Asset Prices with Contingent Risk Preferences*

Stephen Gordon and Pascal St-Amour†

First draft: July 10, 1997.
This version: June 8, 1998

Abstract

This paper develops a consumption-based asset pricing model in which attitudes towards risk are contingent upon the state of the world. For low (high) level of consumption relative to a subjective metric, counter-cyclical (pro-cyclical) risk aversion implies that consumption shocks generate larger fluctuations in marginal utility, against which the agent will hedge in choosing his optimal portfolio. Asset prices are studied using two-state Markov preference regimes where bull and bear markets reflect alternating periods of low and high risk aversion. Joint estimation of bond and stock prices highlights moderate and infrequent movements in risk aversion, and a marked improvement on the model's ability to capture the cyclical nature of observed asset prices. We also study implications for returns under Itô preference states, and show why contingent risk preferences have the potential to resolve the empirical anomalies for asset returns. Exact likelihood estimation of joint diffusion processes using market excess returns data also point toward realistic and counter-cyclical estimates of risk aversion.

Keywords: Asset pricing models, Bayesian analysis, excess volatility, exact likelihood estimation of diffusion processes, Markov chain, regime switching, risk aversion, state-dependent preferences.

JEL classification: C11, D80, G12

*This is a revised version of 'A Preference Regime Model of Bull and Bear Markets', Cahier de recherche #9712, Département d'économique, and #97-36, CRÉFA, Université Laval. This research is supported by a grant from the Social Sciences and Humanities Research Council of Canada. We are thankful for comments made by seminar participants at the Graduate School of Business, University of Chicago and at the Centre Interuniversitaire de Recherche en Analyse des Organisations (CIRANO).

†Département d’Économique and Centre de Recherche en Économie et Finance Appliquées, Université Laval, Cité Universitaire, Québec, Canada, G1K 7P4, e-mail: stephen.gordon@ecn.ulaval.ca and pascal.st-amour@ecn.ula.ca
1 Introduction

The main proposition of the consumption-based capital asset pricing model (C-CAPM) is that a risk-averse representative agent with state- and time-separable preferences allocates consumption and portfolio choices so as to smooth out risks to the inter-temporal marginal rates of substitution (IMRS) over uncertain consumption streams (Lucas 1978, Breeden 1979). The intuition behind this result is straightforward: assets that are expected to pay high returns when marginal utility is high will tend to be valued more by the agent than those whose returns are negatively correlated with anticipated marginal utility. Hence, under the C-CAPM, stock prices reflect the present value of anticipated dividends, where the discount factor is the IMRS between consumption in the current period and consumption in the periods where dividends are paid out; bond prices are based on the conditional mean of IMRS. Furthermore, if within-period utility is iso-elastic, the model predicts that expected excess returns for risky assets are equal to their covariance with consumption (the quantity of risk) times the coefficient of relative risk aversion (the price of risk).

The empirical difficulties of the C-CAPM are well documented (Kocherlakota 1996, Campbell 1996, Cochrane 1997, provide enlightening surveys). First, the moderate quantity of consumption risk does not warrant the high average excess returns we observe, unless an unreasonably large price of risk is assumed (the equity premium puzzle). Secondly, a high curvature coefficient for within-period utility reduces the average MRS; because observed bond prices are high, the subjective discount factor must be increased to levels greater than one (the risk-free rate puzzle). Third, stock prices are quite cyclical relative to the discounted dividends stream while the excess returns are distressingly predictable, even allowing for time-varying quantities of risk (the predictability puzzle).

Recent research focuses on time-varying prices of risk as a potential source of mis-specification. In particular, counter-cyclicality in risk aversion is introduced through wealth dependencies (Bakshi and Chen 1996) and through time-varying habits in which the bliss point is a function of past consumption (Campbell and Cochrane 1995). Both models can be seen as special cases of state-contingent preferences in which attitudes are determined by contemporary state variables, whose evolution may or may not be affected by
the agent’s decisions. These preferences have been advocated in other settings for the study of changing, or apparently excessive risk aversion, e.g., Karni (1983, 1987). Unfortunately, their empirical success within the context of the C-CAPM has been mixed. Implementation of the Bakshi and Chen (1996) model requires uncontroversial proxies for aggregate wealth - which are difficult to obtain - and the degree of non-linearities involved forces the use of restricted cases, thus abandoning much of the potential for improvement over the standard model. Moreover, relying on consumption volatility to explain time-varying risk aversion also limits the potential gains: since aggregate consumption remains frustratingly smooth, time-varying habits - albeit useful for addressing the risk-free rate puzzle - do not generate sufficient additional risk to reconcile the high premia with reasonable prices of risk.

We adopt what might be considered an unrestricted approach to time-varying prices of risk: the state variable that underlies the agent’s preferences is treated as a latent variable. In our model, within-period utility is conditionally (upon the prevailing state of the world) iso-elastic, and we ask what pattern of attitudes toward risk might be consistent with observed movements of asset prices and of asset returns. When risk aversion is state-contingent, low (high) consumption and counter-cyclical (pro-cyclical) curvature imply much more dramatic movements in marginal utility. These fluctuations affect the agent’s choice over risk-free and risky assets in hedging against IMRS risk.

In our discussion of asset prices, we develop a simple discrete-time model in which risk aversion follows a two-state Markov process. This specification has an intuitive interpretation: the celebrated ‘bear’ and ‘bull’ market classification of asset market activity is produced as attitudes toward risk shift between preference regimes. The results suggest a marked improvement in reproducing first and second moments of stock and bond prices, as well as being extremely useful in addressing the cyclical nature of equity prices. These estimated parameters point toward (i) realistic levels of risk aversion, and (ii) infrequent, moderate - though significant - and counter-cyclical shifts in curvature. The main result therefore comes as good news for the standard model: within a given regime, the conventional C-CAPM performs surprisingly well in reproducing asset prices; its inability to reproduce the financial market cycle might stem from the inadequate restriction that risk aversion is constant throughout the sample.
In our treatment of asset returns, we find it useful to make use of a continuous-time model in which the preference states are generated by a diffusion process. We show that specifying a joint Itô process for consumption and conditional risk aversion has the potential to successfully resolve the equity premium, the risk-free rate and the predictability puzzles. First, excess returns reflect two independent contributors to IMRS risk: consumption and curvature risk (i.e. the covariance between risk preferences and returns); this second element justifies high risk premia without requiring excessive risk aversion. Furthermore, the loading for the concavity risk is a function of consumption, which has the potential for explaining cyclical movements in returns. Finally, counter-cyclicality of risk aversion, as well as the additional volatility in the IMRS induce a precautionary demand for risk-free assets, which would justify high bond prices without requiring negative discounting of future utility.

In estimating our models, we make use of Bayesian methods of analysis. Since risk aversion is treated as a latent variable, Markov chain Monte Carlo (MCMC) and data augmentation techniques are especially useful in estimating this class of models; a by-product of our algorithm is an estimate for the realized risk aversion series.

We conclude with the familiar calls for further research: relaxing the hypothesis of fixed attitudes toward risk appears to be a step in the right direction. We use the two extremes of a two-state Markov and a diffusion processes to illustrate the potential that this line of research has to offer; it is presumably the case that better parameterizations lie somewhere in between. The rest of the paper is organized as follows. Section 2 develops the state-contingent model and studies the implications for IMRS risk. Asset prices are discussed in Section 3, where the two-state parametrization is used. Section 4 focuses on the diffusion processes in addressing returns puzzles. Finally, a conclusion reviews the main findings as well as suggestions for future research.
2 Asset Pricing with Contingent Preferences

We consider the case where a representative agent solves

\[
J(\omega_{0,-1}, \omega_{1,-1}, \ldots, \omega_{n,-1}) = \max_{\{C_{t,0}, \omega_{0,0}, \omega_{1,0}, \ldots, \omega_{n,0}\}} E_0 \sum_{t=0}^{\infty} \beta^t U[C_{t,0}, \gamma(S_{t,0})],
\]

subject to

\[
C_t + P_{0,t} \omega_{0,t} + \sum_{i=1}^{n} P_{i,t} \omega_{i,t} \leq \omega_{0,t-h} + \sum_{i=1}^{n} (P_{i,t} + D_{i,t}) \omega_{i,t-h} + W_t.
\]

Here, \(C_t\) is consumption; \(P_{i,t}\) is the (ex-dividends) price of a generic risky asset \(i \in \{1, \ldots, n\}\), conditional on state of the world \(S_t\), with dividends \(D_{i,t}\); while \(P_{0,t}\) is the price of a real risk-free bond paying one unit of the numeraire in the following period. Correspondingly, \(\omega_{i,t}\) is the number of shares of the risky asset \(i\), and \(\omega_{0,t}\) is the quantity of bonds held by the agent. Finally, a unit of labour is supplied inelastically by the agent at all periods, earning a real wage of \(W_t\). All markets are frictionless, with trading intervals given by \(h\).

