Belief-dependent Utilities, Aversion to State-Uncertainty and Asset Prices*

Pietro Veronesi

University of Chicago, NBER and CEPR

January 8, 2004

Abstract

This paper introduces the concepts of *"belief dependent" utility functions* and *aversion to "state-uncertainty,"* and it shows that these preferences help to explain both the unconditional and the conditional properties of asset returns. To solve for asset prices and returns under general conditions, the paper also develops a discretization methodology to obtain approximate analytical solutions. In a parsimonious parametrization that is consistent with habit utility, the model generates unconditional moments for asset returns that closely match the empirical ones. Finally, the estimated time-variation in the *dispersion* of the conditional posterior distribution on the drift rate of consumption implies a pattern of *conditional* return volatility that matches the properties of the stock index volatility.

JEL Classification Number: D81, G12

Keywords: State-dependent Preferences, Uncertainty Aversion, Asset Pricing

*I thank Federico Bandi, John Y. Campbell, Micaela Della Torre, Paolo Ghirardato, Lars P. Hansen, Fulvio Ortu, Nick Polson, Tano Santos, Costis Skiadas, Per Stromberg, Fernando Zapatero and seminar participants at Harvard University, University of Chicago, University of Southern California, University of Minnesota, University of Amsterdam, Tilburg University, Universita' della Svizzera Italiana at Lugano, HEC Lausanne, the 2001 NBER Asset Pricing meeting in Los Angeles and the 2001 SAET conference in Ischia for comments and suggestions. Special thanks to Marciano Siniscalchi for the lengthy conversations on many of the issues discussed in this paper. Of course, errors are my own. This work is supported by the James S. Kemper Faculty Research Fund at the Graduate School of Business, the University of Chicago. Some results contained in this paper also circulated under the title "Relative Performance under Bayesian Uncertainty." Contact information: Pietro Veronesi, University of Chicago, 1101 E. 58th St., Chicago, IL, 60637, USA. Email: pietro.veronesi@gsb.uchicago.edu.

1 Introduction

This paper introduces the concepts of *"belief dependent" utility functions* and *aversion to "state-uncertainty"* and it applies them to a standard pure-exchange economy to show that this type of preferences helps to explain the various stylized facts of stock returns, including a high equity risk premium, a low risk-free rate, a high return volatility, and volatility clustering.

In a nutshell, a "belief-dependent" utility function is a generalization of the more common "state-dependent" utility function, where the "state" is not fully observable. For example, consider the case of utility functions that depend on the agent's health status, which is a standard setting to study issues related to health and life insurance (see e.g. Zeckhauser (1970, 1973), Arrow (1974), Viscusi and Evans (1990)). For a whole range of diseases or risk factors, such as diabetes, high cholesterol etc., it is more reasonable to assume that agents may not be fully aware of their health status, because the diagnostic techniques are imprecise and check-ups infrequent (see e.g. Cutler and Richardson (1997, p. 253)). In this case, agents possess a probability distribution on their own health status and preferences become "belief-dependent:" Changes to this subjective probability distribution lead to changes in the utility and marginal utility of income or consumption.

Similarly, the recent literature in macroeconomics and finance has focused on state-dependent utilities to explain the behavior of individual consumers/investors and of financial variables. Examples include the works on habit formation (e.g. Sundaresan (1989), Constantinides (1990), Abel (1990), Campbell and Cochrane (1999)), relative social standing (Bakshi and Chen (1996)), stochastic subsistence consumption levels (e.g. Campbell and Viceira (2002)) and loss aversion (e.g. Barberis, Huang and Santos (2000)). In this literature the "state" is always assumed perfectly observable although in many cases it is more realistic to assume that it is only partially observable: For example, when preferences depend on the relative performance with respect to a reference class of agents (as in *external* habit formation models), the lack of knowledge of other agents' income/consumption levels naturally leads to a belief-dependent representation of agents' preferences. Clearly, state-dependent utilities are then recovered as a special case in which agents have a degenerate probability distribution.

In this article, I argue that a re-interpretation of standard representation theorems available in the decision theory literature provides a specific form for belief-dependent utilities – subjective beliefs must enter linearly in the utility functional, as in the standard expected utility framework. In addition, however, the same representation results induce a notion of "*aversion to state-uncertainty*," that is, the aversion to a more diffuse distribution on the unknown state. Intuitively, consider again the case of health-dependent utility and suppose that an agent assigns equal probabilities to the three scenarios that a non life-threatening illness is "serious," "mild" or "does not exist." If upon a doctor visit the agent learns that the illness is only "mild" and he is happy because he can now enjoy current consumption more, then we may think of him as being averse to state-uncertainty.[1] In other words, state-dependency may affect the utility in a non-linear fashion so that a change in the *dispersion* of the probability distribution on the "state" would bring about a change in the agent's utility. If such changes in the dispersion of the probability distribution do not affect the agent's utility, then he/she is neutral to state-uncertainty. I also characterize belief-dependent utility functions with constant relative *risk* aversion and constant aversion to state uncertainty.

I then apply this type of preferences to a pure exchange economy with incomplete information, where for generality I first leave unspecified the nature and the process of the unobservable "state" affecting investors' preferences. Using a new discretization approach, I obtain analytical expressions for prices and returns for both bonds and stocks. These solutions allow me to find general conditions under which aversion to state uncertainty yields higher expected returns and volatility, and lower interest rates. Indeed, these effects take place for example when the "state" is a procyclical variable but inversely related to the marginal utility of consumption, a situation that occurs typically in external habit formation models (see e.g. Campbell and Cochrane (1999)). In addition, however, my model uncovers an additional channel that leads to the time variation of the conditional moments of asset returns, namely, the time variation in the uncertainty about the "state" – the "habit" in the case of habit preferences. This channel is important and it underscores why belief-dependent utilities generate a much more volatile stochastic discount factor than standard state-dependent preferences: since the marginal utility of consumption depends on the whole distribution of beliefs, it is affected by the behavior

[1] Clearly, we should assume that the three health states have "equal distance" under some metric.

of the tails of this distribution. Due to the rational Bayesian updating of these beliefs, the tails of the distribution are more volatile than its mean, thereby increasing the volatility of the stochastic discount factor compared to the standard state dependent preferences.

The empirical section pushes the habit interpretation further. In addition, I assume that occasionally the (unobservable) drift rate of aggregate dividends may experience a discrete jump. From the rules of Bayesian updating, this assumption entails that as agents strive to interpret the new data, their uncertainty over the true drift rate of the economy suddenly increases. Examples where a discrete jump in the drift rate of the economy is likely to have occurred abound. For instance, the oil embargo at the beginning of the 70s, or the inflationary period of the late 70s, or the invasion of Kuwait in 1990 and the start of the Desert Storm military operations, are all events that could lead to a stall of the economy, and generate an abrupt (negative) change in its growth rate. The estimates in the empirical sections confirm that these events were characterized by a high probability of a discrete change in consumption drift and thus an increase in the dispersion of the posterior distribution on the underlying drift rate. Interestingly, so do the last two quarters of 2001.

In this setting, I show by Monte Carlo simulations that when calibrated to consumption data the model is able to replicate many of the stylized facts about the unconditional and conditional moments of asset returns. Moreover, the variation over time of the dispersion of agents' distribution yields a time-variation in expected excess return and volatility, consistently with the data. The estimates of the model also suggest that the dispersion of posterior beliefs about consumption drift is high during low growth periods – a finding that others have obtained using Survey data (see e.g. David and Veronesi (2000), Anderson et al. (2003)). This implies that expected returns and volatility of returns should be especially high during bad times, as suggested by the empirical literature (see e.g. Black (1976) and Schwert (1989)).

In addition to the literature on state dependent utilities, already cited earlier, a number of recent articles explore the implication for economics of belief-dependent preferences (see e.g. Geneakopolos et al. (1989), Caplin and Leahy (2001), Yariv (2002)). Although motivated by similar considerations, that is, the fact that beliefs per se may affect agents' well-being besides their impact on decision making, my approach differs substantially from these other articles, as they depart from the expected utility paradigm. In contrast, since in my model beliefs enter the utility function simply because the utility itself depends on an underlying unobservable states, my approach is consistent with the conventional expected utility framework. In addition, none of these other articles investigate the effects of aversion to state-uncertainty, which is the main topic of this paper. Indeed, on this last issue, the paper is also nominally related to the recent literature on "Knightian uncertainty" and asset prices (see e.g. Epstein and Wang (1995), Maenheut (1999), Hansen et. al (1999), Cagetti et al (2002), Epstein and Schneider (2003)). Following Gilboa and Schmeidler (1989, 1993), in this literature agents are endowed with families of prior distributions on a given state and use the max-min rule to take decisions. However, besides the use of the term "uncertainty" and the intuitive notion that "uncertainty" is *bad* and that agents prefer certainty to uncertainty, "state-uncertainty" is completely different from "Knightian uncertainty," as in my model agents are endowed with a unique belief distribution and they maximize a standard expected utility functional.

Finally, this paper is related to the literature on learning and asset pricing (e.g. Detemple (1986), Gennotte (1986), Dothan and Feldman (1986), Feldman (1986), Timmerman (1993), David (1997), Veronesi (1999, 2000), Brennan and Xia (2001)). None of these articles explores the points made here. In addition, the methodology proposed in this paper generalizes the results contained in many previous articles dealing with asset pricing and learning in Lucas economies. Specifically, I propose a general methodology to obtain pricing and return implications in the presence of a complex filtering problem and highly non-linear dynamics of hidden state variables. Based on the observation that even in the *simplest cases* the effective computation of asset prices still requires the use of numerical integration methods, and hence implicitly the approximation of the state space, this article shows that if we approximate the state space to begin with, asset prices can be solved for analytically even when the filtering problem is very complex. In essence, the discretization of the state space enables me to rewrite the posterior *density* as a posterior *probability distribution*, which can be shown to follow a vector linear process with stochastic volatility. This approximation turns out to be extremely tractable.

The article proceeds as follows: Next section introduces the concept of belief-dependent utility functions and aversion to state uncertainty. Section 3 introduces the asset pricing model and proposes a new discretization approach to obtain closed form solutions for asset prices. Section 4 obtains analytical formulas for stock and bond prices, and it characterizes the properties of stock returns and the interest rate under belief dependent preferences. Section 5 calibrates the model to consumption data. Section 6 concludes.

2 Belief-Dependent Utility Functions

In this first section, I set out the minimum notation necessary to understand the nature of belief-dependent utility functions and aversion to state uncertainty. To do so, it is convenient to recall the standard representation theorems for expected utility. The discussion is taken from Myerson (1991), whose axioms and representation theorems are contained in Appendix A. Let $\mathcal{C}$ be a set of prizes and Θ a set of states. A *lottery* $f : \Theta \longrightarrow \Delta(\mathcal{C})$ is a function assigning a probability distribution $\Delta(\mathcal{C})$ on $\mathcal{C}$ to each state $\theta \in \Theta$. For every event $S \subseteq \Theta$, let us denote by $\succeq_S$ a *conditional preference relation* on the set of lotteries on $\mathcal{C}$. The following holds:

Theorem 1: Let $\succeq_S$ satisfy the (standard) axioms listed in Appendix A. Then (a) there exists a *state-dependent* utility function $u : \mathcal{C}\times\Theta \longrightarrow R$ and a subjective conditional probability function $\pi\left(.|S\right)$ on Θ such that for all lotteries f and g ,

$$f \succeq_S g \iff \sum_{\theta\in S}\pi\left(\theta|S\right)\sum_{c\in\mathcal{C}} f\left(c|\theta\right)u\left(c|\theta\right) \geq \sum_{\theta\in S}\pi\left(\theta|S\right)\sum_{c\in\mathcal{C}} g\left(c|\theta\right)u\left(c|\theta\right) \tag{1}$$

(b) In addition, $\pi\left(.|S\right)$ is a conditional probability distribution on Θ with unit mass on the event S, that satisfies the rules of rational updating, that is, for every $S'' \subseteq S' \subseteq S$, with $S' \neq \emptyset$, Bayes law applies: $\pi\left(S''|S\right) = \pi\left(S''|S'\right)\pi\left(S'|S\right)$.

To better interpret the representation of preferences in (1), consider its specialization to the "constant" lotteries, that is such that $f\left(c|\theta\right) = 1$ for every $\theta \in \Theta$. If we denote such a lottery by $[c]$, then the application of the representation result (1) implies that

$$[c] \succeq_S [c'] \iff \sum_{\theta\in S}\pi\left(\theta|S\right)u\left(c|\theta\right) \geq \sum_{\theta\in S}\pi\left(\theta|S\right)u\left(c'|\theta\right)$$

Since $[c]$ is a constant act and agents "know" $u(c|\theta)$ for all θ, this representation is simply saying that even if an agent obtains a prize c, his/her "subjective" utility from "consuming" c is "belief dependent," in the sense that it depends on the whole subjective distribution $\pi(\theta|S)$ over θ. In other words, since the "uncertainty" over the lottery f may be resolved *before* the uncertainty over the underlying state of nature θ, this approach implies that agents have belief-dependent utility functions.

Examples:

(i) *Health-dependent utility functions* (see e.g. Zeckhauser (1970, 1973), Arrow (1974), Viscusi and Evans (1990)). Utility functions have been shown empirically to depend on agents health status, especially in the case of severe health effects. As mentioned in the introduction, for a large set of diseases and risk factors it is more reasonable to assume that agents have only partial information on their own health status, because diagnostic technologies are imprecise and check-ups infrequent. This generates a belief-dependent utility function as in (1) where θ is the health status and $\pi(\theta)$ the subjective probability of it.

(ii) *Relative Performance: External Habits and Relative Social Standing.* Recent literature in macroeconomics and finance has explored preferences that explicitly incorporate relative performance. This could be at the consumption level (e.g. Abel (1990), Campbell and Cochrane (1999)), income level (see e.g. Chan and Kogan (2002)) and wealth level (Bakshi and Chen (1996)). In all cases, agents are unlikely to have a precise information about the consumption/wealth level of all the other agents within a reference class, but most likely they only have a probability distribution on it. Again, in this case the utility function becomes belief-dependent.

Definition: A *belief dependent utility function* over an act f is given by

$$U(f,\pi) = \sum_{\theta\in\Theta} \pi(\theta) \sum_{c\in\mathcal{C}} f(c|\theta)\, u(c|\theta) \tag{2}$$

In particular, a belief dependent utility function over a prize c is

$$U(c,\pi) = \sum_{\theta\in\Theta} \pi(\theta)\, u(c|\theta) \tag{3}$$

Remarks: (i) Representation (2) stems from axiom about preferences on acts in a static environment. Although the same axioms cannot be invoked for the dynamic economy studied in Section 3, they provide the reason why beliefs enter linearly in (2). In addition, one can interpret the prizes $c \in \mathcal{C}$ as consumption streams, so that $u(c|\theta)$ would represent an intertemporal utility as of time $t = 0$. Although additional axioms are necessary to ensure dynamic consistency and rational learning (see e.g. Epstein and Schneider (2003)), it is reasonable to conjecture that this setting can be extended to a dynamic framework. (*ii*) As discussed in the appendix, the preferences' representation (1) is not unique as there are infinite utility-belief pairs that yield the same representation. Yet, Skiadas (1997) provides a set of axioms able to uniquely identify the conditional probability and the state-dependent utility function. (*iii*) Representation (1) and definition (2) assume that $\mathcal{C}$ and Θ are finite spaces. In the rest of the paper, however, I will assume for convenience that they are infinite (in fact, continuous) spaces. Similar representations exist for this case, although they require much heavier notation. See e.g. Fishburn (1970).

