, \sigma_{M,2}, \sigma_{M,3})$, $\sigma_S = (\sigma_{S,1}, \sigma_{S,2}, \sigma_{S,3}, \sigma_{S,4})$, $k_t = \alpha_0 + \alpha_\theta \theta_t + \alpha_\beta \beta_t$, and the vector $\nu_t = (\beta_t, \theta_t, -k_t, \theta_t)^T$ follows a five-state regime switching model with the generator matrix $\Lambda$ whose non-zero non-diagonal elements are shown as $\lambda_{i,j}$. The pricing kernel, $M_t$, and the signal $S_t$, is observed by investors but not by the econometrician. Assets are priced using the formulas in Proposition 1. Estimates are obtained from data on the fundamentals as well six price levels and three price volatilities using the SMM methodology described in Appendix 2. Standard errors are in parentheses.
Table 2: Model Implied Transition Probabilities, Investor Expectations, and Prices

<table>
<thead>
<tr>
<th>State</th>
<th>Implied Quarterly Transition Probability Matrix</th>
<th>Implied Five-Year Transition Probability Matrix</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>(LI-HG)</td>
<td>(MI-LG)</td>
</tr>
<tr>
<td>(LI-HG)</td>
<td>0.991</td>
<td>0.000</td>
</tr>
<tr>
<td>(MI-LG)</td>
<td>0.019</td>
<td>0.891</td>
</tr>
<tr>
<td>(MI-HG)</td>
<td>0.000</td>
<td>0.023</td>
</tr>
<tr>
<td>(HI-LG)</td>
<td>0.001</td>
<td>0.067</td>
</tr>
<tr>
<td>(LI-NG)</td>
<td>0.001</td>
<td>0.005</td>
</tr>
<tr>
<td></td>
<td>(LI-HG)</td>
<td>(MI-LG)</td>
</tr>
<tr>
<td>(LI-HG)</td>
<td>0.842</td>
<td>0.019</td>
</tr>
<tr>
<td>(MI-LG)</td>
<td>0.176</td>
<td>0.298</td>
</tr>
<tr>
<td>(MI-HG)</td>
<td>0.046</td>
<td>0.176</td>
</tr>
<tr>
<td>(HI-LG)</td>
<td>0.102</td>
<td>0.315</td>
</tr>
<tr>
<td>(LI-NG)</td>
<td>0.027</td>
<td>0.051</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>State</th>
<th>Implied Stationary Probabilities and Asset Prices</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>$\pi^*$  $P/E$  $i_{0.25}$  $i_1$  $i_5$  $V$  $P/C$</td>
</tr>
<tr>
<td>(LI-HG)</td>
<td>0.353  17.696  0.036  0.037  0.039  0.145  2.137</td>
</tr>
<tr>
<td>(MI-LG)</td>
<td>0.150  10.085  0.081  0.090  0.091  0.337  0.782</td>
</tr>
<tr>
<td>(MI-HG)</td>
<td>0.306  12.324  0.081  0.082  0.083  0.109  3.464</td>
</tr>
<tr>
<td>(HI-LG)</td>
<td>0.132  8.393   0.150  0.144  0.129  0.222  0.704</td>
</tr>
<tr>
<td>(LI-NG)</td>
<td>0.057  27.721  0.037  0.038  0.042  0.326  1.477</td>
</tr>
</tbody>
</table>

The top and middle panels report the quarterly and five-year implied transition probability matrix between the five states implied from the generator matrix elements displayed in Table 1. Rows may not sum to one due to rounding. The bottom panel report the implied stationary probabilities and implied prices of the variables used in the SMM estimation procedure in the five states. $\pi^*$ is the stationary probability of each state, $P/E$ is the price-earnings ratio, $i_T$ is the Treasury yield with maturity $T$ in each state, $V$ is the at-the-money implied volatility, and PCR is the ratio of 5% OTM put-to-call implied volatilities. The $P/E$ ratio and bond yields are computed as shown in Proposition 1. Implied Volatility and Put-Call Ratio are as calculated in Appendix 2.
We display the fits of the variables used in our SMM procedure: the fundamentals, and the 6 overidentifying conditions. For the two fundamentals we provide the regression results for the equation $x(t) = \alpha + \beta E[x|\mathcal{F}_t] + \epsilon(t)$, where $x(t)$ is the realized growth and $E[x|\mathcal{F}_t]$ is investors’ conditional expected growth of the fundamental under consideration. The conditional expected growth is obtained from the filtered probabilities $\pi(t)$ displayed in Figure 1, and for earnings, for example, is given by $\sum_t \theta_i \pi_i(t)$. For the price (volatility) series, we present the regression results for the equation $p(t) = \alpha + \beta p(\pi(t)) + \epsilon(t)$, where $p(t)$ and $p(\pi(t))$ are the realized and model price (volatilities) conditional on investors’ beliefs at $t$ respectively. The three covariances are not used in the set of overidentifying conditions, and their out-of-sample fits are similarly reported. T-statistics are in parenthesis and are adjusted for heteroskedasticity and autocorrelation. The term EC stands for ex-crash to denote that the fourth quarter of 1987 is removed from the sample.