Within-period (local) utility is assumed to be strictly concave, and conditionally iso-elastic:

\[
U[C_t, \gamma(S_t)] = \Theta \left(\Theta^{-1} C_t\right)^{1-\gamma_t} \frac{1}{1-\gamma_t}
\]

where \(\Theta > 0\) is a constant preference parameters, and where \(\gamma_t = \gamma(S_t)\) is a continuous, strictly positive state-dependent curvature index that can be interpreted as the Arrow-Pratt coefficient of relative risk aversion for static lotteries.

Risk aversion will display time variation in dynamic settings following movements in consumption and in \(\gamma_t\). Such shifts in attitudes could be related to numerous factors. A natural contender would be wealth. Robson (1992) introduces preference over status (which is proxied by relative wealth) to reproduce concave-convex-concave utility functions. Building on this idea, Balsi and Chen (1996) allow for wealth-dependent attitudes toward risk in deriving consumption and portfolio choices. Campbell and Cochrane (1995) are
able to generate counter-cyclical attitudes toward risk by introducing a ‘slow-moving’ habit formation specification: when consumption falls toward a time-varying bliss parameter during downturns, risk aversion increases. Moreover, state-contingent preferences have also been used in other settings for behavior characterized by apparently excessive, and/or changing risk aversion e.g. Karni (1983, 1987).

The specification in (2.3) is a special state-contingent case of the hyperbolic absolute risk aversion (HARA) functionals $U(C) = \Theta(\Theta^{-1}C - \eta)^{1-\gamma}/(1-\gamma)$, with bliss point $\eta$ set to zero. When concavity is time-varying, the parameter $\Theta$ affects marginal rates of substitution. In particular, the effects of fluctuations in concavity on MRS depend crucially on the level of consumption, relative to $\Theta$, as well as on the covariance between innovations in consumption and curvature. To see this, we plot marginal utility $U_c(C, \gamma_t)$ for $\gamma_t = \gamma_h$ and for $\gamma_t = \gamma_l < \gamma_h$.

![Diagram](image)

When consumption is low, e.g. $C = C_l < \Theta$, if risk aversion covaries negatively with consumption, an unanticipated decline in $C$ is associated with an increase in $\gamma$ from $\gamma_l$ to $\gamma_h$, which rotates marginal utility clockwise from $U_c(C, \gamma_l)$ to $U_c(C, \gamma_h)$. Similarly, for consumption at high levels $C_h$, if risk aversion is pro-cyclical, an unanticipated decrease in consumption is associated with a reduction in $\gamma$ from $\gamma_h$ to $\gamma_l$. Both shocks result in much larger fluctuations in MRS’s than a standard, iso-elastic model would suggest – which would correspond to movements in consumption in the vicinity of $\Theta$. Depending on the direction
of co-movements and levels of consumption, state-dependent risk aversion can generate greater MRS risk, against which the agent will want to hedge in choosing optimal leverage and portfolio weights.

Returning to the agent’s problem, standard arguments yield the equilibrium \( n + 1 \) Euler equations for the solution to (2.1):

\[
1 = \beta^h E_t \left[ \frac{\left( \Theta^{-1} C_{t+h} \right)^{-\gamma_{t+h}}}{\left( \Theta^{-1} C_t \right)^{-\gamma_t}} \left( \frac{P_{i,t+h} + D_{i,t+h}}{P_{i,t}} \right) \right],
\]

for the \( n \) risky assets, while the risk-free asset satisfies:\(^1\)

\[
1 = \beta^h E_t \left[ \frac{\left( \Theta^{-1} C_{t+h} \right)^{-\gamma_{t+h}}}{\left( \Theta^{-1} C_t \right)^{-\gamma_t}} \left( \frac{1}{P_{0,t}} \right) \right].
\]

The implications of the two valuation equations (2.4a) and (2.4b) for asset prices and returns depend crucially on the law of motion for \( \gamma(S_t) \). In the next section, we analyze predicted bond and stock prices when concavity follows a two-state Markov process. In Section 4, we let the curvature index follow a diffusion process to study the equity premium and the risk-free rate puzzles.

### 3 Implications for Asset Prices: Markov Preferences

#### 3.1 Financial Cycles and Preference Regimes

Figure 1 plots the fitted asset prices generated by the standard model along the realized series. It is well known that asset prices are much more volatile than the model would predict (LeRoy and Porter 1981, Shiller 1981, Cochrane 1992, among others). Furthermore, stock prices remain ‘too high’ or ‘too low’ for periods of several years, episodes that are often referred to as an asset market cycle of ‘bull’ and ‘bear’ markets.

One approach to reproduce the cyclical nature of asset prices is to specify the growth of endowments as a Markov process as in Cecchetti, Lam and Mark (1990, 1993). However, this does not compensate for the fact that consumption and dividend series remain comparatively smooth: predicted asset prices have cycles,

\(^1\)Note that the presence of labour income \( W_t \) in our model implies that strict correspondence between consumption and dividends à la Lucas (1978), Mehra and Prescott (1985), or Abel (1988) is not imposed. See Abel (1994) for more on this issue.
but the amplitude of the fitted cycle is much smaller than that of observed asset prices. In any case, our data are poorly served by a Markov switching model for endowments; a formal model comparison exercise decisively rejects this specification in favor of a VAR alternative.\footnote{\text{The log Bayes factor in favor of the VAR model is 50, which implies overwhelming rejection of the Markov-switching model. Bonomo and García (1994) also reject the two-state Markov switching model imposed on endowments, when tested against a more general class of Markov structures.}}

Suppose instead that risk aversion $\gamma_t$ follows a discrete-time (i.e., the trading interval $h = 1$) two-state Markov process:

$$
\gamma_t = \begin{cases} 
\gamma_0 \geq 0, & \text{if } S_t = 0 \\
\gamma_1 \geq 0, & \text{if } S_t = 1 
\end{cases}
$$

(3.1)

where $\mathcal{P}(S_{t+1} = s|S_t = s) = \pi_s, s = 0, 1$. We assume also that dividends and consumption are jointly log-normal, and define $d_t^j \equiv [d_{1,t}, \ldots, d_{n,t}, c_t]^T$ such that $d_t^j \sim N(\mu_t^j, \Sigma_t^j)$. The parameters $\mu_t^j$ and $\Sigma_t^j$ respectively denote conditional mean and variance of $d_t^{j+1}$, and have generic elements $\mu_{j,t}$ and $\sigma_{j,k,t}$ for $j, k = 1, \ldots, n, c$.

Appendix A demonstrates that state-dependent stock and bond prices are:

$$
P_{s,t}^j = B_{s,t}^j (\Theta^{-1}C_t)^{\gamma_t} \quad j = 0, 1, \ldots n; \quad s = 0, 1. 
$$

(3.2a)

For the risky assets $i = 1, \ldots, n$, the coefficients $B_{s,t}^j$ solve

$$
\begin{bmatrix} B_{0,t}^j \\ B_{1,t}^j \end{bmatrix} = \beta \begin{bmatrix} \pi_0 & 1 - \pi_0 \\ 1 - \pi_1 & \pi_1 \end{bmatrix} \begin{bmatrix} E_tB_{0,t+1}^j \\ E_tB_{1,t+1}^j \end{bmatrix} + \begin{bmatrix} M_{0,t}^j \\ M_{1,t}^j \end{bmatrix},
$$

(3.2b)

where

$$
M_{j,t}^i \equiv E_t\{\exp[d_{t+1} - \gamma_s(c_{t+1} - \theta)]\}
$$

(3.2c)

$$
= \exp[\mu_{i,t} - \gamma_s(\mu_{c,t} - \theta) + 0.5(\sigma_{i,t} - 2\gamma_s\sigma_{c,t} + \gamma^2_s\sigma_{c,c})].
$$

For the risk-free asset, we obtain:

$$
\begin{bmatrix} B_{0,t}^i \\ B_{1,t}^i \end{bmatrix} = \beta \begin{bmatrix} \pi_0 & 1 - \pi_0 \\ 1 - \pi_1 & \pi_1 \end{bmatrix} \begin{bmatrix} \exp[-\gamma_0(\mu_{c,t} - \theta) + 0.5\gamma^2_0\sigma_{c,c}] \\ \exp[-\gamma_1(\mu_{c,t} - \theta) + 0.5\gamma^2_1\sigma_{c,c}] \end{bmatrix}.
$$

(3.2d)
The weighting matrix $\pi$ and the presence of the subjective discount factor $\beta \in (0, 1)$ guarantee that the loadings $B_t$ are bounded if the conditional means $M_{i,t}$ do not grow too rapidly. Moreover, the structure can be readily adapted to accommodate time-varying transition matrices $\pi_t$ or a larger-dimensional state space.