2.1 Aversion to State-Uncertainty

Characterization (3) naturally leads to a definition of aversion to "state-uncertainty." In fact, given a prize c we can vary the distribution π over Θ and obtain various levels of "utility." Of interest to us are the changes in the "dispersion" of the probability π while keeping its expected value constant. To this end, it is often used the concept of "mean-preserving spread" (see Ingersoll (1987)) to do comparative statics exercises. Assuming that Θ is a metric space, so that mean-preserving spreads are well defined, I define aversion to state uncertainty as follows:

Definition: *(a)* Let π and c be given. A belief-dependent utility $U(c, \pi)$ displays *aversion to state-uncertainty at c* if a mean preserving spread $\widehat{\pi}$ on the distribution π yields

$$U(c, \widehat{\pi}) < U(c, \pi) \tag{4}$$

(b) A belief dependent utility displays *aversion to state uncertainty* if (4) holds for all c.
(c) Similarly, a belief dependent utility function displays *neutrality to state uncertainty* if $U(c, \widehat{\pi}) = U(c, \pi)$ for all c.

Examples revisited: In the context of example (*i*), consider an agent who has been diagnosed a possibly bad, but not life threatening, disease. If news that rule out both a bad health but also the perfect health makes the agent happy, then he/she is averse to "health-state" uncertainty. In the context of the relative-performance example (*ii*) suppose an agent received a $5,000 payraise. If her "utility" from $5,000 decreases by learning that some of her peers received a $10,000 payraise, her utility is "state-dependent." Suppose now this agent ignores what her peers' payraises were. If shifts in her *beliefs* over these payraises change her "utility," then her utility is "belief-dependent." If finally she is happy to find out that everyone else also got $5,000 when she only believed that the *average* was $5,000, then she is averse to state uncertainty.

2.2 Constant Aversion to Risk and Aversion to State Uncertainty

In the context of belief-dependent utility functions we can define the usual notion of relative risk aversion $\gamma(\pi, c)$ as

$$\gamma(\pi, c) = -\frac{c\partial^2 U(c, \pi)/\partial c^2}{\partial U(c, \pi)/\partial c} \tag{5}$$

This is the analogous notion of relative risk aversion as in the case of state-independent utility function. Since for given distribution π, the utility function $U(c) = U(c, \pi)$ is a standard Von-Neuman Morgenstern utility function with respect to state-independent lotteries (of which the constant lotteries are a special case), formula (5) reflects the *local* curvature of the utility function that is necessary and sufficient to generate "aversion" to fair bets (in relative terms). Given its importance in finance applications, I now characterize the belief dependent utility function for the case of constant relative risk aversion:

Proposition 1: $\gamma(c, \pi) = \gamma$ constant if and only if there are functions $k_1, k_2 : \Theta \to \mathcal{R}$ with $k_2(\theta) > 0$ for all $\theta \in \Theta$ such that

$$U(c, \pi) = E_t[k_1(\theta)] + E_t[k_2(\theta)]\frac{c^{1-\gamma}}{1-\gamma} \tag{6}$$

Proof: See Appendix. ■

It is immediate to see from Jensen's inequality that the characteristics of the functions $k_1(.)$ and $k_2(.)$ determine whether the utility function displays aversion, loving or neutrality to state-uncertainty. In particular, it is possible to use a definition of certainty equivalent on states to establish a notion of constant aversion to state uncertainty. In particular, for given prize c, and 'local' distribution π with $E^{\pi}(\theta) = \widetilde{\theta}$, let θ^* be the certain state that is utility-equivalent. That is, such that $U(c, \pi) = u(c|\theta^*)$. An agent displays *constant aversion to state uncertainty* if $\kappa = \widetilde{\theta} - \theta^*$ is independent of c and of $\widetilde{\theta}$.

Proposition 2: An agent with constant relative risk aversion preferences (6) with $\gamma > 1$ displays constant aversion to state uncertainty if and only if $k_1(\theta)$ and $k_2(\theta)$ are given by

$$k_i(\theta) = \alpha_i e^{-\rho\theta}$$

with $\alpha_i, \rho > 0$.

Proof: See the Appendix. ■

Aversion to state uncertainty manifests itself also on the marginal utility of consuming c, which is the key quantity when one studies asset pricing implications. Since the marginal utility of an agent with constant relative risk aversion is given by

$$U_c(c, \pi) = E_t[k_2(\theta)]\, c^{-\gamma} \tag{7}$$

Proposition 2 then implies that for $\gamma > 1$ an increase in uncertainty yields an increase in the marginal utility of c. In a dynamic economy, this effect will determine the characteristics of the equilibrium stochastic discount factor and thus the properties of asset prices.

3 A Pure Exchange Economy

I now apply the belief-dependent utility functions to a Lucas (1978) pure exchange economy and derive asset pricing implications under a general set-up for the processes for dividends and the "state." To fix ideas, one can think of the latter as an external habit as in examples (*ii*) above (see e.g. Abel (1990), Campbell and Cochrane (1999)), or as the state of medical technology, which in turn affects the population health level and hence their utility, as in example (*i*). The goal of the next two sections is to obtain general results about the link between belief dependent utilities and stock returns. The empirical section will focus on one particular application – the case of habit formation – where the magnitudes of these effects can be investigated.

Let $\mathbf{W}_t = (W_{1,t}, W_{2,t})$ be a 2-dimensional Wiener process defined on a complete probability space $(\quad , \mathcal{P}^0, \mathcal{F}^0)$. The usual regularity conditions are assumed throughout (see e.g. Duffie (1996), Karatzas and Shreve (1998)). I make the following assumptions about the economy:

Assumption 1: *Real* log-dividends $\delta_t = \log(D_t)$ evolve according to the stochastic differential equation

$$d\delta_t = g_t dt + \sigma dW_{1,t} \tag{8}$$

where g_t is unobservable and its dynamics is described below.

Assumption 2: As in section 2, the representative investor's utility function is belief-dependent and it has the following form

$$U(c_t, p_{\theta,t}, t) = \int_{\mathcal{R}} p_\theta(\theta|\mathcal{F}_t)\, u(c_t, t|\theta)\, d\theta \tag{9}$$

where θ is an unobservable state whose dynamics is described below, $p_{\theta,t} = p_\theta(\theta|\mathcal{F}_t)$ is the posterior marginal density on θ conditional on investors' information at time t, $\mathcal{F}_t$, and $u(c_t, t|\theta)$ is an instantaneous state dependent utility such that the marginal utility of consumption can be written as

$$u_c(c_t, t|\theta) = e^{-\phi t} k(\theta)\, c_t^{-\gamma} \tag{10}$$

where ϕ is the subjective discount rate.

Assumption 3: $\boldsymbol{\nu}_t = (g_t, \theta_t)$ follows any continuous time, stationary Markov process, independent of $\mathbf{W}_t$. Let $p(\boldsymbol{\nu}, t; \boldsymbol{\nu}', t')$ denote the transition probability density that characterizes its law of motion, with $t' > t$.

Assumption 4: Agents observe a noisy signal on θ_t, given by

$$ds_t = \theta_t dt + \sigma_s dW_{2,t}. \tag{11}$$

Finally, consistently with the representation in Theorem 1 (b), the beliefs $p_{\theta,t}$ must be updated using Bayes rule:

Assumption 5: Given a prior belief at time $t = 0$, $p(\boldsymbol{\nu}|\mathcal{F}_0) = p_0(\boldsymbol{\nu})$, investors rationally update their posterior distribution on $\boldsymbol{\nu}_t$, $p(\boldsymbol{\nu}|\mathcal{F}_t)$, by using Bayes law.

Remarks: (i) Standard state-dependent utility functions can be obtained in the limiting case as $\sigma_s \to 0$, so that θ_t is essentially revealed through the signal (11). (ii) Assumption 3 on the stochastic behavior of the state variables $\boldsymbol{\nu}_t = (g_t, \theta_t)$ is very general. In particular, I do not require them to follow continuous path processes. (iii) Assumption 5 is needed to ensure that the results in this paper do not arise because of an increase in degrees of freedom: Assuming that beliefs must be rational and, in the empirical section, derived from the observation of past data (dividends) makes sure that these are taken exogenously and are not just chosen to match the moments of asset returns.

3.1 Equilibrium Pricing

Although the economy is characterized by many sources of risk, namely $\mathbf{W}_t$ and the shocks characterizing $\boldsymbol{\nu}_t$, when we *condition* on the information set $\mathcal{F}_t = \{\delta_\tau, s_\tau : 0 \leq \tau \leq t\}$ we find that all "shocks" to the economy are captured by the innovation process

$$d\widetilde{\mathbf{W}}_t = \mathbf{\Sigma}^{-1} \left[\begin{pmatrix} d\delta_t \\ ds_t \end{pmatrix} - E \begin{pmatrix} d\delta_t \\ ds_t \end{pmatrix} |\mathcal{F}_t \right) dt \right] \qquad (12)$$

where $\mathbf{\Sigma} = diag(\sigma,\sigma_s)$. It turns out that $\widetilde{\mathbf{W}}_t$ is a standard Wiener process with respect to the filtration $\{\mathcal{F}_t\}$ generated by $\{\delta_\tau, s_\tau : 0 \leq \tau \leq t\}$ (see Lemma 1 below). We shall denote by $(\ , \mathcal{P}, \mathcal{F}, \{\mathcal{F}_t\})$ the filtered probability space induced by the $d\widetilde{\mathbf{W}}_t$ on the original probability space. Taking the latter as the primitive probability space and redefining the processes for fundamentals from (12) as

$$\begin{aligned} d\delta_t &= E_t [g|\mathcal{F}_t] \, dt + \sigma d\widetilde{W}_{1,t} \\ ds_t &= E_t [\theta|\mathcal{F}_t] \, dt + \sigma_s d\widetilde{W}_{2,t} \end{aligned}$$

we obtain a standard pure-exchange economy set-up (see Duffie (1996), Karatzas and Shreve (2000)). In this case, it is known that we can always assume the existence of a sufficient number of assets (at least two) in zero net supply to make the markets complete. It follows that any asset paying the stochastic stream $\{q_\tau\}$ of consumption good has a price given by

$$P_t = E_t \left[\int_t^\infty \frac{U_c(c_\tau, p_{\theta,\tau}, \tau)}{U_c(c_t, p_{\theta,t}, t)} q_\tau d\tau \right] = E_t \left[\int_t^\infty e^{-\phi(\tau - t)} \frac{c_\tau^{-\gamma} \int_{\mathcal{R}} p_{\theta,\tau}(\theta) \, k(\theta) \, d\theta}{c_t^{-\gamma} \int_{\mathcal{R}} p_{\theta,t}(\theta) \, k(\theta) \, d\theta} q_\tau d\tau \right] \qquad (13)$$

Unfortunately, solving for this expectation is rather a daunting exercise, even for simple processes for θ_t and g_t and belief-*in*dependent utility functions, and numerical techniques must be employed (see e.g. Brennan and Xia (2001)). Next section provides a viable analytical alternative to the numerical approach.

3.2 A Discretization Approach

I solve directly for the expectation in (13) by relying on a fine approximation of $\mathcal{R}^2$. Besides obtaining an analytical expression for prices, this approach also yields an analytical solution for returns, which can then be fully characterized theoretically.

Consider a discretization of the state space $\mathcal{V} = \left\{\mathbf{v}^1, ..., \mathbf{v}^n\right\}$ where n is "large" (but finite) and each $\mathbf{v}^i = \left(g^i, \theta^i\right)$ is a point on $\mathcal{R}^2$. Hence, by definition, we may have $\mathbf{v}^i$ and $\mathbf{v}^j$ such that either $g^i = g^j$ or $\theta^i = \theta^j$. Since I assumed $\boldsymbol{\nu}_t$ to be stationary, we can always choose the set $\mathcal{V}$ such that $\Pr(\boldsymbol{\nu}_t \notin \mathcal{V})$ is negligible. $\mathcal{V}$ will be the state space on which I define a continuous time Markov chain process $\{\mathbf{v}_t\}$ which *approximates* the original $\{\boldsymbol{\nu}_t\}$. The process $\{\mathbf{v}_t\}$ can be fully described by its infinitesimal generator $\boldsymbol{\Lambda}$, where $[\boldsymbol{\Lambda}]_{ij} = \lambda_{ij} \geq 0$ for $i \neq j$ and $[\boldsymbol{\Lambda}]_{ii} = \lambda_{ii} = -\sum_{j \neq i} \lambda_{ij}$ (see Karlin and Taylor (1975)). Standard results then show that the transition probability to move from $\mathbf{v}^i$ to $\mathbf{v}^j$ during the time interval τ is given by

$$\Pr\left(\mathbf{v}_{t+\tau} = \mathbf{v}^j | \mathbf{v}_t = \mathbf{v}^i\right) = \mathbf{e}_i' \cdot \mathbf{U} \cdot \exp\left(\mathbf{A} \times \tau\right) \cdot \mathbf{U}^{-1} \cdot \mathbf{e}_j \tag{14}$$

where $\mathbf{A}$ is the diagonal matrix with the eigenvalues of $\boldsymbol{\Lambda}$ on its principal diagonal, $\mathbf{U}$ is the matrix of the associated eigenvectors (see Lemma 2 below), $\mathbf{e}_i$ is the i-th column of the identity matrix and "$\exp(\mathbf{A} \times \tau)$" denotes the diagonal matrix with ii-th element given by $\exp(\mathbf{A}_{ii} \times \tau)$. Equation (14) prompts the following definition:

Definition: Let $\mathcal{V} = \left\{\mathbf{v}^1, .., \mathbf{v}^n\right\}$ be given with interval size equal to h and let Δ be a (small) time interval. The continuous time, discrete-state process $\{\mathbf{v}_t\}$ with infinitesimal generator $\boldsymbol{\Lambda}$ *approximates* the continuous time, continuous state Markov process $\{\boldsymbol{\nu}_t\}$ defined by its transition density $p\left(\boldsymbol{\nu}_t, t; \boldsymbol{\nu}_{t'}, t'\right)$ if

$$\Phi_p\left(\mathbf{v}^i, t; \mathbf{E}\left(\mathbf{v}^j\right), t + \Delta\right) \approx \mathbf{e}_i' \cdot \mathbf{U} \cdot \exp\left(\mathbf{A} \times \Delta\right) \cdot \mathbf{U}^{-1} \cdot \mathbf{e}_j \tag{15}$$

where $\Phi_p\left(\mathbf{v}^i, t; ., t + \Delta\right)$ is the probability measure induced on $\mathcal{R}^2$ from the density $p\left(\mathbf{v}^i, t; ., t + \Delta\right)$

and $\mathbf{E}\left(\mathbf{v}^j\right) \subset \mathbf{R}^2$ is a rectangular interval of size h^2 centered in $\mathbf{v}^j$.

In other words, on the discretized grid $\mathcal{V}\times[\Delta, 2\Delta, ...]$ the transition probabilities induced by the original processes and the one approximated through the infinitesimal generator $\mathbf{\Lambda}$ are "close." By choosing n sufficiently large and Δ sufficiently small, for all practical purposes we have that the two processes coincide.[2]

Given this approximation, we can now make use of the following lemma by Liptser and Shyriayev (1977). Let us denote investors' subjective probability that the state is $\mathbf{v}^i$ at time t given their information $\mathcal{F}_t$ by

$$\pi_t^i = \Pr\left(\mathbf{v}_t = \mathbf{v}^i | \mathcal{F}_t\right)$$

This is the discretized version of the density $p(\boldsymbol{\nu}|\mathcal{F}_t)$ introduced earlier. Differently from the latter, the process for the vector $\boldsymbol{\pi}_t = \left(\pi_t^1, ..., \pi_t^n\right)$ can be easily described as a vector diffusion.