<table>
<thead>
<tr>
<th>Variable</th>
<th>$\alpha$</th>
<th>$\beta$</th>
<th>$R^2$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Inflation</td>
<td>-0.002</td>
<td>2.135</td>
<td>0.467</td>
</tr>
<tr>
<td></td>
<td>[-2.978]</td>
<td>[7.230]</td>
<td></td>
</tr>
<tr>
<td>Earnings</td>
<td>-0.004</td>
<td>2.693</td>
<td>0.095</td>
</tr>
<tr>
<td></td>
<td>[-1.601]</td>
<td>[3.046]</td>
<td></td>
</tr>
<tr>
<td>P/E Ratio</td>
<td>-0.007</td>
<td>1.444</td>
<td>0.727</td>
</tr>
<tr>
<td></td>
<td>[-3.503]</td>
<td>[9.871]</td>
<td></td>
</tr>
<tr>
<td>Three-Month Yield</td>
<td>-0.008</td>
<td>1.062</td>
<td>0.618</td>
</tr>
<tr>
<td></td>
<td>[-1.043]</td>
<td>[7.888]</td>
<td></td>
</tr>
<tr>
<td>One-Year Yield</td>
<td>-0.008</td>
<td>1.143</td>
<td>0.639</td>
</tr>
<tr>
<td></td>
<td>[-0.958]</td>
<td>[7.714]</td>
<td></td>
</tr>
<tr>
<td>Five-Year Yield</td>
<td>-0.007</td>
<td>1.206</td>
<td>0.670</td>
</tr>
<tr>
<td></td>
<td>[-0.705]</td>
<td>[6.270]</td>
<td></td>
</tr>
</tbody>
</table>
Table 4: Explaining At-The-Money Implied Volatility (1986:Q1 – 2006:Q2)

<table>
<thead>
<tr>
<th>No.</th>
<th>Constant</th>
<th>Lag</th>
<th>ATM&lt;sup&gt;M&lt;/sup&gt;</th>
<th>P/E</th>
<th>i&lt;sub&gt;0,25&lt;/sub&gt;</th>
<th>NBER</th>
<th>R&lt;sub&gt;t-1&lt;/sub&gt;</th>
<th>σ&lt;sub&gt;S&lt;/sub&gt;</th>
<th>R&lt;sup&gt;2&lt;/sup&gt;</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.</td>
<td>-0.011</td>
<td>0.989</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>0.502</td>
</tr>
<tr>
<td></td>
<td>[-0.377]</td>
<td></td>
<td>[6.854]</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2.</td>
<td>5.179</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>0.72</td>
<td>0.467</td>
</tr>
<tr>
<td></td>
<td>[2.442]</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>[7.964]</td>
<td></td>
</tr>
<tr>
<td>3.</td>
<td>7.569</td>
<td></td>
<td></td>
<td>0.605</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>0.220</td>
</tr>
<tr>
<td></td>
<td>[2.396]</td>
<td></td>
<td></td>
<td>[4.542]</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>4.</td>
<td>19.335</td>
<td></td>
<td></td>
<td>-0.029</td>
<td></td>
<td></td>
<td></td>
<td>0.000</td>
<td>0.000</td>
</tr>
<tr>
<td></td>
<td>[5.981]</td>
<td></td>
<td></td>
<td>[-0.052]</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>5.</td>
<td>0.235</td>
<td></td>
<td></td>
<td>-0.047</td>
<td></td>
<td></td>
<td></td>
<td>0.042</td>
<td>0.042</td>
</tr>
<tr>
<td></td>
<td>[46.515]</td>
<td></td>
<td></td>
<td>[-3.509]</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>6.</td>
<td>0.178</td>
<td></td>
<td></td>
<td>-2.009</td>
<td></td>
<td></td>
<td></td>
<td>0.329</td>
<td>0.329</td>
</tr>
<tr>
<td></td>
<td>[17.279]</td>
<td></td>
<td></td>
<td>[-7.996]</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>7.</td>
<td>5.751</td>
<td>0.705</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>0.501</td>
<td>0.501</td>
</tr>
<tr>
<td></td>
<td>[3.645]</td>
<td>[0.070]</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>8.</td>
<td>-0.531</td>
<td>0.441</td>
<td>0.548</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>0.574</td>
<td>0.574</td>
</tr>
<tr>
<td></td>
<td>[-0.274]</td>
<td>[4.614]</td>
<td>[4.726]</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>9.</td>
<td>-0.008</td>
<td>0.469</td>
<td>0.516</td>
<td>0.001</td>
<td>0.002</td>
<td>-0.024</td>
<td>0.467</td>
<td>0.587</td>
<td>0.587</td>
</tr>
<tr>
<td></td>
<td>[-0.244]</td>
<td>[4.663]</td>
<td>[4.264]</td>
<td>[0.852]</td>
<td>[0.958]</td>
<td>[-1.819]</td>
<td>[1.153]</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