In the standard model, current prices reflect the discounted stream of expected dividends, and the inter-temporal marginal rates of substitution between the current period and the period at which dividends are paid out to shareholders serves as the discount factor. If the IMRS are (preference) state-independent, the expression for prices is also state-independent. When Markov preferences are introduced, equations (3.2a)–(3.2c) imply that equilibrium prices must follow:

$$
\begin{bmatrix}
P^0_{i,t} \\
P^1_{i,t}
\end{bmatrix} = E_t \sum_{j=1}^{\infty} D_{i,t+j} \begin{bmatrix}
\psi_{00,j} \text{IMRS}_{i,t+j}^{0,0} + \psi_{01,j} \text{IMRS}_{i,t+j}^{0,1} \\
\psi_{10,j} \text{IMRS}_{i,t+j}^{1,0} + \psi_{11,j} \text{IMRS}_{i,t+j}^{1,1}
\end{bmatrix},
$$

where $\text{IMRS}_{i,t+j}^{k,l} \equiv \beta^j (\Theta^{-1} C_{i+j})^{-\gamma_1} / (\Theta^{-1} C_i)^{-\gamma_s}$, for $k,l \in \{0,1\}$, and $\psi_j \equiv \pi^j$. Conditional upon the current state, the discount factors are a weighted average of two IMRS. If the elements of the transition matrix are all strictly positive, all elements of $\psi_j$ will be positive for all $j$, so the agent must always take into account the non-zero probability for the event that the preference state that prevails at the purchase of the asset will not be the same as the one that will occur when dividends are paid out.

The implications for predicted stock prices volatility are intuitive. If risk aversion is counter-cyclical, current consumption is low, and if the off-diagonal terms in the transition matrix $\pi$ are non-zero, there is a positive probability that the agent will reap dividends when marginal utility is much lower. As a result, holding stock entails an additional risk over and above any positive covariance between dividends and consumption, and the price he is willing to pay is reduced. Conversely, if high current consumption coincides with a low curvature index, the possibility that the agent may be in a high marginal utility state when dividends are paid out reduces the risk of holding the risky asset, which increases the price he is willing to pay for the stock.
In this framework, a shift from the good state to the bad state, followed by an anticipated return to the good state would signal the beginning of a bear market. This sequence of movements from one preference state to another forms the basis for an explanation for the cyclical movements in stock prices and for the persistent errors of the standard model. Perhaps more significantly, these movements are consistent with rational behavior, and need not be ascribed to some unpleasant modeling defect such as a violation of the transversality condition, or a persistent dichotomy between subjective and objective distributions (Cecchetti, Lam and Mark 1998). Finally, we should emphasize that the regime shifts in the model should not be interpreted as permanent structural breaks for preferences. Indeed, the large cyclical fluctuations in stock prices following a regime shift are generated by the belief that preferences will eventually return to the initial state.

3.2 Estimation

This section specifies our statistical model and sets out our prior beliefs for the parameters. For implementation purposes, we will consider henceforth the case of a single composite risky asset with price $P_t$, and dividends $D_t$, while we denote bond prices by $Q_t$.

3.2.1 Stochastic Specification

We suppose that the first differences in logs of $C_t$ and $D_t$ are fairly well-described by a VAR(1) model:

$$
\begin{bmatrix}
    d_t - d_{t-1} \\
    c_t - c_{t-1}
\end{bmatrix} = \begin{bmatrix}
    \phi_d \\
    \phi_c
\end{bmatrix} + \begin{bmatrix}
    \phi_{dd} & \phi_{dc} \\
    \phi_{cd} & \phi_{cc}
\end{bmatrix} \begin{bmatrix}
    d_{t-1} - d_{t-2} \\
    c_{t-1} - c_{t-2}
\end{bmatrix} + \begin{bmatrix}
    \varepsilon_{d,t} \\
    \varepsilon_{c,t}
\end{bmatrix},
$$

(3.4)

where $[\varepsilon_{d,t}, \varepsilon_{c,t}]' \sim iid N(0, \Sigma)$. In order to simplify the exposition below, we denote the joint density for a sample of observations generated by (3.4) by $P(d,c|\phi, \Sigma)$, where the dependence on the pre-sample values of $d_0$ and $c_0$ has been suppressed, $d$ and $c$ represent the vector of $T$ time series observations for $d_t$ and $c_t$, and where $\phi \equiv [\phi_d, \phi_{dd}, \phi_{dc}, \phi_{cd}, \phi_{cc}]'$. If $S$ represents the time series of $T$ realizations for the states, let $P(S|\pi, S_0)$ denote the probability of observing this sequence.
Given sequences of $B^0_{i,t}$ and $B^1_{i,t}$ that satisfy (3.2b) and (3.2d), we model the state-dependent asset prices by

$$
p_t = S_t[b^1_{p,t} + \gamma_1(c_t - \theta)] + (1 - S_t)[b^0_{p,t} + \gamma_0(c_t - \theta)] + u_{p,t}
$$
$$
q_t = S_t[b^1_{q,t} + \gamma_1(c_t - \theta)] + (1 - S_t)[b^0_{q,t} + \gamma_0(c_t - \theta)] + u_{q,t}
$$

where $b^1_{p,t} = \log[B^1_{p,t}]$, $b^1_{q,t} = \log[B^0_{q,t}]$, and where $u_t = [u_{p,t}, u_{q,t}]'$ is an iid $N(0, \Omega)$ vector of measurement error terms. Let $p$ and $q$ represent the set of $T$ observations of $p_t$ and $q_t$, respectively. Although the parameters $\beta$, $\pi$, $\phi$, $\Sigma$ and the dividend series $d$ do not appear explicitly in (3.5), these values are used to compute $b^1_{p,t}$ and $b^1_{q,t}$ [see Appendix C for details]. In order to clarify the role of these parameters and of the dividend series in determining asset prices, we denote the joint density for asset prices by

$$
P(p, q|d, c, S, \beta, \gamma_0, \gamma_1, \theta, \phi, \Sigma, \pi, \Omega).
$$

Taken together, the conditional densities for $(d, c)$, $S$ and $(p, q)$ form a joint density for asset prices, preference states, dividends and consumption that is used as a basis for estimating our model:

$$
P(p, q, d, c, S|\beta, \gamma_0, \gamma_1, \theta, \phi, \Sigma, \pi, S_0, \omega) = P(p, q|d, c, S, \beta, \gamma_0, \gamma_1, \theta, \phi, \Sigma, \pi, \Omega)P(d, c|\phi, \Sigma)P(S|\pi, S_0).
$$

### 3.2.2 Priors

The structure of our model allows us to make use of a Bayesian MCMC estimation algorithm and to deal with issues such as the unobserved nature of the preference states and the imposition of the regularity conditions required by the structural model. The prior distributions used in the estimation are:

$$
P(\beta) = B(997.5, 2.5) \quad P(\gamma_0, \gamma_1) = N_+(\mu_2, 100^2I_2) \quad P(\theta) = N(0, 300^2)
$$
$$
P(\phi) = N(0, 10^2I_0) \quad P(\Sigma^{-1}) = W(0.025^2I_2, 5) \quad P(\pi_s) = B(95, 5), \ s = 0, 1
$$
$$
P(S_0 = 1) = 0.5 \quad P(\Omega^{-1}) = W(0.025^2I_2, 5)
$$

where $B, W$ and $G$ denote the beta, Wishart and gamma distributions, respectively $I$ is the unit vector.
Although all prior distributions are proper, most are quite diffuse; the priors for \( \gamma_0, \gamma_1, \theta, \phi \) are all locally uniform in the region in which the likelihood function has mass. The Wishart prior distributions for \( \Sigma^{-1} \) and \( \Omega^{-1} \) are centered over \( \sigma_d = \sigma_c = \omega_p = \omega_q \approx 0.01 \) and \( \sigma_{dc} = \omega_{pq} = 0 \), but setting the degrees of freedom parameter at 5 means that the prior will play only a small role in comparison to the information contained in the 390 observations of our sample.