Lemma 1: Let the prior distribution $\boldsymbol{\pi}_0 = \widehat{\boldsymbol{\pi}}$, with $\sum_{i=1}^n \widehat{\pi}^i = 1$, be given. Then the vector $\boldsymbol{\pi}_t = \left(\pi_t^1, ..., \pi_t^n\right)'$ evolves according to the system of n stochastic differential equations

$$d\boldsymbol{\pi}_t = \mathbf{\Lambda}' \cdot \boldsymbol{\pi}_t \; dt + \boldsymbol{\pi}_t \odot \boldsymbol{\sigma}\left(\boldsymbol{\pi}_t\right) \; d\widetilde{\mathbf{W}}_t \tag{16}$$

where $\odot$ indicates the element-by-element multiplication, $\boldsymbol{\sigma}\left(\boldsymbol{\pi}_t\right)$ is a $n\times 2$ vector whose $i-th$ component is

$$\boldsymbol{\sigma}_i\left(\boldsymbol{\pi}_t\right) = \left(\mathbf{v}^i - \overline{\mathbf{v}}_t\right)' \; \left(\mathbf{\Sigma}'\right)^{-1}$$

[2] I refer the reader to Kushner and Dupuis (2001) for a detailed treatment on the approximation of general continuous time, continuous state processes by continuous time, discrete state processes (see Ch. 4.3, 5.6). Their definition of approximation relies on the notion of "*local consistency*," that is, over the time inverval Δ, the approximating process should have the same local properties as the original process (see pages 71 and 129). The final conclusion, however, is that on the grid $\mathcal{V}\times[\Delta, 2\Delta, ...]$ the transition probabilites coincide. I take the latter directly as a condition of "approximation," since their construction is rather lengthy and not particular useful for the conclusions of this paper.

One can also consult Israel et al. (2001) on issues related to existence, approximations, computations and references. In particular, it turns out that often an *exact* infinitesimal generator for a continuous time process that replicates a discrete-time transition matrix does not exist (see Theorem 3.1 in Israel et al. (2001)). Nonetheless, an *approximate* generator can generally be computed.

and where $\overline{\mathbf{v}}_t = \sum_{j=1}^{n} \pi_t^j \mathbf{v}^j$ and $\mathbf{\Sigma} = diag\,(\sigma, \sigma_s)$. In addition, $\widetilde{\mathbf{W}}_t$ defined in (12) is a Wiener process defined on the filtered probability space $(\ \ , \mathcal{P}, \mathcal{F}, \{\mathcal{F}_t\})$ where $\{\mathcal{F}_t\}$ is the filtration generated by (δ_τ, s_τ).

Proof: See Liptser and Shyriayev (1977). ■

Remark. Equation (16) underscores the gain from the discretization approach: Although the original process for the original hidden state $\boldsymbol{\nu}_t = (g_t, \theta_t)$ could be very complex and highly non-linear, the process for the discretized posterior probability distribution $\boldsymbol{\pi}_t$ is simple, being it a vector linear process with stochastic volatility. As next Lemma shows, this property yields analytical formulas for conditional expectations of variables that are relevant for pricing:

Lemma 2: Let β be a constant and define the matrix $\overline{\mathbf{\Lambda}}_\beta = \mathbf{\Lambda} + \beta \times diag\left(g^1, ..., g^n\right) + \frac{1}{2}\beta^2\sigma^2\mathbf{I}$. If $\overline{\mathbf{\Lambda}}_\beta$ admits n distinct real eigenvalues, then for any $\tau > 0$ and for any $i = 1, ..., n$ we have

$$E\left[c_{t+\tau}^{\beta}\pi_{t+\tau}^{i}|\mathcal{F}_t\right] = c_t^{\beta}\boldsymbol{\pi}_t' \cdot \mathbf{G}(\beta,\tau) \cdot \mathbf{e}_i \tag{17}$$

where $\mathbf{G}(\beta, \tau)$ is the $n-$dimensional matrix $\mathbf{G}(\beta,\tau) = \mathbf{U}_\beta \cdot e^{\mathbf{A}_\beta\,\tau} \cdot \mathbf{U}_\beta^{-1}$, where $\mathbf{A}_\beta$ is the diagonal matrix with the eigenvalues of $\overline{\mathbf{\Lambda}}_\beta$ on its principal diagonal and $\mathbf{U}_\beta$ is the matrix of the associated eigenvectors, and $\mathbf{e}_i$ is the $i-$th column of the identity matrix.

Proof: See Appendix.■

Remarks: (i) If $\overline{\mathbf{\Lambda}}_\beta$ has either complex or multiple eigenvalues, a solution to the expectation (17) can still be found and it is still linear in the current $\boldsymbol{\pi}_t$. The main difference is that the coefficients of the $\boldsymbol{\pi}_t$ will be dependent on the horizon and possibly oscillatory. (ii) The case $\beta = 0$ implies $\overline{\mathbf{\Lambda}}_\beta = \mathbf{\Lambda}$ and the result in (14) is recovered as a special case.

Given that under the discretization approach, the probability density $p(\boldsymbol{\nu}|\mathcal{F}_t)$ is approximated by the vector of probabilities $\boldsymbol{\pi}_t$, I will denote the utility function as dependent on $\boldsymbol{\pi}_t$ rather than p_t. For notational convenience, let $k_i = k\left(\mathbf{v}^i\right) = k\left(\theta^i\right)$ and $\mathbf{k} = (k_1, .., k_n)$. The marginal utility of consumption implied by the belief-dependent utility is now given by $U_c\left(c_t, \boldsymbol{\pi}_t, t\right) = e^{-\phi t}c_t^{-\gamma}\boldsymbol{\pi}_t' \cdot \mathbf{k}$.

4 Asset Prices

I now apply Lemmas 1 and 2 to obtain closed form solutions for bonds and stocks. To fully gauge the usefulness of Lemma 2, it is instructive to go through the steps to obtain the price of a zero-coupon bond paying one unit of consumption good at time $t+\tau$. From the pricing formula (13) we see that

$$\begin{aligned} Q_t(\tau) &\equiv Q(\tau,\boldsymbol{\pi}_t) = E_t\left[\frac{U_c(c_{t+\tau},\boldsymbol{\pi}_{t+\tau},t+\tau)}{U_c(c_t,\boldsymbol{\pi}_t,t)}\right] \\ &= \frac{1}{e^{-\phi t}c_t^{-\gamma}\cdot\boldsymbol{\pi}_t'\cdot\mathbf{k}}E_t\left[e^{-\phi(t+\tau)}\sum_{i=1}^n k_i\pi_{t+\tau}^i c_{t+\tau}^{-\gamma}\right] \qquad (18) \\ &= \frac{e^{-\phi\tau}}{c_t^{-\gamma}\boldsymbol{\pi}_t'\cdot\mathbf{k}}\sum_{i=1}^n k_i E_t\left[\pi_{t+\tau}^i c_{t+\tau}^{-\gamma}\right] \qquad (19) \\ &= \frac{e^{-\phi\tau}\boldsymbol{\pi}_t'\cdot\mathbf{G}(-\gamma,\tau)\cdot\mathbf{k}}{\boldsymbol{\pi}_t'\cdot\mathbf{k}} \qquad (20) \end{aligned}$$

where $\mathbf{G}(-\gamma,\tau)$ is the $n\times n$ matrix defined in Lemma 2. Note that by evaluating $Q(\tau,\boldsymbol{\pi}_t)$ at $\boldsymbol{\pi}_t=\mathbf{e}_i$, the i-th column of the identity matrix, we obtain the price of the bond *conditional* on $\mathbf{v}^i$ being the true state. This yields the more interpretable formula

$$Q_t(\tau) = \overline{\boldsymbol{\pi}}_t'\cdot\mathbf{Q}(\tau) \qquad (21)$$

where $\mathbf{Q}(\tau) = (Q(\tau,\mathbf{e}_1),...,Q(\tau,\mathbf{e}_n))$, and

$$\overline{\pi}^i = \frac{\pi_t^i k^i}{\boldsymbol{\pi}_t'\cdot\mathbf{k}} \qquad (22)$$

is a probability distribution on $\mathcal{V}$ that is weighted by the marginal utility across the states $\mathbf{v}^i$'s. Formula (21) generalizes earlier results by Yared (1999) and Veronesi and Yared (2000). Hence, with belief-dependent utility functions the value of one unit of consumption in the future is a weighted average of its discounted value conditional on each of the $\mathbf{v}^i$, where the weights are probabilities $\overline{\pi}^i$ that are *adjusted* to reflect the different levels of marginal utility across the unobservable states $\mathbf{v}^i$.[3]

Using a similar approach for stocks, the appendix shows:

[3] Clearly, the distribution in formula (22) can be interpreted as a state-price density, where the states are the $\mathbf{v}^i$'s in $\mathcal{V}$.

Proposition 3: Define the $n-$dimensional vector

$$\mathbf{B} = \mathbf{k}^{-1} \odot \left(\phi \mathbf{I} - \overline{\mathbf{\Lambda}}_{1-\gamma}\right)^{-1} \cdot \mathbf{k} \tag{23}$$

where $\odot$ denotes "element-by-element" product, $\mathbf{k}^{-1} = \left(k_1^{-1}, ..., k_n^{-1}\right)$ and $\overline{\mathbf{\Lambda}}_{1-\gamma}$ is defined in Lemma 2. Then:

(a) The price of the risky asset is

$$P_t = D_t \, \overline{\boldsymbol{\pi}}_t' \cdot \mathbf{B} \tag{24}$$

(b) The real rate of interest is

$$r_t = \phi + \gamma \overline{\boldsymbol{\pi}}' \cdot \mathbf{g} - \frac{1}{2}\gamma^2\sigma^2 - \overline{\boldsymbol{\pi}}' \cdot \mathbf{C}^* \tag{25}$$

where $\mathbf{C}^* = (C_1, ..., C_n)$ with $C_j^* = \sum_{i=1}^n \lambda_{ji} \frac{k_i}{k_j}$.

Proof: See Appendix ■

These asset pricing formulas have a number of properties that I discuss in the next few pages. In the following, I will refer to the vector $\mathbf{B}$ in (23) as *conditional price-dividend ratios*, because each element of it is the price-dividend ratio that would occur if there was perfect certainty on the underlying state. Indeed, from (22) we see immediately that if $\pi_t^\ell = 1$ for some ℓ, then $\overline{\pi}^\ell = 1$ and hence $P_t/D_t = B_\ell$. To better understand the effect of belief dependent utilities, and aversion to state uncertainty, it is useful to obtain the prices and interest rates under the benchmark case of belief independent utility function. This is a special case $k(\theta) = k = 1$.

Corollary 1: In the case of state-independent utilities, let $\overline{\mathbf{B}} = \left(\phi \mathbf{I} - \overline{\mathbf{\Lambda}}_{1-\gamma}\right)^{-1} \cdot \mathbf{1}_n$ and $\mathbf{Q}^B(\tau) = e^{-\phi\tau}\mathbf{G}(-\gamma, \tau) \cdot \mathbf{1}_n$. Then:

$$P_t^B = D_t \, \boldsymbol{\pi}_t' \cdot \overline{\mathbf{B}} \tag{26}$$

$$Q_t^B(\tau) = \boldsymbol{\pi}_t' \cdot \mathbf{Q}^B(\tau) \tag{27}$$

$$r_t^B = \phi + \gamma \, \boldsymbol{\pi}_t' \cdot \mathbf{g} - \frac{1}{2}\gamma^2\sigma^2 \tag{28}$$

Comparing the pricing formulas obtained in Proposition 3 with the corresponding ones in Corollary 1 for the benchmark case, we see that the effect of a belief-dependent utility function shows itself in two terms in the pricing function (24) compared to the benchmark case (26):

First, belief-dependent utilities affects the "conditional price-dividend ratios" B_i as it can be seen by comparing (23) with the same expression in Corollary 1. Intuitively, the marginal utility of consumption is now belief-dependent. Hence, even when we condition on a particular state (i.e. we set $\pi_t^i = 1$ for some i), the state-dependent marginal utility affects the comparison between current and future marginal utilities, thereby affecting the conditional price-dividend ratio B_i. The impact is large: Figure 1 plots the conditional price-dividend ratios B_i under the special case discussed in Section 5 for calibrated parameter values. The conditional price dividend ratios are plotted for three values of a parameter ρ that, as in proposition 2, regulates the "aversion" to state uncertainty, $\rho = 0, 40, 80$. The case $\rho = 0$ corresponds to the benchmark case in Corollary 1.

A few remarks are in order here: First, in the Benchmark case $\rho = 0$ (and $\gamma > 1$) Figure 1 shows that conditional price-dividend ratios $\overline{B}_i$'s are typically *decreasing* with g^i, i.e. a *higher* growth rate of the economy is associated with a *lower* price-dividend ratio. This is due to the low elasticity of intertemporal substitution that occurs if $\gamma > 1$ (see e.g. Campbell (1999), Veronesi (2000) for a discussion).[4] I point out that since this problem is due to the low elasticity of intertemporal substitution (EIS), Epstein-Zin-Weil preferences that disentangle risk aversion from EIS do not provide help unless one is ready to assume that EIS>1. Macroeconomic studies seem to agree that EIS<1 (see e.g. Campbell (1999) for a discussion). Second, as we increase the parameter regulating aversion to state uncertainty ρ, the conditional price-dividend ratio moves from being negatively sloped with respect to g^i to being positively sloped. This has the important implication of generating a positive covariance between expected consumption growth and returns, thereby increasing the equity premium.

A second reason why the price dividend ratio under belief dependent preferences differs from the Benchmark case is that the conditional price-dividend ratios B_i are weighted by the probabilities $\overline{\pi}_t^i$ rather than the original π_t^i. Intuitively, the probabilities $\overline{\pi}_t^i$ are now adjusted for the impact that each state θ^i has on the marginal utility of consumption, namely for k_i. In other words, if $k_i > k_j$, then $\overline{\pi}_t^i$ becomes relatively bigger than $\overline{\pi}_t^j$: Stock prices now reflect more those states characterized by higher marginal utility of consumption.

[4] This is not absolutely true and depends on the transition probabilities λ_{ij}. However, this holds under the assumptions made in section 6, where the model is taken to the data. See also Figure 1.

Turning to the risk free rate, a similar comment applies: in the benchmark case, a high average interest rate realizes when agents have positive expectation of consumption growth $E_t[g] = \boldsymbol{\pi}_t' \cdot \mathbf{g}$, and a low intertemporal elasticity of substitution EIS, which translates into a high $\gamma = 1/\text{EIS}$. Proposition 3 (b) shows that although a low EIS would also generate a high risk free rate with belief dependent utility, now the expectation of future consumption growth is taken *under the distorted beliefs* $\overline{\boldsymbol{\pi}}_t$, $\overline{E}_t[g] = \overline{\boldsymbol{\pi}}_t' \cdot \mathbf{g}$. This quantity will turn out to be lower than the its counterpart under the true beliefs. In addition, the risk free rate is also affected by the term $\overline{\boldsymbol{\pi}}_t' \cdot \mathbf{C}^*$ which is generally different from zero. This measures the predictable changes in state (through the transition matrix $\boldsymbol{\Lambda}$), which yield a predictable variation in future marginal utilities (through the $k(\theta)$ function) which again will generally call for an increase in the demand for bonds.

4.1 Stock Returns

Let us denote the *excess* return as

$$dR_t = \frac{dP_t + D_t dt}{P_t} - r_t dt$$

Proposition 4: The excess return dR_t follows the process $dR_t = \mu_{R,t}dt + \boldsymbol{\sigma}_{R,t}d\widetilde{\mathbf{W}}_t$ with

$$\mu_{R,t} = \gamma\left(\sigma^2 + \Delta V_t^g\right) - V_t^{k,g}\left(1 + \sigma^{-2}\Delta V_t^g\right) - \sigma_s^{-2} V_t^{k.\theta} \Delta V_t^\theta \tag{29}$$

$$\boldsymbol{\sigma}_{R,t} = \left(\sigma + \Delta V_t^g \sigma^{-1}, \Delta V_t^\theta \sigma_s^{-1}\right) \tag{30}$$

where for $y = k, kB$ and $x = g, \theta$, we have

$$V_t^{y,x} = \frac{\sum_i \pi_t^i y_i \left(x_i - \overline{x}_t\right)}{\sum_i \pi_t^i y^i} = \frac{Cov_t(y,x)}{E_t[y]} \tag{31}$$

$$\Delta V_t^x = V_t^{kB,x} - V_t^{k,x} \tag{32}$$

Proof: See Appendix. ■

The properties of stock returns are determined by the quantities $V_t^{k,x}$ and ΔV_t^x, for $x = g, \theta$, which I characterize in the next proposition. To state it, it is convenient to return to the notation introduced in Section 3. That is, let $p(\theta, g|\mathcal{F}_t)$ denote the posterior density on (θ, g), and $B(\theta, g)$ denote the conditional price dividend ratio at the value (θ, g).