The table reports the quarterly time series regressions

\[
\text{ATM}(t) = \beta_0 + \beta_1 \text{ATM}(t-1) + \beta_2 \text{ATM}^M(t) + \beta_3 \text{P/E}(t) + \beta_4 \text{i}_{0.25}(t) + \beta_5 \text{NBER}(t) + \beta_6 R_{t-1} + \beta_7 \sigma_S(\pi_t) + \epsilon(t).
\]

In different lines some of the \(\beta_i\) are set to zero. ATM(t) is the at-the-money implied volatility on S&P 500 index options traded on the CBOE with approximately three months to maturity. ATM<sup>M</sup> is the at-the-money implied volatility implied by our model and calculated as shown in Appendix 2. The historical and model implied series are shown in the top panel of Figure 4. The latter are calculated conditional on investors’ beliefs of fundamental drift states that are extracted and displayed in Figure 1. P/E is the price to operating income ratio of S&P 500 firms, i<sub>0,25</sub> is the 3-month Treasury Bill rate, NBER is the quarterly expansion indicator created by the NBER, and R<sub>t-1</sub> is the one quarter lagged returns in periods when it is negative on the S&P 500 index. T-Statistics are in parenthesis and are adjusted for heteroskedasticity and autocorrelation using the Newey and West (1987) method. The symbol * indicates statistical significance at the 5% level.
<table>
<thead>
<tr>
<th>No.</th>
<th>Constant</th>
<th>Lag</th>
<th>Model</th>
<th>Trader Sentiment</th>
<th>Investor Sentiment</th>
<th>$R^2$</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.</td>
<td>0.226</td>
<td>0.846</td>
<td></td>
<td></td>
<td></td>
<td>0.474</td>
</tr>
<tr>
<td></td>
<td>[1.776]</td>
<td>[8.339]</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2.</td>
<td>0.429</td>
<td>0.665</td>
<td></td>
<td></td>
<td></td>
<td>0.451</td>
</tr>
<tr>
<td></td>
<td>[3.805]</td>
<td>[7.356]</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3.</td>
<td>1.219</td>
<td>-0.925</td>
<td></td>
<td></td>
<td></td>
<td>0.080</td>
</tr>
<tr>
<td></td>
<td>[32.779]</td>
<td>[-1.777]</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>4.</td>
<td>0.233</td>
<td>0.837</td>
<td>-0.058</td>
<td></td>
<td></td>
<td>0.461</td>
</tr>
<tr>
<td></td>
<td>[1.776]</td>
<td>[7.614]</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>5.</td>
<td>0.208</td>
<td>0.357</td>
<td>0.476</td>
<td>-0.470</td>
<td></td>
<td>0.565</td>
</tr>
<tr>
<td></td>
<td>[2.801]</td>
<td>[3.406]</td>
<td>[4.367]</td>
<td>[-1.408]</td>
<td></td>
<td></td>
</tr>
<tr>
<td>6.</td>
<td>1.327</td>
<td>-0.413</td>
<td></td>
<td></td>
<td></td>
<td>0.129</td>
</tr>
<tr>
<td></td>
<td>[29.729]</td>
<td>[-2.424]</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>7.</td>
<td>0.308</td>
<td>0.792</td>
<td></td>
<td>-0.116</td>
<td></td>
<td>0.470</td>
</tr>
<tr>
<td></td>
<td>[2.658]</td>
<td>[8.923]</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>8.</td>
<td>0.178</td>
<td>0.355</td>
<td>0.527</td>
<td>-0.042</td>
<td></td>
<td>0.548</td>
</tr>
<tr>
<td></td>
<td>[2.104]</td>
<td>[3.342]</td>
<td>[4.801]</td>
<td>[-0.373]</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