Stronger priors are adopted for the subjective rate of discount and the transition probabilities. The prior for \( \beta \) is consistent with a mean annual rate of time preference of 0.03 and a prior standard deviation of about 0.02. The prior probability of staying in a given state is represented by a beta distribution that is consistent with a mean prior expected duration of roughly two years for each state, and with a prior standard deviation of 16 months; the size of the fictitious prior sample is 100, which is roughly one-fourth the size of the sample. Equal prior weight is assigned to both \( S_0 = 0 \) and \( S_0 = 1 \). The conditions \( 0 < \beta < 1 \) and \( \gamma_0, \gamma_1 > 0 \) ensure the existence of a general equilibrium, and are imposed by the priors. The additional restriction \((\gamma_1 - \gamma_0)(c_t - \theta) > 0, \forall t\) identifies \( S_t = 1 \) as the state with lower marginal utility of consumption (good state).

### 3.3 Results

The Markov model of preference regimes is estimated using US monthly series for stock prices and dividends (S&P composite stock indices); aggregate consumption (real per-capita expenditures on nondurables and services), and bonds (3-months T-Bills). The sampling period is 1960:2 to 1992:9 (390 observations).

The complex nature of the data density (3.6) and the fact that we do not observe the latent state variable \( S \) poses significant problems for classical estimation of our model. For example, Cecchetti, Lam and Mark (1993) are obliged to make use of a mixture of estimation and calibration techniques in order to apply their model to data, instead of their preferred, likelihood-based approach (Cecchetti et al. 1993, p.30, fn. 12).

The development of Bayesian MCMC and data augmentation techniques (Chib and Greenberg 1996, provide a survey) permits us to make exact (finite-sample) inferences based on the exact likelihood function associated with (3.6). In the Markov chain used here (see Appendix D for details), we make use of the Gibbs
sampler and data augmentation to generate a sequence of autocorrelated draws from the joint posterior distribution

\[ P(\beta, \gamma_0, \gamma_1, \theta, \phi, \Sigma, \pi, \omega, S_0, S|p, q, d, c). \]  

(3.8)

These sequences are used to compute the posterior moments of the various parameters reported in Table 1 and to estimate the posterior probabilities that \( S_t = 1 \) at each data point.

Table 1: Posterior Moments: Markov Preferences Model

<table>
<thead>
<tr>
<th>parameter</th>
<th>mean</th>
<th>std. deviation</th>
</tr>
</thead>
<tbody>
<tr>
<td>( \beta )</td>
<td>0.9965</td>
<td>0.0001</td>
</tr>
<tr>
<td>( \gamma_0 )</td>
<td>1.8511</td>
<td>0.0896</td>
</tr>
<tr>
<td>( \gamma_1 )</td>
<td>1.4204</td>
<td>0.0877</td>
</tr>
<tr>
<td>( [\gamma_0 - \gamma_1] )</td>
<td>0.4307</td>
<td>0.1233</td>
</tr>
<tr>
<td>( \theta )</td>
<td>9.2492</td>
<td>0.2891</td>
</tr>
<tr>
<td>( \phi_d )</td>
<td>3.65e-05</td>
<td>2.10e-04</td>
</tr>
<tr>
<td>( \phi_c )</td>
<td>-1.61e-04</td>
<td>1.12e-04</td>
</tr>
<tr>
<td>( \phi_{dd} )</td>
<td>-0.0471</td>
<td>0.0177</td>
</tr>
<tr>
<td>( \phi_{dc} )</td>
<td>0.0742</td>
<td>0.0082</td>
</tr>
<tr>
<td>( \phi_{cd} )</td>
<td>0.1506</td>
<td>0.0179</td>
</tr>
<tr>
<td>( \phi_{cc} )</td>
<td>-0.0132</td>
<td>0.0647</td>
</tr>
<tr>
<td>( \Sigma_{dd} )</td>
<td>8.37e-05</td>
<td>0.61e-05</td>
</tr>
<tr>
<td>( \Sigma_{dc} )</td>
<td>6.44e-06</td>
<td>3.66e-06</td>
</tr>
<tr>
<td>( \Sigma_{cc} )</td>
<td>6.20e-05</td>
<td>0.45e-05</td>
</tr>
<tr>
<td>( \sigma_d )</td>
<td>0.0091</td>
<td>0.0003</td>
</tr>
<tr>
<td>( \rho_{dc} )</td>
<td>0.0893</td>
<td>0.0500</td>
</tr>
<tr>
<td>( \sigma_c )</td>
<td>0.0079</td>
<td>0.0003</td>
</tr>
<tr>
<td>( \pi_0 )</td>
<td>0.9898</td>
<td>0.0012</td>
</tr>
<tr>
<td>( \pi_1 )</td>
<td>0.9929</td>
<td>0.0015</td>
</tr>
<tr>
<td>( \Omega_{pp} )</td>
<td>0.0111</td>
<td>0.0010</td>
</tr>
<tr>
<td>( \Omega_{pq} )</td>
<td>-6.4e-05</td>
<td>4.37e-05</td>
</tr>
<tr>
<td>( \Omega_{qq} )</td>
<td>2.89e-05</td>
<td>2.53e-06</td>
</tr>
<tr>
<td>( \omega_p )</td>
<td>0.1052</td>
<td>0.0046</td>
</tr>
<tr>
<td>( \rho_{pq} )</td>
<td>-0.1113</td>
<td>0.0727</td>
</tr>
<tr>
<td>( \omega_q )</td>
<td>0.0054</td>
<td>0.0002</td>
</tr>
</tbody>
</table>

Note: posterior means and standard deviations for parameters of Markov preferences model (3.6).

The risk aversion coefficients \( \gamma_0 \) and \( \gamma_1 \) – for which weak priors were used – are well within the range that is usually considered reasonable. Moreover, our estimates suggest that risk aversion is counter-cyclical: the state with high asset prices is also associated with low risk aversion. In Section 2, it was noted that
this would be true for values of $\Theta^{-1}C$ that are less than one. In our sample, $\log[C]$ runs from 8.0 to 8.4, so our estimate for $\theta \equiv \log[9]$ is consistent with this interpretation. Although the difference between $\gamma_0$ and $\gamma_1$ does not appear to be large, the implications for marginal utility are quite dramatic: the transition from the good state to the bad state increases marginal utility by about 50%. The posterior mean of the discount factor $\beta$ – for which we used fairly strong priors – corresponds to an annual discount rate of 4.3%.

Both preference states are quite persistent, with a slightly higher probability of staying in state 1. The mean expected duration of a bear market $S_t = 0$ is 8.3 years, while that of the bull market is roughly 12.3 years. The unconditional probability that a bull market will prevail in a given month is 0.6, and our estimate for the sample mean of $S_t$ is 2/3. Although our prior means for both $\pi_0$ and $\pi_1$ are set at the fairly large value of 0.95, the posterior means are even larger, suggesting that the high estimated persistence is not an artifact of our priors; the data have the effect of revising upwards the probability of staying in a given state. If weaker priors are used, these probabilities are revised even further upwards.

Figure 2 plots observed prices with the fitted prices generated by the Markov preferences model: the improvement over the fit of the fixed preference model in Figure 1 is striking. The square root of the mean squared error (RMSE) of the fitted log stock price series is reduced from 0.21 to 0.10, and the coefficient of correlation between the fitted and actual series increases from 0.65 to 0.93. An interesting feature of our model is that within a given phase, its predictions look much like those of the standard model; the improvements are almost entirely due to the model’s ability to track the sharp price movements around the oil price shock and at the beginning of the bull market in the late 1980’s.

Even though our model is able to reproduce the strong cyclical movements in stock prices, it is also able to capture the weak variability of bond prices as well. The mean of the fitted log bond price series (-0.0019) is within the range of the observed value (-0.0014), and the standard deviation of the predicted series (0.0036) is close to that of the observed series (0.0039). The RMSE of the fitted log bond price series is 0.0045, and the coefficient of correlation between the actual and fitted series is 0.31.