Proposition 5: From Proposition 1, recall $k(\theta) > 0$ for all θ. Let $p(\theta, g|\mathcal{F}_t)$ be non-degenerate. Then:

(a) $V_t^{k,\theta} > (<) 0$ if $k(\theta)$ is monotonically increasing (decreasing) in θ. $V_t^{k,\theta} = 0$ if $k(\theta) = k$, a constant;

(b) $V_t^{k,g} > (<) 0$ if $k(\theta)$ is monotonically increasing (decreasing) in θ, and (θ, g) have positive correlation. Viceversa if (θ, g) have negative correlation. $V_t^{k,g} = 0$ if $k(\theta) = k$, a constant, or (θ, g) are uncorrelated.

(c) Both ΔV_t^g and ΔV_t^θ are non negative (non positive) if $B(\theta, g)$ is monotonically non decreasing (non increasing) in (θ, g).

Proof: See Appendix. ■

This proposition yields a simple intuition for $V_t^{k,x}$ and $V_t^{kB,x}$. Consider the case where $x = \theta$, for example. Then, $V_t^{k,\theta}$ can be interpreted to measure two important components of the belief-dependent utility function. First, the degree of "state-dependency" as measured by the behavior of the function $k(\theta)$: If $k(\theta) = k = \text{constant}$, then $V_t^{k,\theta} = 0$. Second, the amount of uncertainty on θ_t: If the distribution $\boldsymbol{\pi}_t$ gives probability one to one particular state θ^ℓ, then again $V_t^{k,\theta} = 0$. Similarly, for $x = g$, we have that $V_t^{k,g} = 0$ if either there exists ℓ such that $\pi_t^\ell = 1$, or $k(\theta) = k = \text{constant}$, or θ_t and g_t are uncorrelated. A similar intuitive argument holds also for $V_t^{kB,x}$, with the additional caveat that it will depend also on the behavior of the *equilibrium* conditional price-dividend ratio.

Proposition 5 has a few immediate implications which I collect in the following Corollary:

Corollary 2: (i) If $\boldsymbol{\pi}_t$ is degenerate (perfect certainty), then $\mu_{R,t} = \gamma\sigma^2$, and $\boldsymbol{\sigma}_R \boldsymbol{\sigma}'_R = \sigma^2$. (ii) In the benchmark case (corollary 1), if $B(g)$ is decreasing in g then $\mu_{R,t} < \gamma\sigma^2$. (iii) If $k(\theta)$ is decreasing, θ is procyclical ($cov_t(\theta, g) > 0$), and $B(g, \theta)$ is non decreasing in both arguments, then $\mu_{R,t} > \gamma\sigma^2$, and $\boldsymbol{\sigma}_{R,t} \boldsymbol{\sigma}'_{R,t} > \sigma^2$.

Proof: Immediate from Propositions 4 and 5. ■

Case (i) is the classic case where there is no uncertainty, while case (ii) shows that uncertainty alone cannot yield a high equity premium (see Veronesi (2000)). Recall from the earlier discussion (see Figure 1) that $\gamma > 1$ typically implies a decreasing conditional price dividend ratio $B(g)$, which makes it impossible to match the equity premium, even for high values of γ. Case (iii) shows instead that belief dependent utilities go in the right direction, under some conditions about the nature of the "state." Next section shows that these conditions are met by habit persistence preferences, for instance.

Finally, notice that even if the unobservable state θ is unrelated to the drift rate g of the economy, equilibrium stock returns would still be affected by the beliefs on θ, as in this case we simply have $V_t^{k,g} = 0$ but all the other terms in (29) remain. For instance, in a model with a stochastic subsistence level of consumption (see e.g. Campbell and Viceira (2002, Ch. 6)) correlated with an unobservable health status (see example (i) in Section 2), we would still obtain an effect on the equity premium, although aggregate health and aggregate economy are uncorrelated (see Cutler and Richardson (1997)).

4.2 Habit Formation

It is useful to interpret the results in the previous sections in light of a popular external habit formation model. Consider the utility index

$$u\left(c_t^i, t|X_t\right) = e^{-\phi t}\frac{\left(c_t^i - X_t\right)^{1-\gamma}}{1-\gamma} \tag{33}$$

where X_t is a slow-moving external habit. As discussed in Section 2, since X_t is an *external* index of the "Joneses" standard of living (see e.g. Abel (1990), Detemple and Zapatero (1991), Campbell and Cochrane (1999)) it is likely that each investor does not actually observe X_t, but he/she only possess a probability distribution $p_t(X_t)$ on its level, which is obtained from various signals, possibly including the actual behavior of the "Joneses" themselves.[5] From Section 2, preferences must then be described by $E_p\left[u\left(c_t^i, t|X_t\right)\right]$ yielding "belief-dependency".

[5] While doing my PhD in Cambridge, one of my acquantancies had a shining BMW (while I was driving a very old Volvo). I later found out that the BMW was pretty much all this person in fact possessed, and it was certainly not reflecting the overall state of his aggregate consumption. Observable consumption is just a noisy signal about the standard of living.

In addition, the curvature of the utility function (33), implies that a mean preserving spread $\widetilde{p}(X)$ on the density $p(X)$ decreases the agent's utility. That is, for given consumption level c_t^i at time t, $E_{\widetilde{p}}\left[u\left(c_t^i, t|X_t\right)\right] < E_p\left[u\left(c_t^i, t|X_t\right)\right]$. This implies that preferences with habit formation naturally generate an *aversion to state-uncertainty*. Intuitively, this is because with a higher dispersion of beliefs, there is marginally more probability to be close to X_t, which is the point of maximum disutility.

What are the implications for equilibrium returns? Propositions 4 and 5 provide an immediate answer. In fact, borrowing the methodology of Campbell and Cochrane (1999), denote by $S_t = (c_t - X_t)/c_t$ the surplus consumption ratio. Note that uncertainty on X_t implies uncertainty on S_t, and that S_t is a procyclical variable, that is, it increases when g_t increases, if the latter is time varying. In addition, we can rewrite the marginal utility implied by (33) as $u_c(c_t, t|X_t) = e^{-\phi t} S_t^{-\gamma} c_t^{-\gamma}$, which has the same form as in Assumption 2 in Section 3. Thus, identifying S_t with θ_t in the previous notation, we obtain that S_t and g_t are positively related, while $k(S_t) = S_t^{-\gamma}$ is decreasing in S_t. From Proposition 5 (a) and (b), then, $V_t^{k,\theta} < 0$ and $V_t^{k,g} < 0$. For the case where g is (known) constant, Campbell and Cochrane (1999) show that the conditional price dividend ratio $B(S_t)$ is increasing in S_t, which from Proposition 5 (c) again implies $\Delta V_t^{\theta} > 0$. Thus, from (29) and (30) this behavior of S_t, $k(S_t)$ and $B(S_t)$ yields immediately an expected return and a volatility that are higher than the ones one would obtain without belief dependency. In addition to this effect, belief dependency would generate time varying conditional moments as the beliefs on S_t move over time. As I show in the next section for a model with also time varying g_t, I do not need to assume time varying volatility of S_t – as Campbell and Cochrane (1999) do – to generate time varying volatility and a low volatile risk free rate, as these effects are obtained through the time variation in the dispersion of beliefs on S_t.

5 A Calibration

Propositions 4 and 5 fully characterize the impact of belief dependent utility functions on expected returns and volatility. In this section, I calibrate the economy to obtain some quantitative predictions, especially on the time variation of conditional moments, such as volatility.

In order to carry out the exercise without adding too much computational complexity, I focus on a very special case of the more general set up developed in the previous sections, namely, the case where the state affecting individual preferences is a monotonic function of the drift rate of dividends. This is consistent with an extreme case of the habit formation model discussed in section 4.2, where $\theta_t = S_t = (c_t - X_t)/c_t$. In fact, S_t is a procyclical variable that is high when g_t is high.[6] Thus, since the purpose here is to obtain some ballpark numbers of the quantitative effect of belief dependent utilities, an approximation of the model in Section 4.2 is to assume that $S_t = G(g_t)$ for some function $G(.)$ with $G'(g) > 0$. In addition, since investors are learning about g over time by observing consumption realizations, this also implies that the expected surplus $E_t[S_t]$ is perfectly correlated with news in consumption growth, as the habit formation specification requires.

5.1 Structural Breaks

I assume that the drift rate g_t evolves according to the following process

$$dg_t = (\mu - \kappa g_t)\, dt + \sigma_g dW_{2,t} + \zeta_t dQ_t^p \tag{34}$$

where $dW_{2,t}$ is a Brownian motion independent of dW_t, dQ_t^p is a Poisson process with intensity p, and ζ_t is a random variable distributed according to some distribution F_ζ such that $E[\zeta_t] = \mu/\kappa$. The jump component captures the intuition discussed in the introduction that occasionally the growth rate of the economy can change abruptly and discretely.[7] Since g_t is unobservable, the effect is to generate a time varying dispersion in the posterior distribution of g.

To describe further the properties of the conditional price-dividend ratios B_i, I need more structure on the function $k(g)$ and the density F_ζ of the jump variable ζ_t. In line with a notion of constant aversion to state uncertainty (see Proposition 2), it is convenient to assume[8]

$$k(g) = e^{-\rho(g-g^n)}$$

[6] For the case where X_t equals a weighted average of past consumption c_s's, simulations show that S_t has a correlation as high as 80% with g_t for reasonable parameters.

[7] See Veronesi (2000) and Timmerman (2001) for similar models applied to stock prices.

[8] Rescaling by the constant $e^{\rho g^n}$ ensures that the restrictions imposed by the habit formation interpretation of the model are met, that is, that the implied surplus $S_t^{-\gamma} = k(g_t) > 1$.

As discussed in Proposition 2, the parameter ρ can be interpreted as a measure of *aversion to state uncertainty.* A high level of ρ implies that a mean preserving spread on the posterior $p(g|\mathcal{F}_t)$ has a high impact on the representative agent's utility and marginal utility of consumption. The intuition within the habit formation interpretation is straightforward: if the standard of living of the "Joneses" are not observable, periods of high uncertainty on the drift rate of the economy must also be periods where agents have a more difficult time to filter the "Joneses" actual standard of living. This implies a higher uncertainty on the habit level X_t. Finally, note that the case $\rho = 0$ corresponds to the state-independent utility function, whose properties under learning are investigated in Veronesi (2000).

Second, I consider two cases for the jump distribution F_ζ: First, the case where F_ζ is a standard symmetric normal distribution, and second, the case where F_ζ is allowed to be asymmetric. The latter case is motivated by the empirical observation that realized real consumption growth rates are negatively skewed. In fact, using data from 1952 - 2001, I find that the skewness of $\Delta \log(c_t)$ is about -0.4304. A symmetric model such as the one where F_ζ is normally distributed, or the case with no jumps, would generate zero skewness. To allow for this possibility, then, I consider a flexible jump distribution, such as the Weibull distribution

$$F_\zeta(x) = I_{\{x \geq \mu_\zeta\}} \eta_\zeta \left(\frac{x-\mu_\zeta}{\sigma_\zeta}\right)^{\eta_\zeta - 1} e^{-\left(\frac{x-\mu_\zeta}{\sigma_\zeta}\right)^{\eta_\zeta}} \text{ for } \eta_\zeta, \sigma_\zeta > 0 \tag{35}$$

The three parameters appearing in equation (35) are a location parameter μ_ζ, a dispersion parameter σ_ζ and a shape parameter η_ζ. In particular, a low value of η_ζ generates a positively skewed distribution, while a high value generates a negatively skewed distribution.

5.2 Fitting Real Consumption Data

In this section I estimate the consumption model in the previous section using quarterly data on real consumption growth from 1952-2001. Consumption data are from the NIPA tables and include only non-durables and services. Nominal per-capita data have been deflated using the CPI index. In addition, for comparison, I also estimate standard moments of stock prices, returns, and riskless rate, using quarterly return data and the riskless rate data from the CRSP dataset of the University of Chicago.

The discretization approach described in section 3.2 to solve for asset prices turns out to be extremely useful also operationally in order to estimate the parameters of the model, a task that is complicated by the highly non-linear dynamics imposed on the hidden state variable g_t. Specifically, I apply the (approximate) Maximum Likelihood methodology put forward by Kitagawa (1987) or, as a generalized case, by Hamilton (1989), to the discrete-time version of the model. In Appendix C, I obtain the discrete-time counterparts of the pricing formulas (21) and (24) for bonds and stocks, which will be used in simulations in the next section.

More specifically, I estimate the discrete-time model

$$\Delta \log (c_{t+1}) = g_t + \sigma \varepsilon_{t+1}$$

where $\varepsilon_{t+1} \sim \mathcal{N}(0,1)$, and g_t follows the discrete-time counterpart of (34), that is

$$g_{t+1} = \begin{cases} \mu + a g_t + \sigma_g \varepsilon_{g,t+1} & \text{with prob.} \quad 1-p \\ \zeta_{t+1} & \text{with prob.} \quad p \end{cases} \tag{36}$$

and $\zeta_{t+1} \sim F_\zeta$, described earlier. Using the same methodology as in Section 3.2, I transform the continuous state process (36) into a $n-$state, Markov chain model with transition probability $\mathbf{\Psi}$, that is, $\psi_{ij} = \Pr\left(g_t = g^j | g_{t-1} = g^i\right)$. The parameters of the model can be estimated using a standard Maximum Likelihood approach.

Table I contains the parameter estimates for the three cases where the jump probability $p = 0$, $p \neq 0$ and jumps are symmetric, and $p \neq 0$ and jumps may have an asymmetric distribution. The parameters are reported in quarterly units. Across the three cases, the estimates imply an annualized average long-term growth rate of real consumption of about 2.04-2.08%, and a volatility of consumption of 1.13-1.16%. These estimates are in line with the ones obtained by others (see e.g. Campbell, Lo and MacKinlay (1997)). Across the models the estimates also reveal a somewhat strong mean reversion of the hidden state g_t, with a mean reversion parameter a around 79%, for the case without jumps, and up to 89% for the case with jumps. In the case of symmetric jumps (Panel B), the point estimate of the jump probability is $p = 7.7\%$ (quarterly), implying a discrete change in g_t once every 3.25 years in average. Thus, discrete changes are not very frequent, but not very infrequent either. I notice, however, that this jump probability has a rather large standard error. Given the significantly high number of parameters, the low frequency of jumps, and the fact that what is subjects to "jumps" is a hidden state, it may not be too surprising to see a large value of the standard error for the probability p. Notice, however, that the size of the distribution of the jump, σ_ζ is indeed statistically different from zero, and fairly large ($\sigma_\zeta = .5\%$ quarterly). Figure 2 plots the shape of the distribution in this case. Notice that even if a discrete break should occur in average every 3.25 years, they are drawn from a distribution with a large mass around the unconditional mean. Thus, the break may not be easily detectable. In other words, large discrete breaks occur with much less frequency, and it is to this one that agents in my model are more averse.