The table reports the quarterly time series regressions

\[ P/C(t) = \beta_0 + \beta_1 P/C(t-1) + \beta_2 P/C^M(t) + \beta_3 \text{Sentiment}(t) + \epsilon(t). \]

In different lines some of the $\beta_i$ are set to zero. $P/C(t)$ is the ratio of implied volatilities of 5% out-of-the-money put and call options on S&P 500 index options traded on the CBOE with approximately three months to maturity. $P/C^M$ is the put-call ratio implied by our model and calculated as shown in Appendix 2. The historical and model implied series are shown in the bottom panel of Figure 4. The latter are calculated conditional on investors’ beliefs of fundamental drift states that are extracted and displayed in Figure 1. Sentiment is either “Trader Sentiment” or “Investor Sentiment”. Trader sentiment is the net long position of large speculators on S&P 500 index futures obtained from the Commodity Futures Trading Commission’s Commitment of Traders Report. Investor sentiment is the bull-bear spread (proportion of traders bullish less bearish) in Investor’s Intelligence’s survey of investment newsletter writers. T-Statistics are in parenthesis and are adjusted for heteroskedasticity and autocorrelation using the Newey and West (1987) method. The symbol * indicates statistical significance at the 5% level.
The states are numbered as \((\beta_1, \theta_2), (\beta_2, \theta_1), (\beta_3, \theta_2), (\beta_3, \theta_1),\) and \((\beta_1, \theta_3)\), where \(\beta_i, i = 1, 2, 3\) are the low, medium, and high states of inflation, and \(\theta_1, \theta_2,\) and \(\theta_3,\) are the regular low and high states, and the “new economy” rates of earnings growth. The filtered beliefs are obtained from the SMM procedure in Appendix B. The calibrated values of the parameters are shown in Table 1. Shaded areas represent NBER-dated recessions.
Figure 2: Fundamental and Financial Variables: Empirical and Model Fitted (1960-2006:Q2)

Historical values of financial and fundamental variables series (D) are in solid lines and their fitted values (M) from the SMM estimation procedure in Appendix B are in dashed lines. The calibrated values of the parameters are shown in Table 1. The filtered beliefs series of investors used to generate the fitted values are shown in Figure 1. Shaded areas represent NBER-dated recessions.
Figure 3: Strike-Adjusted 3-Month Densities of Stock Returns Under the Risk-Neutral Measures in the Five States

Densities in the five states are calculated as shown in Appendix 2.
Figure 4: Data and Model Fitted At-the-Money Implied Volatility and 5% Out-of-the-Money Put-Call Implied Volatility Ratio (1986:Q2-2006:Q2)
Data and model ATM implied volatility are shown in the top panel of Figure 4. The model volatility of variance is computed using 20 and the filtered belief series in Figure 1. Earnings and inflation uncertainties are computed using (24) and the filtered beliefs.
Figure 6: Relationship Between P/E and Put-Call Ratios

- **Data, 1986:Q2 -- 2006:Q2**: Correlation = -0.24
- **Model, 1986:Q2 -- 2006:Q2**: Correlation = -0.3
- **Data, 1986:Q2 -- 1993:Q4**: Correlation = 0.39
- **Model, 1986:Q2 -- 1993:Q4**: Correlation = 0.76
- **Data, 1994:Q1 -- 2006:Q2**: Correlation = -0.6
- **Model, 1994:Q1 -- 2006:Q4**: Correlation = -0.85
The first and second panels show the historical realized volatility, which in each quarter is the sum of squared daily returns on the S&P 500 index, and its forecasted value from two sets of forecasts based on information available in the previous quarter. The forecast labeled “D” (first panel) is the fitted value from the regression in (25) Similarly, the forecast labeled “M” (second panel) is the fitted value from the regression in (26) The third panel shows the data volatility premium defined as $ATM(t + 1) - Vol^D(t + 1)$, where $ATM(t + 1)$ is the at-the-money implied volatility on S&P 500 index options, and $Vol^D(t + 1)$ is the forecast in (25). Similarly the model volatility premium is difference between the implied volatility from our model shown in Figure 4 and the model forecast of realized volatility in (26). The model volatility premium explains 29.5 % of the variation in the data premium as shown in the regression in (27). The fourth panel plots the model implied volatility premium and its fitted value from the regression in (31) that has the model’s volatility of volatility series calculated from (20) using the beliefs series in Figure 1.