Figure 3 plots the posterior probability that a given observation occurred during a bull market. We noted earlier that our results suggest that bear markets prevailed for roughly a third of our sample; Figure 3
indicates there were two or perhaps three bear markets during the period 1960:3–1992:9. The first episode occurs during the four-month period 1970:5–1970:8, when there is a sharp drop in the bull market probability, although the fact that the lowest probability is 0.77 suggests that this period is best classified as a bull market, albeit a weak one. After that, the first clear bear market starts in 1974:7 and ends 17 months later in 1975:12, while the last bear market of the sample begins in 1977:4, lasting for over 8 years before ending in 1985:11.

It is also interesting to compare our financial cycle chronology with the University of Michigan index of consumer expectations (dashed line) and the NBER reference business cycle recessionary periods (shaded regions). In the first half of the sample, the three series display a high degree of coherence, but from 1976 on, the relationship between the three indices appears to weaken: the bear market of 1977-85 persists through two recoveries, and the bull market that began in 1985 continues right through the sharp recession of 1990-91 and on to the end of the sample. Since the asset price series in Figures 1 and 2 remains low until 1985 and does not appear to be greatly affected by the 1990-91 recession, this puzzle does not indicate an obvious problem with the model’s ability to explain the data.

4 Implications for Asset Returns: Itô Preferences

4.1 Theoretical Model

We noted in the Section 3 that the Markov preference model reduces the RMSE of the standard model by half, and that this improvement is almost entirely due to its ability to reproduce a financial market cycle with phases that last 8 to 12 years. If the model is to reproduce movements in the data at higher frequencies, more flexibility is needed. We find that a continuous-time framework in which risk aversion follows a diffusion process is a convenient framework that allows for a great deal of short-run variability, and that permits us to develop some insights into the implications of contingent preferences for asset returns.

Let lower-case letters denote natural logarithms, and let $X_j(t)$ denote the total real value (including any dividends) of asset $j = 0, 1, \ldots, n$. Furthermore, assume that $c(t)$, $\gamma(t)$, and $x_j(t)$ follow joint Itô processes.
Then, (2.4a) and (2.4b) imply [see Appendix B] that mean excess returns and the risk-free rate must satisfy:

\[ E(t) [dx_l(t) - dx_0(t)] = [c(t) - \theta]\sigma_{/t_l}(t) + \gamma(t)\sigma_{/t_l i}(t) - 0.5\sigma_{/t_l i i}(t) \]  

(4.1a)

and,

\[
dx_0(t) = -\log(\beta)dt - E(t)[du_q(t)] - 0.5\sigma_{uu}(t) \\
= -\log(\beta)dt + \gamma(t)E(t)[dc(t)] + \gamma(t) - \theta]E(t)[d\gamma(t)] + \{1 - \gamma(t)c(t) - \theta\}\gamma_{/t_l}(t) \\
-0.5\gamma(t)^2\sigma_{/t_l /t_l c}(t) - 0.5[c(t) - \theta]^2\gamma_{/t_l /t_l c}(t) 
\]

(4.1b)

where \(\sigma_{/t_l}(t) \equiv E(t)[dc(t)dx_l(t)]\), \(\sigma_{/t_l i}(t) \equiv E(t)[d\gamma(t)dx_l(t)]\), denotes the covariance between returns and instantaneous changes in log consumption and risk aversion; \(\sigma_{/t_l i i}(t) \equiv E(t)[dx_l(t)^2]\) is the risky returns’ variance; \(\sigma_{/t_l /t_l c}(t)\) is the covariance between innovations in risk aversion and log consumption, and \(\sigma_{/t_l /t_l c}(t)\) and \(\sigma_{/t_l /t_l c}(t)\) are the variance of instantaneous changes in risk aversion and log consumption respectively. Consider further the standard unconditional iso-elastic pricing equations obtained by imposing \(\gamma(t) = \gamma, \forall t:\)

\[ E(t) [dx_l(t) - dx_0(t)] = \gamma\sigma_{/t_l i c}(t) - 0.5\sigma_{/t_l i i}(t), \]  

(4.2a)

\[ dx_0(t) = -\log(\beta)dt + \gamma E(t)[dc(t)] - 0.5\gamma^2\sigma_{/t_l /t_l c}(t) \]  

(4.2b)

The empirical difficulties of the standard model with homoscedastic innovations are well-known. Firstly, since estimates for the covariance term \(\sigma_{/t_l i}(t)\) in (4.2a) are so small, the coefficient of relative risk aversion \(\gamma\) must be increased to unrealistic levels to reproduce the mean equity premium \(E(t)[dx_l(t) - dx_0(t)]\) (the equity premium puzzle). Secondly, mean excess returns vary over time, instead of being constant (the predictability puzzle). Finally, since high risk aversion is necessary to justify the risk premium, and since mean consumption growth \(E(t)[dc(t)]\) is also relatively high, the subjective discount rate must be negative (i.e. \(\beta > 1\)) to reproduce the historically low risk-free rates \(dx_0(t)\) (the risk-free rate puzzle).
Introducing state contingencies does not alter the fundamental result of the C-CAPM: the representative agent allocates savings to hedge against fluctuations in \( du_c(t) \). There are, however, some notable differences. In (4.1b), the concavity variance term \( \sigma_{\gamma\gamma}(t) \) increases the volatility of the MRS. Moreover, if risk aversion is counter- (pro-) cyclical, and marginal utility is high (low), then small shocks to consumption are associated with large fluctuations in MRS’s, for reasons discussed in Section 2. This result is reflected in the expression for the risk-free rate, through the term \( \{1 - \gamma(t)[c(t) - \theta]\} \sigma_{\gamma\gamma}(t) \). Both the additional volatility and the cyclical nature of \( \gamma(t) \) induce a precautionary demand for the risk-free asset and therefore reduce its return.

Secondly, it is typically observed that \( \sigma_{ic}(t) > 0 \), i.e. low returns are associated with adverse consumption innovations. For high (low) marginal utility, if returns and risk aversion are negatively (positively) correlated, then \( [c(t) - \theta]\sigma_{\gamma i}(t) > 0 \) in (4.1a). Hence, holding the risky asset entails a larger risk to marginal utility due to the coincidence of low returns, low consumption and high risk aversion, which justifies the high observed premia without having to inflate risk aversion to unrealistic levels. This extra risk is the distinguishing feature of our state-contingent risk preference specification. For example, the preference state variable used in Campbell and Cochrane (1995) is a deterministic function of consumption, so all IMRS risk must be explained by the consumption covariances. Not allowing any independent sources of IMRS risk implies that the slow-moving habit model will be unable to explain observed equity premia with reasonable levels of relative risk aversion.

Finally, notice that even if innovations are homoscedastic, the risk premium is time-varying due to the presence of log consumption as loading for the concavity risk. Since consumption is a smooth series, this may explain why ARCH-type scedastic structures have enjoyed widespread success.

4.2 Empirical Model

Suppose for now that \( \gamma(t) \) is an observable process, and that there is only one risky asset \( p \). Define \( x_e(t) \equiv \int_0^t [dx_p(\tau) - dx_0(\tau)]d\tau \) as the integral measure of excess returns, and suppose that \( Y(t) \equiv [x_e(t)', c(t), \gamma(t)']' \)
is an Itô process following an arithmetic Brownian motion:

\[ dY(t) = [AY(t) + b]dt + \sigma dZ(t) \]  

(4.3)

where \( A = [a_{ij}] \) for \( l, j = p, c, \gamma \) is a square matrix; \( b = [b_j] \) is a column vector; \( \sigma = [\sigma_j] \) is a diagonal matrix; and finally \( dZ(t) = [d\tilde{Z}_j] \) is a vector of correlated Wiener processes, with \( E(t)[d\tilde{Z}_i d\tilde{Z}_j] = \rho_{ij} dt \) and \( \rho_{ij} \equiv \sigma_{ij}/(\sigma_i \sigma_j) \) is the correlation coefficient.