Panel C contains the estimates for the case where the jump distribution is allowed to have a non-symmetric shape. The parameters for the autoregressive part are very similar to the ones obtained for the symmetric jump case in Panel B. The jump probability is now $p = 10.8\%$, implying a discrete change in g_t every 2.3 years in average. Slightly more frequent than before, but the standard errors are large, so the estimates are not significantly different from each other in any case. Most interestingly, however, is that consistently with the evidence that sample growth rates of consumption are negatively skewed, I find a negatively skewed jump distribution, as plotted in Figure 2. As in Panel B, also in this case the parameters describing the shape of the jump distribution are statistically significant.[9]

5.2.1 Time Varying Uncertainty

Figure 3 reports the evolution of the posterior distribution $\boldsymbol{\pi}_t = \left(\pi_t^1, ..., \pi_t^n\right)$ from 1952 to 2001, for the two cases where the drift rate of consumption g_t is not allowed to jump discretely (Panel A), or is allowed to do so with an asymmetric jump distribution (Panel B).[10] Panel A shows that consistently with the Kalman filter, the posterior density obtained in the case where g_t is

[9] The large values for σ_ζ and η_ζ appear unusual, but they are due to the lack of sufficient flexibility in the Weibul distribution to generate a negatively skewed distribution. In fact, data seem to suggest a very strong negatively skewed distribution, that is, a high value for η_ζ. However, a high value of η_ζ also decreases the variance of the distribution. An increase in σ_ζ is then necessary to re-establish a higher value of the variance of the distribution.

[10] The case with symmetric jumps lead to a very similar plot.

not allowed to jump is normally distributed with a constant dispersion. Indeed, as it can be seen it is only the mean of the distribution that changes over time, and not its variance. In contrast, when jumps are allowed (Panel B), the evolution of the posterior distribution shows a good deal of variation in its dispersion around the mean. The intuition is simple: When a jump occurs (or agents believe it occurred) it takes some time to learn the new drift, while the latter is moving back towards its long term mean $\mu/(1-a)$.[11] During this period of time the posterior distribution widens. This effect is even more evident in Figure 4, which plots the time series of mean drift rate $E_t(g) = \sum_{i=1}^{n} \pi_t^i g^i$ (Panel A) and the root mean square error of agents distribution $RMSE_t = \sqrt{E_t(g^2) - E_t(g)^2}$ (Panel B), again for the two cases. Panel A shows that the filtered drift rate of consumption is essentially identical in the two models (with and without jumps), as the solid and the dash-dotted line exactly overlap each other. In contrast, Panel B shows that the two models differ markedly in the dispersion of the posterior belief around the mean. While the model with no jumps implies a constant RMSE, the model where jumps are allowed implies a high variability of state-uncertainty. More specifically, the uncertainty on the drift rate of consumption was high in the 56-58 period, it decreased in the 60s but recovered in the middle of the 70s and especially at the beginning of the 80s. Higher uncertainty was again realized in the 90-95 period, and, interestingly, in the year 2001.

5.3 Implications for Asset Prices

Table II reports simulation results for the most important unconditional moments of returns and interest rates. The simulations are performed as follows: I simulate 4000 quarters of consumption data according to the three models estimated in Table I. Given each sequence of consumption data the posterior probabilities are computed using Bayes law, via the updating

[11] I also estimated a "pure jump" process, where $g_{t+1} = g_t$ with probability $1-p$ and $g_{t+1} = \zeta_{t+1}$ with probability p. The implications for returns and volatility are even stronger than the one we will see below. However, p came out too high, about $p = .2$, implying a change in drift every 1.25 years. Given the frequency of breaks, the estimated model, although more intuitive, failed to conform to the intuition of rare discrete breaks in g_t. Results are available upon request.

formula

$$\pi_i(t+1) = \frac{e^{-\frac{1}{2\sigma^2}\left(\Delta c(t+1)-g^i\right)^2}\left[\boldsymbol{\pi}(t)'\cdot\boldsymbol{\Psi}\right]_i}{\sum_{j=1}^n e^{-\frac{1}{2\sigma^2}\left(\Delta c(t+1)-g^j\right)^2}\left[\boldsymbol{\pi}(t)'\cdot\boldsymbol{\Psi}\right]_j}$$

From the sequence of consumption growth and probabilities $\{\boldsymbol{\pi}(t)\}$, a time-series of prices for the risky stock and bonds are then computed, by using the convenient formulas (46), and (47) in Appendix C. Results are reported for various values of the preference parameters γ, ρ and ϕ. Specifically, I consider the cases where $\rho = 0$, corresponding to the standard power utility, and $\rho = 40$ and 80. The implied P/D ratios for these three values of ρ and for $\gamma = 3$ and $\phi = .01$ are in Figure 1.

Panel A of Table II contains the unconditional moments of asset returns in the data, for the sample 1952 - 2001. These values are those that make up the various puzzles in finance, such as the high equity premium (7.19%), the high volatility of stock returns (16.51%) compared to consumption (1.12%), the low average real rate (1.53%), where expected inflation is computed using a standard AR(1) process for inflation, the low volatility of the risk-free rate (1.05%), the high Sharpe ratio $E[R]/\sigma(R) = 43.54\%$ and the average price-dividend ratio $\overline{P/D} = 32$. This latter value is heavily affected by the last 5 years of the nineties, which saw unprecedented values for the P/D ratio. Using data up to 1995, the corresponding value of the average P/D would be around 27.

Panel B - D contain the results of simulations for the three models described. It is useful to have a first look at the case where $\rho = 0$ and thus agents have a standard power utility.[12] First, the equity premium is small and, in addition, it becomes negative as γ increases. This is due to the a natural hedging effect that is generated by the case $\gamma > 1$. As explained in Section 4.1 after corollary 2, in this case the conditional price dividend ratio $B(g)$ is decreasing in g. This implies that bad news in consumption growth increase the price dividend ratio, thereby generating a smaller – or possibly even negative – covariance between returns and consumption growth. This in turn yields a low or negative expected return on stock (see Veronesi (2000)).

Across the three Panels, as we increase ρ, the coefficient of aversion to state uncertainty, the equity premium and volatility increase, and the interest rate level decreases, although its

[12]This case is very similar to Veronesi (2000), who studies the case of power utility under the assumption that g_t follows a pure jump process.

volatility increases. Thus, the Sharpe ratio increases. The most important effect, however, is the fact it is possible to obtain a high expected return and a high volatility of returns, without increasing the coefficient γ, which regulates the intertemporal elasticity of substitution. Thus, "aversion to state uncertainty" is able to generate a sizable equity premium and volatility, without making the level of interest rates unboundedly high. Although the model finds it hard to match the very high Sharpe ratio (43%) observed in the data, it still produces a value of up to 32% ($\gamma = 3$, $\rho = 80$, $\phi = .01$ in Panel D) with a reasonable real risk-free rate around to 3.2%. The volatility of risk-free rate results instead higher than in the data (4.5% against 1%). However, it is far lower than the typical calibrated value obtained in other models with state-dependent preferences, such as the habit formation model of Abel (1990), who reports a volatility of the risk free rate of about 18%, or, more recently, the habit formation model of Boldrin et al. (2001), who report a level of volatility for the riskless rate of 24%. The exception is the model of Campbell and Cochrane (1999), who managed to obtain a constant risk free rate. However, this is accomplished by choosing the volatility of the surplus consumption ratio ad hoc with the sole intent to achieve this result.

As Panels B - D reveal, the three models yield similar results at the unconditional level, with the asymmetric jump case performing slightly better. The three models mostly differ in their implications about the time-series of volatility, to which I now turn. Since the results are similar across the various parameter levels, I choose the parameters $\gamma = 3$, $\rho = 80$ and $\phi = .01$ for the three models. These are the ones for which unconditional moments better reflect the unconditional moments in the data.

5.3.1 Time-Varying Volatility

The results in Section 4.1 suggest that time varying uncertainty – such as the one in Figures 3 and 4 – should generate time varying volatility as it affects the quantities $V_t^{k,x}$ which enter the volatility formula (30). In addition, those results also show that if g_t has no jumps but follows a simple AR(1) process, we should expect little or no time-varying volatility, as the uncertainty is constant over time (see, again, Figure 3 A and Figure 4). This is exactly what I find in the simulations.

For each simulated series, I fit a GARCH(1,1) and a EGARCH(1,1) model, where returns are conditionally normally distributed $R_{t+1} = \mu + \sigma_t \varepsilon_{t+1}$, $\varepsilon_{t+1} \sim \mathcal{N}(0,1)$, but with volatility given by, respectively:

$$\text{GARCH(1,1)} \quad : \quad \sigma_{t+1}^2 = \omega + \beta \sigma_t^2 + \alpha R_{t+1}^2 \tag{37}$$

$$\text{EGARCH(1,1)} \quad : \quad \log(\sigma_{t+1}) = \omega + \beta \log(\sigma_t) + \alpha \left[|\varepsilon_{t+1}| - c\varepsilon_{t+1} \right] \tag{38}$$

In the EGARCH(1,1) model, the coefficient c regulates the "leverage effect": If $c = 0$, then positive and negative innovations in returns have equal impact on the volatility. A positive $c > 0$ implies that negative innovations increase the volatility by more than positive innovations. A value $c > 1$ implies that while negative innovations increase volatility, positive innovations in fact decrease it.

Table III reports the results of the simulations. For convenience, Panel A reports the results for quarterly data for the sample 1952 - 2001. I notice that for quarterly data it is harder to detect strong persistence in volatility, as it is known that volatility clustering is much stronger at shorter horizons, such as daily or monthly. Yet, the empirical results still reveal some volatility clustering, with an autoregressive parameter β equal to .63 (GARCH) and .91 (EGARCH). In addition, the coefficient c is significantly positive, revealing that a leverage effect is present even at the quarterly frequency, and significantly less than 1, revealing that both bad and good news increase volatility.

Panels B - D report the results for the three models. As discussed earlier and intuitively, the smooth autoregressive model for g_t shows little or no time variation in volatility. The parameter α is essentially zero in both cases, implying that innovations in returns have no effect on the volatility. Recall that the sample in simulation is of 4000 quarters, and thus finding insignificant coefficients in α must reflect that no action stems from this variable. In contrast, the models with discrete jumps in the drift rate g_t demonstrates a good deal of time variation in volatility, and of about the same magnitude as in the data. The main difference between the symmetric and asymmetric model has to do with the EGARCH(1,1) model: The model with symmetric jump implies essentially no asymmetry between good news or bad news as c is small, and not significantly different from zero. In contrast, the model with asymmetric jump generates an asymmetry quite similar to the one in the data. In fact, the negatively skewed jump distribution implies that it is more likely to see an increase in the dispersion of the distribution $\boldsymbol{\pi}_t$ after bad news than after good news. Thus, since bad news imply low returns, we have contemporaneously a negative return with an increase in volatility (see equation (30)), and hence, the leverage effect.

5.3.2 High Volatility in Recessions

Related to the point above, the model with asymmetric jump distribution yields a higher volatility of returns during recessionary phases of the economy. Intuitively, consider the case where times are good and uncertainty is low. A sequence of negative innovations in consumption leads agents to believe that g_t dropped discretely to a lower level. This implies a high dispersion of beliefs initially, as agents try to figure out "how bad" the bad state is. Thus, the sequence of negative innovations in consumption are followed by a high volatility. As agents learn, the mean reversion of g_t leads it to the stationary value, therefore implying also the end of the recession together with lower volatility. A sequence of unusually high innovations in consumption, however, do not lead to an equal increase in volatility, as in this case there is less probability that g_t is extremely high. Thus, the dispersion is bounded above, in a sense. This implies that the volatility is lower during good (better) times, than during bad (worse) times.

The simulation results confirm this. Without reporting the results in a separate table to save space, I find in the simulations that the correlation between the expected growth rate of the economy, $E_t[g_t]$, and the theoretical volatility $\sigma_{R,t}$, the latter computed using equation (30), is 0.2952, -0.0700, and -0.4731, in the three models with no jumps, symmetric jumps and asymmetric jumps, respectively.

6 Conclusions

In this article I re-interpreted standard axioms in choice theory delivering state-dependent utility functions to introduce the notion of a "*belief*-dependent" utility function, that is, a utility function that depends on the subjective beliefs on an underlying partially observable state of nature. I show that this interpretation naturally leads to a notion of "*aversion to state-uncertainty,*" that is, aversion to a wider subjective distribution on the underlying (unobservable) state.

I then apply this type of preferences to a standard pure exchange economy under a rather general process for the dividend drift and the "state." Using a new discretization approach, I obtain analytical expressions for prices and returns for both bonds and stocks and find conditions under which aversion to state uncertainty yields higher expected returns and volatility, and lower interest rates. Indeed, these effects take place, for example, if the state is positively correlated with the drift rate of consumption and negatively correlated with the agents' marginal utility, a situation that occurs for example in external habit formation models.

In the empirical section, I specialize the model to the case where the "state" and the drift rate of dividends are perfectly correlated, a situation that is consistent with standard external habit formation models where agents are unaware of other agents' consumption. In addition, I assume that the economy is hit by structural breaks, which yield a time-variation in the dispersion of the agents posterior density on the underlying drift rate of dividends.

In this set up, when fitted to real consumption data, I show that posterior distributions contain a good deal of uncertainty on the current drift of consumption, a finding which strengthen the notion that investors may be "averse" to this dispersion. Beside matching the unconditional moments of stock returns and the interest rate process, I also show that the time-series of model-generated conditional volatility and price-dividend ratios are broadly consistent with the ones observed in the financial data.

7 References

Abel, Andrew (1990) "Asset prices under habit formation and catching up with the Joneses," *American Economic Review Papers and Proceedings*, 80, 38-42.

Andersen, Torben, Tim Bollerslev, Francis X. Diebold and Clara Vega (2003) "Micro effects of macro-announcements: Real-time price discovery in foreign exchange," *American Economic Review*

Arrow, Kenneth J. (1974) "Optimal insurance and generalized deductible" *Scandinavian Actuarial Journal*, 1-42. Reprinted in *Collected Papers of Kenneth J. Arrow, Volume 3: Individual Choice under Certainty and Uncertainty,* Cambridge University Press, 1984, Cambridge: MA.

Bakshi, Gurdip and Zhiwu Chen (1996) "The spirit of capitalism and stock market prices" *American Economic Review,* 86, 1, 133-157.

Barberis, Nicholas, Ming Huang and Tano Santos (2002) "Prospect theory and asset prices," *Quarterly Journal of Economics.*

Black, Fisher (1976) "Studies of stock price volatility changes," Proceeding of the 1976 meetings of the Business and Economic Statistics Section, American Statistical Association.

Boldrin, Michele, Lawrence J. Christiano and Jonas D. Fisher (2001) "Habit persistence, asset returns, and the business cycle", *American Economic Review*, 91.

Brennan, Michael and Yihong Xia (2001) "Stock price volatility and the equity premium" *Journal of Monetary Economics.*

Cagetti Marco, Lars P. Hansen, Thomas Sargent and Noha Williamn (2002), "Robustness and pricing with uncertain growth," *Review of Financial Studies*, Special Issue.

Campbell, John Y. (1999) "Asset prices, consumption and the business cycle" in John Taylor and Michael Woodford eds. *Handbook of Macroeconomics*, Vol. 1 (North Holland, Amsterdam).

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

Campbell John Y., Andrew W. Lo, and A. Craig MacKinlay (1997) *The Econo-*

metrics of Financial Markets, Princeton University Press, Princeton, NJ.

Campbell John Y. and Luis Viceira (2002) *Strategic Asset Allocation*, Oxford University Press, New York.

Caplin A. J. and J. Leahy (2001) "Psychological expected utility theory, and anticipatory feelings," *Quarterly Journal of Economics,* 106, 1, 55-80.

Chan Y. L. and Leonid Kogan (2002) "Catching up with the Joneses: Heterogeneous Preferences and the Dynamics of Asset Prices" *Journal of Political Economy.*

Constantinides, George (1990), "Habit Formation: A Resolution of the Equity Premium Puzzle," *Journal of Political Economy*, 98, 519-543.

Cutler, David M. and Elizabeth Richardson (1997) "Measuring the Health of the U.S. Population," *Brooking Papers on Economic Activity*, Microeconomics, Volume 1997, 217-271.

David, Alexander (1997) "Fluctuating confidence in stock markets: Implications for returns and volatility," *Journal of Financial and Quantitative Analysis*, 32, 4, 427-462.

David, Alexander, and Pietro Veronesi (2001) "Inflation and earnings uncertainty and the volatility of asset prices: An empirical investigation," CRSP Working paper, University of Chicago.

Detemple, Jerome B., (1986), Asset pricing in a production economy with incomplete information, *Journal of Finance* 41, 383-392.

Detemple, Jerome B. and Fernando Zapatero, (1991), "Asset Prices in an Exchange Economy with Habit Formation," *Econometrica*, Vol. 59, pp. 1633-1657.

Dothan, Michael U., and David Feldman, (1986), Equilibrium interest rates and multiperiod bonds in a partially observable economy, *Journal of Finance 41*, 369-382.

Duffie, Darrell (1996), *Dynamic Asset Pricing Theory*, Princeton University Press, Princeton, New Jersey.

Epstein, Lawrence and Martin Schneider (2003) "Recursive Multiple Priors", Journal of Economic Theory, 113, 1-31.

Epstein, Lawrence and Tan Wang (1994) "Intertemporal Asset Pricing Under Knightian Uncertainty," *Econometrica*, 62, 2, 283-322.

Epstein, Lawrence and Stanley Zin (1989) "Substitution, risk aversion, and the temporal behavior of consumption and asset returns," *Econometrica*, 57, 937-969.

Feldman, David, 1989, "The term structure of interest rates in a partially observable economy," *Journal of Finance* 44, 789-812

Ferson, Wayne and George Constantinides, (1991), "Habit Persistence and Durability in Aggregate Consumption: Empirical Tests," *Journal of Financial Economics,* 29, 199-240.

Fishburn Peter, (1970), Utility theory for decision making, John Wiley and Sons, New York.

Geneakoplos, J., D. Pearce and E. Stacchetti (1989) "Psychological games and sequential rationality" *Games and Economic Behavior*, 1, 60-79.

Gennotte, Gerard, (1986), Optimal portfolio choice under incomplete information, *Journal of Finance* 41, 733-746.

Gilboa Itzack and David Schmeidler (1989) "Maxmin Expected Utility with Non-Unique Prior," *Journal of Mathematical Economics*, 18, 141-153.

Gilboa Itzack and David Schmeidler (1993) "Updating Ambiguous Beliefs," *Journal of Economic Theory*, 59, 33-49.

Hamilton, James D. (1989), "A new approach to the analysis of nonstationary returns and the business cycle." *Econometrica*, 57, 357-384.

Hansen, Lars P., Thomas Sargent and Thomas Tallarini Jr. (1999) "Robust Permanent Income and Pricing," *Review of Economic Studies*, 66, 873-907.

Ingersoll, Jonathan E., (1987), *Theory of Financial Decision Making* (Rowman & Littlefield).

Israel Robert B., Jeffrey S. Rosenthal and Jason Z. Wei (2001) "Finding Generators for Markov Chains via Empirical Transition Matrices, with Applications to Credit Ratings," *Mathematical Finance*, 11,2, 245-265.

Karatzas, Ioannis and Steven E. Shreve (1998), *Methods of Mathematical Finance*, Springer-Verlag, New York.

Karlin S., and H.M. Taylor (1975), *A First Course in Stochastic Processes*, Academic Press, New York.

Karni Edi, (1985) *Decision Making under Uncertainty: The Case of State-Dependent*

Preferences, Harvard University Press, Cambridge, MA.

Kitagawa, Genshiro (1987a) "Non-Gaussian State-Space Modeling of Nonstationary Time Series" *Journal of the American Statistical Association: Theory and Methods*, 82, 400, 1032 - 1041

Kitagawa, Genshiro (1987b) "Non-Gaussian State-Space Modeling of Nonstationary Time Series: A Rejoinder" *Journal of the American Statistical Association: Theory and Methods*, 82, 400, 1060 - 1063

Kushner Harold J. and Paul Dupuis (2001), *Numerical Methods for Stochastic Control Problems in Continuous Time*, Second Edition, Springer, New York.

Liptser, Robert.S. and Albert N. Shiryayev (1977), *Statistics of Random Processes: I, II*, Springer-Verlag, New York.

Lucas, Robert E. (1978) "Asset prices in an exchange economy," *Econometrica*, 46, 1429-1445.

Maenhout, Pascal (1999) "Robust portfolio choice and asset pricing," Working paper, Harvard University.

Myerson, Roger B. (1991) *Game Theory: Analysis of Conflicts*, Harvard University Press, Cambridge, Massachusetts.

Schwert, G. William (1989) "Why does stock market volatility change over time?" *Journal of Finance*, 44, 1115 - 1153.

Skiadas, Costis (1997) "Subjective Probability under Additive Aggregation of Conditional Preferences" *Journal of Economic Theory*, 76, 2, 242-271.

Sundaresan, Suresh (1989), "Intertemporal Dependent Preferences and the Volatility of Consumption and Wealth," *The Review of Financial Studies*, 2, 73-88.

Timmermann, Allan (1993) "How learning in financial markets generates excess volatility and predictability of stock returns" *Quarterly Journal of Economics*, 108, 1135-1145.

Timmermann, Allan (2001) "Structural Breaks, Incomplete Information and Stock Prices" *Journal of Business and Economic Statistics*, (forthcoming).

Veronesi, Pietro (1999) "Stock market overreaction to bad news in good times: A rational expectations equilibrium model," *Review of Financial Studies,* Vol. 12, 5.

Veronesi, Pietro (2000) "How does information quality affect stock returns?" *Journal of Finance*.

Veronesi, Pietro and Francis Yared (1999) "Short and Long Horizon Term and Inflation Risk Premia in the US Term Structure: Evidence from an Integrated Model for Nominal and Real Bond Prices under Regime Shifts," CRSP Working Paper, University of Chicago - GSB.

Viscusi Kip W. and William N. Evans (1990) "Utility Functions That Depend on Health Status: Estimates and Economic Implications," *American Economic Review*, 80, 3, 353 - 374.

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

Yared, Francis (1999) "Path-Dependence in Expected Inflation: Evidence from a New Term Structure Model," Ph.D. dissertation, University of Chicago.

Yarid, Leeat (2001) "Believe and let believe: Axiomatic foundations for belief dependent utility," Cowles Foundation Discussion Paper 1344, Yale University.

Zechauser, Richard J. (1970) "Medical insurance: A case study of the trade off between risk spreading and appropriate incentives" *Journal of Economic Theory*, 2, 1, 10-26.

Zechauser, Richard J. (1973) "Coverage for Catastrophic Illness" *Public Policy*, 21, 149-172.

8 Appendix A: Axiomatic Foundation of State-dependent Utility

There are a number of axiomatic approaches to state-dependent preferences (see e.g. Karni (1985)). I will refer to the Myerson (1991) set of axioms. I start by introducing some notation: For every finite set Z, let $\Delta(Z)$ be the set of probability distributions over Z, that is

$$\Delta(Z) = \left\{ q : Z \to R^{|Z|} | \sum_{y \in Z} q(y) = 1 \text{ and } q(y) \geq 0 \right\}$$

Let $\mathcal{C}$ be the set of possible prizes (= consequences = consumption) the decision maker could get. Let Θ be the set of possible states. I define a lottery to be any function f that specifies a nonnegative real number $f(c|\theta)$ for every $c \in \mathcal{C}$ and for every $\theta \in \Theta$, such that $\sum_{c \in \mathcal{C}} f(c|\theta) = 1$.

That is

$$L = \{f : \Theta \to \Delta(\mathcal{C})\}$$

I will denote by $[c]$ the lotteries giving probability one to the prize $c \in \mathcal{C}$.

I will assume that conditional on each event $S \in \Theta$, the agent will be able to rank lotteries *conditional* on the event S being true. That is, given any two lotteries f and $g \in L$, I will denote $f \succ_S g$ to mean that the agent strictly prefers the lottery f to the lottery g if the event S were true. Similarly, $f \succeq_S g$ denotes weak preference. I will denote Ξ the set of all events in Θ.

Finally, for every two lotteries f and g and scalar $\alpha \in [0,1]$ I will denote by $f\alpha g = \alpha f + (1-\alpha)g$ the lottery assigning probability $\alpha f(c|\theta) + (1-\alpha)g(c|\theta)$ to every $c \in \mathcal{C}$ and for every $\theta \in \Theta$.

The following are Myerson (1991) axioms

Axiom 1.1: (a - *Completeness*) $f \succeq_S g$ or $g \succeq_S f$ and (b - *Transitivity*) $f \succeq_S g$ and $g \succeq_S h$ then $f \succeq_S h$.

Axiom 1.2: (*Relevance*) If $f(.|\theta) = g(.|\theta)$ for all $\theta \in S$, then $f \sim_S g$.

Axiom 1.3: (*Monotonicity*) If $f \succ_S h$ and $0 \leq \beta < \alpha \leq 1$, then $f\alpha h \succ_S f\beta h$.

Axiom 1.4: (*Continuity*) If $f \succeq_S g$ and $g \succeq_S h$, then there exists $\alpha \in [0,1]$ such that $f\alpha h \sim_S g$.

Axiom 1.5: (a - *Objective Substitution*) If $e \succeq_S f$ and $g \succeq_S h$ and $\alpha \in [0,1]$, then $e\alpha g \succeq_S f\alpha h$. (b - *Strict Objective Substitution*) If $e \succ_S f$ and $g \succeq_S h$ and $\alpha \in (0,1]$, then $e\alpha g \succ_S f\alpha h$.

Axiom 1.6: (a - *Subjective Substitution*) If $f \succeq_S g$ and $f \succeq_T g$ and $S \cap T = \varnothing$, then $f \succeq_{S\cup T} g$; (b - Strict Subjective Substitution) If $f \succ_S g$ and $f \succeq_T g$ and $S \cap T = \varnothing$, then $f \succ_{S\cup T} g$;

Axiom 1.7: (*Interest*) For every state in $\theta \in \Theta$, there exist prizes y and z such that $[y] \succ_\theta [z]$

Before stating the representation theorem, I need the following definition:

Definition: A *Conditional Probability Function* on Θ is any function $\pi : \Xi \to \Delta(\Theta)$ such that for every $S \in \Xi$, $\pi(.|S)$ is a well defined probability function, such that $\pi(\theta|S) = 0$ if $\theta \notin S$ and $\sum_{\theta \in S} \pi(\theta|S) = 1$.

The following representation theorem is proved by Myerson (1991), among others.

Theorem 1: Axioms 1.1 - 1.7 are satisfied if and only if there exists a utility function $u : \mathcal{C} \times \Theta \rightarrow R$ and a conditional probability function $\pi : \Xi \longrightarrow \Delta(\Theta)$ such that

(I) $\max_{c \in \mathcal{C}} u(c, \theta) = 1$ and $\min_{c \in \mathcal{C}} u(c, \theta) = 0$

(II) For all R, S, T such that $R \subseteq S \subseteq T \subseteq \Theta$ and $S \neq \varnothing$ we have

$$\pi(R|T) = \pi(R|S)\,\pi(S|T)$$

(III) For all $f, g \in L$ and for all $S \in \Xi$ we have

$$f \succeq_S g \Longleftrightarrow \sum_{\theta \in S} \pi(\theta|S) \sum_{c \in \mathcal{C}} f(c|\theta)\, u(c|\theta) > \sum_{\theta \in S} \pi(\theta|S) \sum_{c \in \mathcal{C}} g(c|\theta)\, u(c|\theta) \tag{39}$$

Proof: See Myerson (1991). ■

For completeness, I also state the axiom that provides state-independent utility functions and the representation theorem:

Axiom 1.8: (*State Neutrality*) For every two states θ and θ', if $f(.|\theta) = f(.|\theta')$, $g(.|\theta) = g(.|\theta')$ and $f \succeq_\theta g$ then $f \succeq_{\theta'} g$.

In this case, we have the following:

Theorem 2: Axioms 1.1 - 1.8 are satisfies if an only if there exists a utility function $u : \mathcal{C} \rightarrow R$ and a conditional probability function $\pi : \Xi \longrightarrow \Delta(\Theta)$ such that (I) and (II) in Theorem 1 are satisfied and in addition

(IV) For all $f, g \in L$ and for all $S \in \Xi$ we have

$$f \succeq_S g \Longleftrightarrow \sum_{\theta \in S} \pi(\theta|S) \sum_{c \in \mathcal{C}} f(c|\theta)\, u(c) \geq \sum_{\theta \in S} \pi(\theta|S) \sum_{c \in \mathcal{C}} g(c|\theta)\, u(c)$$

Proof: See Myerson (1991). ■

One important caveat is that the representation (39) is not unique in the sense that also a conditional probability system $\widehat{\pi}(.|.)$ and a state dependent utility function $\widehat{u}(.|.)$ represents the same conditional preferences over lotteries if (and only if) there exists a positive number A and a function $B : S \rightarrow R$ such that

$$\widehat{\pi}(\theta|S)\,\widehat{u}(c|\theta) = A\pi(\theta|S)\,u(c|\theta) + B(\theta)$$

(see Myerson (1991, Theorem 1.2)). However, Skiadas (1997) provides a set of axioms able to uniquely identify the conditional probability and the state-dependent utility function (see his

Theorem 1, point (a) for a representation as in (39) and point (b) for the uniqueness of the probability and utility representation).

9 Appendix B: Proofs of Propositions

Proof of Proposition 1: Using the definition we have

$$-c\frac{\partial^2 U(c,\pi)}{\partial c^2} = \gamma \frac{\partial U(c,\pi)}{\partial c} \tag{40}$$

Hence, (40) can be rewritten

$$\sum_{\theta} \pi(\theta) \left(\frac{\partial u(c|\theta)}{\partial c}\gamma + c\frac{\partial^2 u(c|\theta)}{\partial c^2} \right) = 0$$

Since this must hold for all θ, we must have

$$\frac{\partial u(c|\theta)}{\partial c}\gamma + c\frac{\partial^2 u(c|\theta)}{\partial c^2} = 0$$

Let $V(c|\theta) = \frac{\partial U(c|\theta)}{\partial c}$, so that $\frac{\partial V(c|\theta)}{\partial c} = -\frac{V(c|\theta)}{c}\gamma$. The solution to this differential equation is $V(c|\theta) = k_2(\theta)\, c^{-\gamma}$. Hence, integrating $V(c|\theta)$ over c we obtain

$$U(c|\theta) = \begin{cases} k_1(\theta) + k_2(\theta)\frac{c^{1-\gamma}}{1-\gamma} & \text{if} \quad \gamma \neq 1 \\ k_1(\theta) + k_2(\theta)\ln(c) & \text{if} \quad \gamma = 1 \end{cases}$$

■

Proof of Proposition 2: Consider an agent with constant relative risk aversion, whose preferences are then given by equation (6). Since θ^* is determined by the equality $U(c,\pi) = E^{\pi}[u(c|\theta)] = u(c|\theta^*)$, we can substitute the functional form for constant relative risk aversion to find the equality

$$E_t^{\pi}[k_1(\theta)] + E_t^{\pi}[k_2(\theta)]\frac{c^{1-\gamma}}{1-\gamma} = k_1(\theta^*) + k_2(\theta^*)\frac{c^{1-\gamma}}{1-\gamma}$$

Pulling terms together, we have

$$k_1(\theta^*) - E_t^{\pi}[k_1(\theta)] = \frac{c^{1-\gamma}}{1-\gamma}\left(E_t^{\pi}[k_2(\theta)] - k_2(\theta^*)\right)$$

This equality is satisfied for all distributions π, and all c if and only if

$$\begin{aligned} E_t^{\pi}[k_1(\theta)] &= k_1(\theta^*) \\ E_t^{\pi}[k_2(\theta)] &= k_2(\theta^*) \end{aligned} \tag{41}$$