Under these set of assumptions, the restrictions implied by the model are \( a_{pp} = 0, a_{pc} = \sigma_{pr}, a_{pc} = \sigma_{pc}, b_p = -\theta \sigma_{pr} - 0.5 \sigma_{pp} \). The risk premium will then have the same form as in (4.1a); the only difference being that the covariance terms are no longer functions of time:

\[ E(t) [dx_p(t) - dx_0(t)] = [c(t) - \theta] \sigma_{pr} + \gamma(t) \sigma_{pc} - 0.5 \sigma_{pp} \]  

(4.4)

We note two difficulties for estimating this model: the risk aversion series \( \gamma(t) \) is not observed, and our data are not in a continuous-time format. Fortunately, these problems can be dealt with by making use of a MCMC data augmentation algorithm based on the exact likelihood function for time-averaged data. Suppose that \( Y(t) \) satisfies (4.3) and define \( Y_t \equiv \int_{t-1}^{t} Y(s) ds \). It can then be shown\(^3\) that the discrete time series process \( \{Y_t\} \) satisfies

\[ Y_t = \phi Y_{t-1} + g + \eta_t \]  

(4.5)

where \( \phi \equiv e^A \), \( g \equiv (e^A - I)A^{-1}b \), where \( e^A \) is the matrix exponential of \( A \) and where the error vector \( \eta_t \) is a multivariate MA(1) process. The results below make use of the techniques developed in Gordon and St-Amour (1998) for this class of models.

An identification issue must also be addressed. It is noted in Gordon and St-Amour (1998) that the scale parameter \( \theta \) is not identified in the model defined by (4.3) and (4.4); variations in \( \theta \) simply generate values for \( \sigma_{pr} \) that produce similar estimates for the term \( [c(t) - \theta] \sigma_{pr} \); the other terms in (4.4) are unaffected. In

\(^3\)Bengtstrom (1984).
principle, \( \theta \) can be identified by the information provided by the expression for the risk-free rate in (4.1b); this is the approach taken in Section 3, where prices of risky and risk-less assets are modeled simultaneously. However, this approach does not seem promising in this context: since (4.1b) is a non-linear function of \( c(t) \), analytic expressions for the likelihood function are unavailable. In our results below, we simply fix \( \theta = 0 \).

4.3 Results

Locally diffuse priors based on those used in Gordon and St-Amour (1998) were used in estimating the fixed parameters and the risk aversion series. The estimation results are based on 2000 draws from the Markov chain described in Gordon and St-Amour (1998); 500 burn-in iterations were used to initialize the series. The parameter estimates are listed in Table 2, and Figure 4 plots the estimates for the realized values of \( \gamma_t \). These results are very similar to those obtained in Gordon and St-Amour (1998), who estimate the model for various pairs of risky assets.

<table>
<thead>
<tr>
<th>parameter</th>
<th>mean</th>
<th>std. deviation</th>
</tr>
</thead>
<tbody>
<tr>
<td>( a_{cp} )</td>
<td>9.77e-04</td>
<td>0.0036</td>
</tr>
<tr>
<td>( a_{cc} )</td>
<td>-0.0024</td>
<td>0.0005</td>
</tr>
<tr>
<td>( a_{cr} )</td>
<td>0.0384</td>
<td>0.0058</td>
</tr>
<tr>
<td>( a_{cp} )</td>
<td>-0.0395</td>
<td>0.3847</td>
</tr>
<tr>
<td>( a_{cc} )</td>
<td>0.1287</td>
<td>0.1080</td>
</tr>
<tr>
<td>( a_{cr} )</td>
<td>-4.0156</td>
<td>0.5512</td>
</tr>
<tr>
<td>( b_{c} )</td>
<td>-0.0025</td>
<td>0.0053</td>
</tr>
<tr>
<td>( b_{r} )</td>
<td>1.7679</td>
<td>0.7162</td>
</tr>
<tr>
<td>( \sigma_{p} )</td>
<td>0.0574</td>
<td>0.0020</td>
</tr>
<tr>
<td>( \sigma_{c} )</td>
<td>0.0229</td>
<td>0.0022</td>
</tr>
<tr>
<td>( \sigma_{r} )</td>
<td>2.0992</td>
<td>0.7848</td>
</tr>
<tr>
<td>( \rho_{pc} )</td>
<td>0.0554</td>
<td>0.0163</td>
</tr>
<tr>
<td>( \rho_{pc} )</td>
<td>0.0073</td>
<td>0.0047</td>
</tr>
<tr>
<td>( \rho_{cr} )</td>
<td>-0.9727</td>
<td>0.0099</td>
</tr>
</tbody>
</table>

Note: posterior means and standard deviations for parameters of Itô preferences model (4.5).

Our estimates for the risk aversion series fluctuate in a narrow band around a sample mean of 0.80, and our estimate for the sample standard deviation is 0.38. Both values are consistent with the scale and
variability of the risk aversion coefficients of the two-state model, suggesting again that comparatively small movements in risk aversion are able to generate enough movement in the IMRS to explain asset market behavior using plausible levels of risk aversion.

The main difference between the risk premium in (4.4) and that of the standard model is the term \( [c(t) - \theta] \sigma_p \), which characterizes the excess return required to compensate for the risk associated with movements in marginal utility generated by fluctuations in \( \gamma(t) \). If this expression is positive, then the usual consumption risk term \( \gamma(t) \sigma_{pc} \) is not forced to account for the entire risk premium. Since the posterior probability that \( [c(t) - \theta] \sigma_p > 0 \) is greater than 0.98, the model is able to fit excess returns with reasonable values of risk aversion. Indeed, since the consumption risk term is only 1% of the concavity risk term, it is perhaps unsurprising that the standard model fares so poorly.

Although the risk-free rate is not used in the formal estimation in order to identify \( \theta \), we conducted an ad hoc calibration exercise in which \( c(t) \) and \( \gamma(t) \) are approximated by their time-averaged values, and the parameter estimates in Table 2 are used to compute the various conditional expectations in (4.1b). Two free parameters remain, so we set \( \beta = 1.03^{-1/12} \), and searched for a value of \( \theta \) so that the sample mean of the fitted risk-free rate matches that of the data. This exercise produces a value of \( \theta = 8.8 \), which is remarkably close to the estimated value of 9.25 obtained in the Markov preference model.

In the two-state model, values of \( \theta \) in this range produced a pattern of counter-cyclical risk aversion. In the results in Table 2, innovations in risk aversion have a strong negative correlation with consumption shocks, and when \( \theta \) is set equal to zero, the model generates a weak positive correlation between price and risk aversion innovations. However, if \( \theta \) is set so that \( c(t) - \theta \) is negative throughout the sample, the estimated correlation coefficient must be negative in order to ensure that concavity risk term \( [c(t) - \theta] \sigma_{pc} \) remains positive; Gordon and St-Amour (1998) note that this is precisely what happens when the consumption data are rescaled. Since negative values of \( c(t) - \theta \) are consistent with the risk-free rates we observe, the most plausible conclusion is that increases in risk aversion, which are strongly negatively correlated with consumption, are also negatively correlated with stock price movements.
5 Conclusion

This study shows that the addition of contingent risk preferences to the original C-CAPM model is able to capture many of the salient features of financial market data. The gains stem chiefly from the additional volatility to the IMRS that state dependencies are able to provide.

We study stock and bond prices using a simple two-state Markov process for the parameter associated with risk aversion. Our estimates indicate that small but significant movements in the agent’s concavity index are able to generate the sharp swings in asset prices that are characteristic of financial market data. These shifts are infrequent, and both states display a high degree of persistence. For much of the sample, our estimated chronology of bull and bear markets displays interesting parallels with consumer confidence indices and with the business cycle. A noteworthy finding is that in the latter part of our sample (1977–1992) there is less coherence between the bull and bear markets of the financial market and the expansions and recessions of the goods market.

We also study the other extreme where concavity follows a diffusion process. This approach highlights the potential of contingent preferences in solving the empirical anomalies associated with excess returns and the risk-free rate. Estimation results support this conjecture, with realistic estimates for risk aversion when the model is estimated with market equity premium data. Perhaps even more significantly, estimates for the preference parameters are remarkably similar to those obtained under the simpler Markov preferences model. This suggests that our state-dependent risk aversion model is robust to the choice of the law of motion for the preference state variable.

This paper is more descriptive than explanatory, and it raises perhaps more questions than it answers. For example, how and why the financial cycle lost much of its coherence with the business cycle in the mid-1970’s is an empirical puzzle that should be addressed in future research. Furthermore, this study is agnostic about why preferences might shift: potential explanations include movements in wealth or shifts in relative prices. Future work should address the issue of determining the factors that underlie the movements in risk preferences that are identified in this paper.
A Derivation of Markov preferences pricing kernels

A given realization of the state variable \( S_t \) affects equilibrium asset prices through the (quasi reduced-form) demand schedules \( P_{i,t} = P(C_{i,t}, D_{i,t}; S_t) \); we denote \( P^i_{i,t} \) as the observed equilibrium price for asset \( i \) given the realized state \( S_t = s \in \{0, 1\} \). Substituting the transition probabilities in the Euler equations reveals that risky asset prices must satisfy:

\[
1 = \beta \left\{ \pi_0 E_t \left[ \left( \frac{P^0_{i,t+1} + P^i_{i,t+1}}{P^0_{i,t}} \right) \left( \frac{\Theta(1, C_{i,t+1}) - \gamma}{\Theta(1, C_{i,t}) - \gamma} \right) \right] + (1 - \pi_0) E_t \left[ \left( \frac{P^0_{i,t+1} + P^i_{i,t+1}}{P^0_{i,t}} \right) \left( \frac{\Theta(1, C_{i,t+1}) - \gamma}{\Theta(1, C_{i,t}) - \gamma} \right) \right] \right\},
\]

(A.1)

while bond prices are obtained by:

\[
1 = \beta \left\{ \pi_0 E_t \left[ \left( \frac{1}{P^0_{i,t}} \right) \left( \frac{\Theta(1, C_{i,t+1}) - \gamma}{\Theta(1, C_{i,t}) - \gamma} \right) \right] + (1 - \pi_0) E_t \left[ \left( \frac{1}{P^0_{i,t}} \right) \left( \frac{\Theta(1, C_{i,t+1}) - \gamma}{\Theta(1, C_{i,t}) - \gamma} \right) \right] \right\},
\]

(A.2)

Substitute the candidate solutions \( P^i_{i,t} = B^i_{i,t} (\Theta^{-1} C_{i,t})^{-\gamma} \), in equations (A.1)–(A.2), and use the log-normal properties to solve for \( B^i_{i,t} \) and \( B^0_{i,t} \) as in (3.2b)–(3.2d).

B Derivation of continuous-time returns

Rewrite the discrete-time Euler equations for the risky (2.4a), and the risk-free asset (2.4b) as:

\[
1 = \beta^h E_t \exp \left[ \Delta u_{c,h} + \Delta x_{j,t+h} \right],
\]

(B.1)

where \( \Delta x_{j,t+h} \) is the change in the real value (cum-dividends) of asset \( j = 0, 1, \ldots, n \). Take the limit \( h \to 0 \), and use the properties of Wiener processes to obtain:

\[
E(t)[dx_j(t)] = -\log(\beta) dt - E(t)[du_j(t)] - 0.5 \text{VAR}(t)[du_j(t)] - 0.5 \text{VAR}(t)[dx_j(t)] - \text{COV}(t)[du_j(t), dx_j(t)]
\]

\[
= -\log(\beta) dt - E(t)[du_j(t)] - 0.5 E(t)[du_j(t)^2] - 0.5 E(t)[dx_j(t)^2] - E(t)[du_j(t)dx_j(t)]
\]

(B.2)

Next, from preferences (2.3) it follows directly that \( c(t) \) and \( \gamma(t) \) being Itô processes is sufficient for \( u_c(t) \) being Itô, with diffusion:

\[
du_c(t) = -[c(t) - \theta dy(t) - \gamma(t) dx(t)] - E(t)[dy(t)dx(t)].
\]

Substitute in (B.2) and solve for mean excess returns \( E(t)[dx_i(t) - dx_0(t)] \) in (4.1a) and the risk-free rate \( dx_0(t) \) in (4.1b).

C Computing Equilibrium Prices

The solution\(^4\) to the difference equation (3.2b) can be written as

\[
\begin{bmatrix}
B^0_{i,t} \\
B^1_{i,t}
\end{bmatrix}
= \sum_{j=1}^{K-1} \beta^j \pi^j \begin{bmatrix}
M^0_{i,t+j} \\
M^1_{i,t+j}
\end{bmatrix} + \sum_{j=K}^{\infty} \beta^j \pi^j \begin{bmatrix}
M^0_{i,t+j} \\
M^1_{i,t+j}
\end{bmatrix}
\]

(C.1)

---

\(^4\)We assume that the relevant transversality condition rules out bubbles.
where $M_{t+j}^i$ is defined as in (3.2c). For a given value of $K$, we can compute the sequence \( \{\mu_{t+t+j}, \mu_{t+t+j}\}_{j=1}^K \) for each $t$ by repeated application of the VAR model in (3.4) for dividend and consumption growth rates in order to compute the partial sum $\sum_{j=1}^{K-1} \beta^j \pi^j M_{t+j}^i$.

The remainder term is not so easily dealt with. One approach would be to set $K$ to be large enough so that $\sum_{j=K}^{\infty} \beta^j \pi^j M_{t+j}^i \simeq t$, but since $\beta$ is close to one and since dividends and consumption grow over time, $K$ must be set to an extremely large value. Since (C.1) must be evaluated at each evaluation of the likelihood, this technique does not appear to be feasible given our current computing facilities.

Instead of choosing $K$ such that the remainder term is zero, we set $K \rightarrow \infty$, so that $\pi^{K+j} \simeq \pi$, $M_{t+k}^i \simeq M_{t+k}^i e^{j(\mu_{t+k}+\gamma_{t+k})}$ and $M_{t+k+k}^i \simeq M_{t+k+k}^i e^{j(\mu_{t+k+k}+\gamma_{t+k+k})}$ for $j = 0, 1, \ldots$. Since $\pi$ is the transition matrix of an ergodic Markov chain, $\pi^t$ converges to a constant matrix of the unconditional state probabilities. If the VAR process in (3.4) is stationary (a condition we impose in the estimation, even though this constraint does not appear to be binding), then predictions for dividend and consumption growth rates will converge to their unconditional means $\mu_{t+k}$ and $\mu_{t+k}$, respectively.

The appropriate choice for $K$ depends on the desired level of precision and the rate of convergence of the VAR and Markov state processes. Let $\lambda_2$ denote the second-largest eigenvalue (since the process is ergodic, the largest eigenvalue is equal to one) of $\pi$, and let $\lambda_1$ and $\lambda_2$ denote the eigenvalues of the matrix of autocorrelation coefficients in (3.4), and let $\lambda = \max\{|\lambda_1|, |\lambda_2|\}$. Given the precision of our data (three or four significant digits), $K$ is set so that $\lambda^K < 0.00001$. In the estimation, $K$ is reset to satisfy this condition at every evaluation of the likelihood.

Given $K$ and the growth rates, the terms in the remainder are simply the sums of geometric series. Adding this term to the partial sum yields the value of $B_{t+i}$ used in the estimation.

## D Estimation Details

Estimates for the model (3.6) are based on draws generated by iterating through the following sequence, which combines the Gibbs sampler with data augmentation:

\[
\begin{align*}
\beta^i & \sim \mathcal{P}(\beta|\gamma_0^i, \gamma_1^i, \phi^i, \Sigma^i, \Omega^i, S_0^i, S_1^i, p, q, d, c) \quad (D.1) \\
(\gamma_0^i, \gamma_1^i, \theta^i) & \sim \mathcal{P}(\gamma_0, \gamma_1, \theta|\beta^i, \phi^i, \Sigma^i, \Omega^i, S_0^i, S_1^i, p, q, d, c) \quad (D.2) \\
\phi^i & \sim \mathcal{P}(\phi|\beta^i, \gamma_0^i, \gamma_1^i, \theta^i, \Sigma^i, \Omega^i, S_0^i, S_1^i, p, q, d, c) \quad (D.3) \\
\Sigma^i & \sim \mathcal{P}(\Sigma|\beta^i, \gamma_0^i, \gamma_1^i, \theta^i, \phi^i, \pi^i, \Omega^i, S_0^i, S_1^i, p, q, d, c) \quad (D.4) \\
\pi^i & \sim \mathcal{P}(\pi|\beta^i, \gamma_0^i, \gamma_1^i, \phi^i, \Sigma^i, \Omega^i, S_0^i, S_1^i, p, q, d, c) \quad (D.5) \\
\Omega^i & \sim \mathcal{P}(\Omega|\beta^i, \gamma_0^i, \gamma_1^i, \phi^i, \Sigma^i, \pi^i, S_0^i, S_1^i, p, q, d, c) \quad (D.6) \\
S_0^i & \sim \mathcal{P}(S|\beta^i, \gamma_0^i, \gamma_1^i, \phi^i, \Sigma^i, \pi^i, S_1^i, p, q, d, c) \quad (D.7) \\
S_1^i & \sim \mathcal{P}(S|\beta^i, \gamma_0^i, \gamma_1^i, \phi^i, \Sigma^i, \pi^i, \Omega^i, S_0^i, p, q, d, c) \quad (D.8)
\end{align*}
\]

In the Gibbs sampler steps (D.1)-(D.7), parameter values are simulated from their “full conditional” distributions, i.e., conditional on the observed data, the states $S$ and the other parameters. The data augmentation step (D.8) is the key element of the algorithm, since it avoids the computational burden of direct evaluation of the likelihood function, which would involve integrating $S$ out of the data density (3.6).