We can use a standard 'local' argument to show that $k_i(\theta) = \alpha_i e^{-\rho\theta}$. Specifically, for e.g. $i = 2$ consider a second order Taylor expansion around $\widetilde{\theta} = E^{\pi}[\theta]$ on both sides of (41)

$$\begin{aligned} k_2(\theta) &= k_2\left(\widetilde{\theta}\right) + k_2'\left(\widetilde{\theta}\right)\left(\theta - \widetilde{\theta}\right) + \frac{1}{2}k_2''\left(\widetilde{\theta}\right)\left(\theta - \widetilde{\theta}\right)^2 \\ k_2(\theta^*) &= k_2\left(\widetilde{\theta}\right) + k_2'\left(\widetilde{\theta}\right)\left(\theta^* - \widetilde{\theta}\right) + \frac{1}{2}k_2''\left(\widetilde{\theta}\right)\left(\theta^* - \widetilde{\theta}\right)^2 \end{aligned}$$

The equality $E_t^{\pi}\left[k_2(\theta)\right] = k_2(\theta^*)$ then implies

$$\frac{1}{2}k_2''\left(\widetilde{\theta}\right)V^{\pi}(\theta) = k_2'\left(\widetilde{\theta}\right)\left(\theta^* - \widetilde{\theta}\right) + \frac{1}{2}k_2''\left(\widetilde{\theta}\right)\left(\theta^* - \widetilde{\theta}\right)^2$$

By definition of constant aversion to state uncertainty, we must have $\kappa = \widetilde{\theta} - \theta^*$, a constant. Substituting

$$k_2''\left(\widetilde{\theta}\right)\left(V^{\pi}(\theta) - \kappa^2\right) = -k_2'\left(\widetilde{\theta}\right)2\kappa$$

Assume $\left(V^{\pi}(\theta) - \kappa^2\right) = \sigma_\theta^2 - \kappa^2 \neq 0$. Then

$$k_2''\left(\widetilde{\theta}\right) = k_2'\left(\widetilde{\theta}\right)\left(-\frac{2\kappa}{\sigma_\theta^2 - \kappa^2}\right)$$

is satisfied by $k_2\left(\widetilde{\theta}\right) = \alpha_0 e^{-\rho\widetilde{\theta}}$ with $\rho = \frac{2\kappa}{\sigma_\theta^2 - \kappa^2}$. Finally, for $\gamma > 1$, $c^{1-\gamma}/(1-\gamma) < 0$. Thus, aversion to state uncertainty occurs if $k_2(\theta)$ is globally convex, implying that we need $\rho > 0$. ■

In the following proofs, it is convenient to use the vector notation $\boldsymbol{\sigma} = (\sigma, 0)$ and $\boldsymbol{\sigma}_s = (0, \sigma_s)$. This implies that $\boldsymbol{\Sigma} = diag(\sigma, \sigma_s)$ in Lemma 1 can be written as $\boldsymbol{\Sigma} = (\boldsymbol{\sigma}', \boldsymbol{\sigma}_s')'$.

Proof of Lemma 2: For all $i = 1, .., n$, denote $n_t^{\beta,i} = c_t^{\beta}\pi_t^i$ where β is a constant. Define also the matrix $\overline{\boldsymbol{\Lambda}}_\beta = \boldsymbol{\Lambda} + \beta \times diag\left(g^1, ..., g^n\right) + \frac{1}{2}\beta^2\boldsymbol{\sigma}\boldsymbol{\sigma}'\ \mathbf{I}$. Then, we need to prove that for any $u > t$ we have

$$E\left[n_u^{\beta,i}|\mathcal{F}_t\right] = \sum_{\ell=1}^{N} n_t^{\beta,\ell}\sum_{j=1}^{N} w(\beta)_{jk}^{-1}\, w_{ij}(\beta)\, e^{\omega_j(\beta)(u-t)}$$

where $\omega_j(\beta)$ are the eigenvalues of $\overline{\boldsymbol{\Lambda}}_\beta'$ and $w_{ij}(\beta)$ are associated eigenvectors and $w(\beta)_{ij}^{-1} = [\mathbf{W}^{-1}]_{ij}$. Since $c_t = D_t = \exp(\delta_t)$, we have

$$\frac{dc_t}{c_t} = \left(\sum_{j=1}^{n}\pi_t^j g^j + \frac{1}{2}\boldsymbol{\sigma}\boldsymbol{\sigma}'\right)dt + \boldsymbol{\sigma}d\widetilde{\mathbf{W}}_t = \mu_c(\boldsymbol{\pi}_t)\,dt + \boldsymbol{\sigma}d\widetilde{\mathbf{W}}_t$$

It is convenient to work with the *row* vector $\mathbf{n}_t^{\beta} = \left(n_t^{\beta,1}, .., n_t^{\beta,n}\right)$. By Ito's lemma

$$\begin{aligned}
dn_t^{\beta,i} &= \left\{\left[\mathbf{n}_t^{\beta}\cdot\boldsymbol{\Lambda}\right]_i + \beta n_t^{\beta,i}\mu_c(\boldsymbol{\pi}_t) + \frac{1}{2}\beta(\beta-1)\,n_t^{\beta,i}\boldsymbol{\sigma}\boldsymbol{\sigma}'\right\}dt \\
&\quad + n_t^{\beta,i}\left(\mathbf{v}^i - \overline{\mathbf{v}}_t\right)'\left(\boldsymbol{\Sigma}'\right)^{-1}d\widetilde{\mathbf{W}}_t + \beta n_t^{\beta,i}\boldsymbol{\sigma}d\widetilde{\mathbf{W}}_t + \beta n_t^{\beta,i}\left(\mathbf{v}^i - \overline{\mathbf{v}}_t\right)\left(\boldsymbol{\Sigma}'\right)^{-1}\boldsymbol{\sigma}'dt \\
&= \left\{\left[\mathbf{n}_t^{\beta}\cdot\boldsymbol{\Lambda}\right]_i + \beta n_t^{\beta,i}\mu_c(\boldsymbol{\pi}_t) + \frac{1}{2}\beta(\beta-1)\,n_t^{\beta,i}\boldsymbol{\sigma}\boldsymbol{\sigma}'\right\}dt \\
&\quad + n_t^{\beta,i}\left(\mathbf{v}^i - \overline{\mathbf{v}}_t\right)'\left(\boldsymbol{\Sigma}'\right)^{-1}d\widetilde{\mathbf{W}}_t + \beta n_t^{\beta,i}\boldsymbol{\sigma}d\widetilde{\mathbf{W}}_t + \beta n_t^{\beta,i}\left(\mathbf{v}^i - \overline{\mathbf{v}}_t\right)(1,0)'\,dt
\end{aligned}$$

where the last line stems from the definition of $\boldsymbol{\Sigma} = (\boldsymbol{\sigma}', \boldsymbol{\sigma}_s')'$ so that $(\boldsymbol{\Sigma}')^{-1}\boldsymbol{\sigma}' = \left((\boldsymbol{\sigma}', \boldsymbol{\sigma}_s')\right)^{-1}(\boldsymbol{\sigma}', \boldsymbol{\sigma}_s')(1,0) = \mathbf{I}\times(1,0) = (1,0)$. From $\mu_c(\boldsymbol{\pi}_t) = \sum_{j=1}^{n}\pi_t^j g^j + \frac{1}{2}\boldsymbol{\sigma}\boldsymbol{\sigma}'$ and $\left(\mathbf{v}^i - \overline{\mathbf{v}}_t\right)(1,0)' = g^i - \sum_{j=1}^{n}\pi_t^j g^j$ we obtain

$$dn_t^{\beta,i} = \left\{\left[\mathbf{n}_t^{\beta}\cdot\boldsymbol{\Lambda}\right]_i + \frac{1}{2}\beta^2 n_t^{\beta,i}\boldsymbol{\sigma}\boldsymbol{\sigma}' + \beta n_t^{\beta,i}g^i\right\}dt + n_t^{\beta,i}\left(\boldsymbol{\sigma} + \left(\mathbf{v}^i - \overline{\mathbf{v}}_t\right)'\left(\boldsymbol{\Sigma}'\right)^{-1}\right)d\widetilde{\mathbf{W}}_t \qquad (42)$$

We can write this in vector form:

$$d\mathbf{n}_t^{\beta\,\prime} = \left(\mathbf{n}_t^{\beta}\overline{\mathbf{\Lambda}}_\beta\right)' dt + \mathbf{n}_t^{\beta\,\prime} \odot \mathbf{\Sigma}\left(\boldsymbol{\pi}_t\right) d\widetilde{\mathbf{W}}_t \tag{43}$$

where $\overline{\mathbf{\Lambda}}_\beta$ is given in the claim of the Lemma and $\mathbf{\Sigma}\left(\boldsymbol{\pi}_t\right)$ is some bounded $n \times 2$ matrix. Let

$$\widetilde{\mathbf{n}}_u^{\beta} = E\left[\mathbf{n}_u^{\beta}|\mathcal{F}\left(t\right)\right]$$

We can write (43) in integral form

$$\mathbf{n}_u^{\beta\,\prime} = \mathbf{n}_t^{\beta\,\prime} + \int_t^u \left(\mathbf{n}_s^{\beta}\overline{\mathbf{\Lambda}}_\beta\right)' ds + \int_t^u \mathbf{n}_s^{\beta\,\prime} \odot \mathbf{\Sigma}\left(\boldsymbol{\pi}_s\right) d\widetilde{\mathbf{W}}_s$$

Taking expectations on both sides and using the fact that the stochastic integral has zero expectations, we then have (by also transposing)

$$\widetilde{\mathbf{n}}_u^{\beta} = \mathbf{n}_t^{\beta} + \int_t^u \widetilde{\mathbf{n}}_s^{\beta}\overline{\mathbf{\Lambda}}_\beta ds$$

This can we rewritten as

$$d\widetilde{\mathbf{n}}_t^{\beta} = \widetilde{\mathbf{n}}_t^{\beta}\overline{\mathbf{\Lambda}}_\beta dt$$

Assuming the matrix $\overline{\mathbf{\Lambda}}_\beta$ has n distinct real eigenvalues, the solution to this system of ordinary differential equations with initial condition $\widetilde{\mathbf{n}}_t^{\beta} = \mathbf{n}_t^{\beta}$ is

$$\widetilde{\mathbf{n}}_u^{\beta} = \mathbf{n}_t^{\beta}\ \mathbf{U}_\beta e^{\mathbf{A}_\beta (u-t)} \mathbf{U}_\beta^{-1}$$

where $[\mathbf{A}_\beta]_{jj}$ are the eigenvalues of $\overline{\mathbf{\Lambda}}_\beta$ and $\mathbf{U}_\beta$ are associated column eigenvectors. We then find

$$E\left[n_u^{\beta,i}|\mathcal{F}_t\right] = \sum_{\ell=1}^{N} n_t^{\beta,\ell} \sum_{j=1}^{N} \left[\mathbf{U}_\beta\right]_{\ell j} \left[\mathbf{U}_\beta^{-1}\right]_{ji} e^{\left[\mathbf{A}_\beta\right]_{jj}(u-t)}$$

This concludes the proof of the lemma.■

Proof of proposition 3: We shall need the following technical condition:
Condition A1: Assume

$$E\left[\int_0^\infty e^{-\phi\tau} |D_\tau^{1-\gamma}| d\tau\right] < \infty$$

Since the good is perishable, it is always suboptimal to consume less than D_t and consuming more is not feasible. Hence, we can impose the market clearing condition that $c_t = D_t$ for all $t > 0$. Usual arguments imply that we can use the marginal utility of consumption to discount future consumption. Hence, the price of an asset must satisfy

$$P_t = E_t\left[\int_t^\infty \frac{U_c\left[c_s, s, \boldsymbol{\pi}_s\right]}{U_c\left[c_t, t, \boldsymbol{\pi}_t\right]} D_s ds\right] = \frac{1}{U_c\left[c_t, t, \boldsymbol{\pi}_t\right]} E_t\left[\int_t^\infty e^{-\phi s} \sum_i k_i \pi_s^i c_s^{1-\gamma} ds\right]$$

Noticing that for all s, $c_s^{1-\gamma}\pi_s^i \leq c_s^{1-\gamma}$, invoking condition A1 and Fubini's theorem, we can use the result in Lemma 2 to obtain the value of P_t

$$P_t = \frac{1}{e^{-\phi t}\sum_{i=1}^n k_i \pi_t^i c_t^{-\gamma_i}} \left\{ \int_t^\infty e^{-\phi s} \sum_i k_i E_t \left[c_s^{1-\gamma}\pi_s^i \right] ds \right\}$$

I now apply the result of Lemma 2 and substitute for each i

$$E_t \left[c_s^{1-\gamma}\pi_s^i \right] = E_t \left[n_s^{\beta,i} \right] = \sum_{\ell=1}^N n_t^{\beta,\ell} \sum_{j=1}^N \left[\mathbf{U}_\beta \right]_{\ell j} \left[\mathbf{U}_\beta^{-1} \right]_{ji} e^{\left[\mathbf{A}_\beta \right]_{jj}(s-t)}$$

where $\beta = 1-\gamma$. This yields

$$\begin{aligned}
P_t &= \frac{1}{\sum_{i=1}^n k_i \pi_t^i c_t^{-\gamma}} \left\{ \sum_{i=1}^N k_i \sum_{\ell=1}^N n_t^{\beta,\ell} \sum_{j=1}^N \left[\mathbf{U}_\beta \right]_{\ell j} \left[\mathbf{U}_\beta^{-1} \right]_{ji} \int_t^\infty e^{\left(\left[\mathbf{A}_\beta \right]_{jj} - \phi \right)(s-t)} ds \right\} \\
&= \frac{1}{\sum_{i=1}^n k_i \pi_t^i c_t^{-\gamma}} \left\{ \sum_{i=1}^N \sum_{\ell=1}^N k_i c_t^{1-\gamma} \pi_t^\ell \sum_{j=1}^N \left[\mathbf{U}_\beta \right]_{\ell j} \left[\mathbf{U}_\beta^{-1} \right]_{ji} \frac{1}{\phi - \left[\mathbf{A}_\beta \right]_{jj}} \right\} \\
&= c_t \frac{\sum_{\ell=1}^N k_\ell \pi_t^\ell}{\sum_{i=1}^n k_i \pi_t^i} \left\{ \sum_{i=1}^N \frac{k_i}{k_\ell} \sum_{j=1}^N \left[\mathbf{U}_\beta \right]_{\ell j} \left[\mathbf{U}_\beta^{-1} \right]_{ji} \frac{1}{\phi - \left[\mathbf{A}_\beta \right]_{jj}} \right\} \\
&= c_t \sum_{\ell=1}^N \overline{\pi}_t^\ell B_\ell
\end{aligned}$$

where

$$\begin{aligned}
\overline{\pi}_t^\ell &= \frac{k_\ell \pi_t^\ell}{\sum_{i=1}^n k_i \pi_t^i} \\
B_\ell &= \sum_{i=1}^N \frac{k_i}{k_\ell} \sum_{j=1}^N \left[\mathbf{U}_\beta \right]_{\ell j} \left[\mathbf{U}_\beta^{-1} \right]_{ji} \frac{1}{\phi - \left[\mathbf{A}_\beta \right]_{jj}}
\end{aligned}$$

I finally show that

$$\sum_{j=1}^N \left[\mathbf{U}_\beta \right]_{\ell j} \left[\mathbf{U}_\beta^{-1} \right]_{ji} \frac{1}{\phi - \left[\mathbf{A}_\beta \right]_{jj}} = e_\ell \left(\phi \mathbf{I} - \overline{\mathbf{\Lambda}}_\beta \right)^{-1} e_i$$

which then implies the claim

$$B_\ell = \sum_{i=1}^N \frac{k_i}{k_\ell} e_\ell \left(\phi \mathbf{I} - \overline{\mathbf{\Lambda}}_\beta \right)^{-1} e_i$$