Under fairly weak conditions that are satisfied in this application, the sequence of draws generated according to (D.1)-(D.8) forms an ergodic Markov chain whose stable distribution is the joint posterior distribution (3.8). Once the chain has converged to its stable distribution, it produces a sequence of correlated draws from (3.8). Given a sufficiently large number of draws, the posterior moments of the various parameters can be estimated with an arbitrarily high degree of accuracy.

---

5See Roberts and Smith [1994]. The algorithm satisfies the sufficient condition that each of the full conditional distributions (D.1)-(D.8) have positive mass everywhere in the admissible region.
It is possible to simulate draws for two parameters using standard techniques. Since \( \Omega \) appears only in (3.5), its full conditional distribution can be derived using well-known results for the bivariate normal model with known variance and a conjugate inverse-Wishart prior. If the rest of the parameters and the states are known, we can recover the series \( u_{p,t} \) and \( u_{q,t} \) from (3.5). Given these values and the inverse-Wishart prior, the full conditional posterior for \( \Omega^{-1} \) is also an inverse-Wishart density; see Poirier (1995, p. 300).

Similarly, since \( S_0 \) appears only in \( P(S|\pi, S_0) \), the full conditional density for \( S_0 \) makes use only of the realized value for \( S_1 \) and the transition probabilities. Given the prior \( P(S_0) \) and \( S_1 \) the full conditional probability the \( S_0 = 1 \) is obtained by using Bayes’ rule; see Albert and Chib (1993).

### D.1 “Metropolis-within-Gibbs”

Estimation of the other parameters is complicated by the fact that they appear in two of the three data densities that form (3.6). Although direct draws are not available, it is possible to implement a fairly simple Metropolis-Hastings algorithm (MHA) by exploiting the structure of our model.

It is perhaps useful to provide a brief summary of the MHA; a more detailed survey is provided by Chib and Greenberg (1995). Suppose that we wish to simulate draws from a “target density” \( f(\lambda) \), and define \( q(\lambda, \lambda') \) to be a known density from which it is easy to simulate a candidate \( \lambda' \) given \( \lambda \), the value generated by the previous iteration if the algorithm.

Define

\[
\alpha(\lambda, \lambda') \equiv \max \left[ \frac{f(\lambda')q(\lambda', \lambda)}{f(\lambda)q(\lambda, \lambda')}, 1 \right]
\]  

(D.9)

The MHA returns the candidate \( \lambda' \) with probability \( \alpha \); if \( \lambda' \) is rejected, then the MHA returns \( \lambda \). Typically, some extra work is required in order to identify suitable forms for \( q(\lambda, \lambda') \); see Chib and Greenberg (1995).

In estimating our model, we are able to make use of the special case in which the target density can be written as the product of two densities \( g \) and \( h \), so that \( f(\lambda) = g(\lambda)h(\lambda) \). Suppose further that we can simulate draws from \( h \), and set \( q(\lambda, \lambda') = h(\lambda') \). Substituting into (D.9) yields

\[
\alpha(\lambda, \lambda') \equiv \max \left[ \frac{g(\lambda')}{g(\lambda)}, 1 \right]
\]  

(D.10)

Here, \( h \) is used to generate new candidates, and \( g \) is used to decide whether or not they are to be accepted.

In our application, we make use of this special case in simulating draws for \( \beta, \phi, \Sigma \) and \( \pi \):

- **\( g(\beta) \):** Proportional to \( P(p, q|d, c, S, \beta, \gamma_0, \gamma_1, \phi, \Sigma, \pi, \Omega) \) considered as a function of \( \beta \) and conditional on everything else.

- **\( h(\beta) \):** \( P(\beta) \)

NOTE: \( \beta \) appears only in (3.5), and the tight prior for \( \beta \) plays a relatively important role in the posterior.

- **\( g(\phi) \):** Proportional to \( P(p, q|d, c, S, \beta, \gamma_0, \gamma_1, \phi, \Sigma, \pi, \Omega) \) considered as a function of \( \phi \) and conditional on everything else.

- **\( h(\phi) \):** \( P(\phi|\Sigma, d, c) \propto P(d, c|\phi, \Sigma)P(\Sigma) \)

NOTE: The full conditional distribution of \( \phi \) is proportional to the product of the prior and the data densities \( P(p, q|d, c, S, \beta, \gamma_0, \gamma_1, \theta, \phi, \Sigma, \pi, \Omega) \) and \( P(d, c|\phi, \Sigma) \). Given \( \Sigma \), the factor \( h(\phi) \) is simply the posterior for the slope parameters of a SUR model with a known covariance matrix. Chib and Greenberg (1996, section 3.1) note that \( h(\phi) \) is a multivariate normal, and they provide its mean and variance.
whether or not the candidate is accepted is whether or not the new candidate helps the asset price series.

As we noted in Section 3.3, simulating the latent variable $S$ in (D.8) allows us to estimate the model without direct evaluation of the likelihood function. Albert and Chib (1993) note that the full conditional distribution for a given $S_i$ is obtained by applying Bayes’ rule:

$$P(S_t = 1|\beta, \gamma_0, \gamma_1, \theta, \phi, \Sigma, \pi, \Omega, S_0, p, q, d, c) = \frac{P(S_t=1|S_{t-1})P^*(p_t, q_t|S_t=1)P(S_{t+1}|S_t=1)P(S_t=0|S_{t-1})P^*(p_t, q_t|S_t=0)P(S_{t+1}|S_t=0)}{P(S_t=1|S_{t-1})P^*(p_t, q_t|S_t=1)P(S_{t+1}|S_t=1)P(S_t=0|S_{t-1})P^*(p_t, q_t|S_t=0)}$$

(D.11)

where $P^*(p_t, q_t|S_t = 0)$ and $P^*(p_t, q_t|S_t = 1)$ are the densities for the observed $(p_t, q_t)$ pair, given the rest of the model and evaluated at each preference state. Given values for $S_{t-1}$ and $S_{t+1}$, the terms $P(S_t = 1|S_{t-1})$, $P(S_{t+1}|S_t = 1)$, $P(S_t = 0|S_{t-1})$ and $P(S_{t+1}|S_t = 0)$ are the appropriate elements of the transition matrix $\pi$.

The sequence $\{S_t\}_{t=1}^{T-1}$ is generated by recursive simulation from (D.11) for $t = 1, 2, \ldots, T - 1$. Values for $S_T$ are simulated from

$$P(S_T = 1|\beta, \gamma_0, \gamma_1, \theta, \phi, \Sigma, \pi, \Omega, S_0, p, q, d, c) = \frac{P(S_T=1|S_{T-1})P^*(p_T, q_T|S_T=1)P(S_{T+1}|S_T=1)P(S_T=0|S_{T-1})P^*(p_T, q_T|S_T=0)}{P(S_T=1|S_{T-1})P^*(p_T, q_T|S_T=1)P(S_{T+1}|S_T=1)P(S_T=0|S_{T-1})P^*(p_T, q_T|S_T=0)}$$

(D.12)
References


Figure 1: Actual and fitted (state-independent preferences) prices

Note: Logarithm of actual S&P composite stock price (dotted line) and fitted stock prices (solid line) for state-independent model $P_t = B_t C_t^\gamma$; $B_t = \beta [E_t B_{t+1} + M_t]$; $M_t \equiv E_t \exp(d_{t+1} - \gamma c_{t+1})$. Log consumption and dividends follow the VAR(1) process (3.4).
Note: Logarithm of actual S&P composite stock price (dotted line) and fitted stock prices (solid line) for Markov preferences model (3.4) and (3.5).
Note: Posterior probability that $S_t = 1$ (solid line); University of Michigan's index of consumer expectations (dashed line), and NBER's index of recessions (shaded areas).
Figure 4: Estimated Risk Aversion Coefficients: Itô Preferences