In fact, we know that $\left(\phi \mathbf{I} - \overline{\mathbf{\Lambda}}_\beta \right) = \mathbf{U}_\beta \left(\phi \mathbf{I} - \mathbf{A}_\beta \right) \mathbf{U}_\beta^{-1}$ which implies

$$\left(\phi \mathbf{I} - \overline{\mathbf{\Lambda}}_\beta \right)^{-1} = \mathbf{U}_\beta \left(\phi \mathbf{I} - \mathbf{A}_\beta \right)^{-1} \mathbf{U}_\beta^{-1}$$

Hence

$$e_\ell' \left(\phi\mathbf{I}-\overline{\mathbf{\Lambda}}_\beta\right)^{-1} e_i = e_\ell \mathbf{U}_\beta \left(\phi\mathbf{I} - \mathbf{A}_\beta\right)^{-1} \mathbf{U}_\beta^{-1} e_i = \sum_{j=1}^{n} \frac{\left[\mathbf{U}_\beta\right]_{\ell j} \left[\mathbf{U}_\beta^{-1}\right]_{ji}}{\phi - \left[\mathbf{A}_\beta\right]_{jj}}$$

(b) We know that the real rate of interest rate is given by $r_t = -E_t\left[\frac{d\mathcal{U}_t}{\mathcal{U}_t}\right]$ where $\mathcal{U}_t$ is the real pricing kernel given by $\mathcal{U}_t = \partial U/\partial c = e^{-\phi t} c_t^{-\gamma} \sum_{i=1}^{n} k_i \pi_t^i$. Using Ito's lemma and equation (42), we find that $d\mathcal{U}/\mathcal{U} = -\mu_{\mathcal{U}} dt + \boldsymbol{\sigma}_{\mathcal{U}} d\widetilde{\mathbf{W}}_t$ with

$$\begin{aligned} \mu_{\mathcal{U}} &= \phi + \frac{1}{\sum_{i=1}^{n} k_i \pi_t^i} \left(\gamma \sum_{i=1}^{n} k_i \pi_t^i g^i - \frac{1}{2}\gamma^2 \sum_{i=1}^{n} k_i \pi_t^i \boldsymbol{\sigma}\boldsymbol{\sigma}' - \sum_{i=1}^{n} k_i \left[\boldsymbol{\pi}_t \mathbf{\Lambda}\right]_i \right) \\ \boldsymbol{\sigma}_{\mathcal{U}} &= -\gamma\boldsymbol{\sigma} + \frac{\sum_{i=1}^{n} k_i \pi_t^i \left(\mathbf{v}_i - \overline{\mathbf{v}}_t\right)'}{\sum_{j=1}^{n} k_j \pi_t^j} \left(\mathbf{\Sigma}'\right)^{-1} \end{aligned} \tag{44}$$

Hence $r_t = -E_t\left[\frac{d\mathcal{U}_t}{\mathcal{U}_t}\right] = \mu_{\mathcal{U}}$. By redefining variables we obtain expression (25). ■

Proof of Proposition 4: The price of the asset is

$$P_t = D_t \frac{\sum_{j=1}^{n} \pi_t^j k_j B_j}{\sum_{i=1}^{n} k_i \pi_t^i}$$

In this proof only, let $X_t = \sum_i k_i \pi_t^i$. Hence,

$$dX_t = \sum_i k_i d\pi_t^i = \sum_i k_i \left[\boldsymbol{\pi}_t \mathbf{\Lambda}\right]_i dt + \sum_i k_i \pi_t^i \left(\mathbf{v}_i - \overline{\mathbf{v}}_t\right)' \left(\mathbf{\Sigma}'\right)^{-1} d\widetilde{\mathbf{W}}_t$$

or

$$\begin{aligned} \frac{dX_t}{X_t} &= \frac{\sum_i k_i \left[\boldsymbol{\pi}_t \mathbf{\Lambda}\right]_i}{\sum_i k_i \pi_t^i} dt + \frac{\sum_i k_i \pi_t^i \left(\mathbf{v}_i - \overline{\mathbf{v}}_t\right)'}{\sum_i k_i \pi_t^i} \left(\mathbf{\Sigma}'\right)^{-1} d\widetilde{\mathbf{W}} \\ &= \mu_{X,t} dt + \boldsymbol{\sigma}_{X,t} d\widetilde{\mathbf{W}}_t \end{aligned} \tag{45}$$

with

$$\boldsymbol{\sigma}_{X,t} = \frac{\sum_i k_i \pi_t^i \left(\mathbf{v}_i - \overline{\mathbf{v}}_t\right)'}{\sum_i k_i \pi_t^i} \left(\mathbf{\Sigma}'\right)^{-1}$$

It is convenient to define $\widetilde{P}_t = D_t \sum_{i=1}^{n} \pi_t^i k_i B_i$. From Ito's lemma we obtain

$$\begin{aligned} d\widetilde{P}_t &= D_t \sum_i k_i B_i d\pi_t^i + \sum_{i=1}^{n} \pi_t^i k_i B_i dD_t + \sum_{i=1}^{n} k_i B_i d\pi_t^i dD_t \\ &= D_t \sum_i k_i B_i \left[\boldsymbol{\pi}_t \mathbf{\Lambda}\right]_i dt + D_t \sum_i k_i B_i \pi_t^i \left(\mathbf{v}_i - \overline{\mathbf{v}}_t\right)' \left(\mathbf{\Sigma}'\right)^{-1} d\widetilde{\mathbf{W}}_t \\ &\quad + \sum_{i=1}^{n} \pi_t^i k_i B_i D_t \mu_{D,t} dt + \sum_{i=1}^{n} \pi_t^i k_i B_i \boldsymbol{\sigma} d\widetilde{\mathbf{W}}_t \\ &\quad + \sum_{i=1}^{n} k_i B_i D_t \pi_t^i \left(\mathbf{v}_i - \overline{\mathbf{v}}_t\right)' \left(\mathbf{\Sigma}'\right)^{-1} \boldsymbol{\sigma}' dt \end{aligned}$$

Notice now again that by definition $(\boldsymbol{\Sigma}')^{-1}= ((\boldsymbol{\sigma}', \boldsymbol{\sigma}'_s))^{-1}$ which implies $(\boldsymbol{\Sigma}')^{-1}\boldsymbol{\sigma}' = ((\boldsymbol{\sigma}', \boldsymbol{\sigma}'_s))^{-1}(\boldsymbol{\sigma}', \boldsymbol{\sigma}'_s)(1,0)$ $(1,0)$ Hence, we have

$$(\mathbf{v}_i - \overline{\mathbf{v}}_t)'\left(\boldsymbol{\Sigma}'\right)^{-1}\boldsymbol{\sigma}' = \left(g^i - \sum_{j=1}^{n}\pi_t^j g^j\right)$$

which yields

$$d\widetilde{P}_t = \widetilde{P}_t\widetilde{\mu}_{P,t}dt + \widetilde{P}_t\widetilde{\boldsymbol{\sigma}}_{P,t}d\widetilde{\mathbf{W}}_t$$

where

$$\begin{aligned}
\widetilde{\mu}_{P,t} &= \mu_{D,t} + \frac{\sum_{i=1}^{n} k_i B_i \left[\boldsymbol{\pi}_t\boldsymbol{\Lambda}\right]_i}{\sum_{i=1}^{n}\pi_i k_i B_i} + V_t^B \\
\widetilde{\boldsymbol{\sigma}}_{P,t} &= \boldsymbol{\sigma} + \frac{\sum_i k_i B_i \pi_t^i (\mathbf{v}_i - \overline{\mathbf{v}}_t)'}{\sum_i k_i B_i \pi_t^i}\left(\boldsymbol{\Sigma}'\right)^{-1}
\end{aligned}$$

Notice that since $P_t = \widetilde{P}_t / X_t$ we also find

$$\frac{dP_t}{P_t} = \left(\widetilde{\mu}_{P,t} - \mu_{X,t} + \boldsymbol{\sigma}_{X,t}\boldsymbol{\sigma}'_{X,t} - \widetilde{\boldsymbol{\sigma}}_{P,t}\boldsymbol{\sigma}'_{X,t}\right)dt + \left(\widetilde{\boldsymbol{\sigma}}_{P,t} - \boldsymbol{\sigma}_{X,t}\right)d\widetilde{\mathbf{W}}_t$$

and hence

$$\begin{aligned}
\boldsymbol{\sigma}_{P,t} &= \boldsymbol{\sigma} + \left(\frac{\sum_i k_i B_i \pi_t^i (\mathbf{v}_i - \overline{\mathbf{v}}_t)'}{\sum_i k_i B_i \pi_t^i} - \frac{\sum_i k_i \pi_t^i (\mathbf{v}_i - \overline{\mathbf{v}}_t)'}{\sum_i k_i \pi_t^i}\right)\left(\boldsymbol{\Sigma}'\right)^{-1} \\
&= \boldsymbol{\sigma} + \left(V_t^{kB,g} - V_t^{k,g}, V_t^{kB,\theta} - V_t^{k,\theta}\right)\left(\boldsymbol{\Sigma}'\right)^{-1} \\
&= \boldsymbol{\sigma} + \left(\Delta V_t^g, \Delta V_t^\theta\right)\left(\boldsymbol{\Sigma}'\right)^{-1}
\end{aligned}$$

The equilibrium condition requires that the excess return

$$dR_t = \frac{dP_t + D_t dt}{P_t} - r_t dt$$

is such that

$$E_t\left[dR_t\right] = -Cov\left(dR_t, \frac{d\mathcal{U}}{\mathcal{U}}\right) = \boldsymbol{\sigma}_P\boldsymbol{\sigma}'_{\mathcal{U}}$$

Using (44)

$$\begin{aligned}
\boldsymbol{\sigma}_{\mathcal{U}} &= -\gamma\boldsymbol{\sigma} + \left(V_t^{k,g}, V_t^{k.\theta}\right)\left(\boldsymbol{\Sigma}'\right)^{-1} \\
&= -\gamma\boldsymbol{\sigma} + \boldsymbol{\sigma}_{X,t}
\end{aligned}$$

we finally find

$$\begin{aligned}
E_t\left[dR_t\right] &= -\left(\widetilde{\boldsymbol{\sigma}}_{P,t} - \boldsymbol{\sigma}_{X,t}\right)\boldsymbol{\sigma}_{\mathcal{U},t}dt \\
&= -\left(\boldsymbol{\sigma} + \left(\Delta V_t^g, \Delta V_t^\theta\right)\left(\boldsymbol{\Sigma}'\right)^{-1}\right)\left(-\gamma\boldsymbol{\sigma} + \left(V_t^{k,g}, V_t^{k.\theta}\right)\left(\boldsymbol{\Sigma}'\right)^{-1}\right)' \\
&= \gamma\left(\boldsymbol{\sigma}\boldsymbol{\sigma}' + \Delta V_t^g\right) - V_t^{k,g} - \left(\Delta V_t^g, \Delta V_t^\theta\right)\left(\boldsymbol{\Sigma}\boldsymbol{\Sigma}'\right)^{-1}\left(V_t^{k,g}, V_t^{k.\theta}\right)'
\end{aligned}$$

which yields (29) ■

Proof of Proposition 5: (a) and (b) are immediate from the definition of $V_t^{k,\theta}$ and $V_t^{k,g}$. As for part (c), consider the case $x = g$. For $x = \theta$ the proof is analogous. By definition

$$\begin{aligned}\Delta V_t^g &= V_t^{kB,g} - V_t^{k,g} = \frac{E_t\left[k\left(\theta\right)B(\theta,g)g\right]}{E_t\left[k\left(\theta\right)B(\theta,g)\right]} - \frac{E_t\left[k\left(\theta\right)g\right]}{E_t\left[k\left(\theta\right)\right]} \\ &= E_t^{kB}\left[g\right] - E_t^{k}\left[g\right]\end{aligned}$$

where the expectations $E_t^{kB}[.]$ and $E_t^{k}[.]$ are taken with respect the two new densities

$$p_t^{kB}(\theta,g) = \frac{p_t(\theta,g)k(\theta)B(\theta,g)}{E_t\left[k\left(\theta\right)B(\theta,g)\right]} \text{ and } p_t^{k}(\theta,g) = \frac{p_t(\theta,g)k(\theta)}{E_t\left[k\left(\theta\right)\right]}$$

Then, the point wise ratio is

$$\frac{p_t^{kB}(\theta,g)}{p_t^{k}(\theta,g)} = B\left(\theta,g\right)\frac{E_t\left[k\left(\theta\right)\right]}{E_t\left[k\left(\theta\right)B(\theta,g)\right]}$$

Since $B\left(\theta,g\right)$ is non decreasing in both arguments, also the ratio $\frac{p_t^{kB}(\theta,g)}{p_t^{k}(\theta,g)}$ is non-decreasing in both θ and g. Because both densities must integrate to 1, we have that for every θ, $p_t^{kB}\left(\theta,g\right) \leq p_t^{k}(\theta,g)$ for $g \leq g^*$ and $p_t^{kB}(\theta,g) \geq p_t^{k}(\theta,g)$ for $g > g^*$ for some cutoff value g^*. This immediately yields the result. Similarly for $x = \theta$. ■

Appendix C

In this appendix, I show that when time is discrete, the price of the stock is given by

$$P_t = D_t\ \left(\overline{\boldsymbol{\pi}}_t' \cdot \mathbf{B}\right) \tag{46}$$

where $\mathbf{B} = \mathbf{k}^{-1} \odot \left(\mathbf{I} - \overline{\boldsymbol{\psi}}_\beta\right)^{-1}\mathbf{D}_\beta$ with $\overline{\boldsymbol{\psi}}_\beta = \boldsymbol{\psi} \times diag\left(e^{-\phi+\beta g^j+\frac{1}{2}\beta^2\sigma^2}\right)$, $\mathbf{D}_\beta = [D_{\beta,1},..,D_{\beta,n}]$, and $D_{\beta,i} = e^{-\phi+\frac{1}{2}\beta^2\sigma^2}\sum_{j=1}^{n}\lambda_{ij}e^{\beta g^j}k_j$. Similarly, the price of a long-term bond is

$$Q_{t,t+\tau} = \overline{\boldsymbol{\pi}}_t' \cdot \mathbf{Q}_\tau \tag{47}$$

where $\mathbf{Q}_\tau \approx \mathbf{k}^{-1} \odot \left(\eta\overline{\boldsymbol{\Psi}}_{\beta-1}\right)^{\tau} \cdot \mathbf{k}$. Clearly, the one period risk free rate would then be $R_t = 1/Q_{t,t+1}$.

Proof: Let $\beta = 1 - \gamma$. Then, in discrete time the pricing formula is

$$P_t = \frac{c_t}{c_t^\beta \boldsymbol{\pi}_t' \cdot \mathbf{k}} E_t\left[\sum_{\tau=1}^{\infty} e^{-\phi\tau} c_{t+\tau}^{\beta} \boldsymbol{\pi}_{t+\tau}' \cdot \mathbf{k}\right]$$

Using the law of iterated expectations, we can write the expectation as

$$E_t\left[\sum_{\tau=1}^{\infty} e^{-\phi\tau} c_{t+\tau}^{\beta} \boldsymbol{\pi}_{t+\tau}' \cdot \mathbf{k}\right] = E_t\left[\sum_{\tau=1}^{\infty} e^{-\phi\tau} c_{t+\tau}^{\beta} E_{t+\tau}\left[k\left(\theta_{t+\tau}\right)\right]\right] = c_t^\beta E_t\left[\sum_{\tau=1}^{\infty} e^{-\phi\tau+\beta(\log(c_{t+\tau})-\log(c_t))} k\left(\theta_{t+\tau}\right)\right]$$

Since

$$\log(c_{t+\tau}) - \log(c_t) = \sum_{j=1}^{\tau} g_{t+j} + \sigma \sum_{j=1}^{\tau} \varepsilon_{t+j}$$

and substituting $k(\theta_{t+\tau}) = \alpha e^{-\rho\theta_{t+\tau}}$ it is clear that the last conditional expectation can be written as

$$E_t\left[\sum_{\tau=1}^{\infty} e^{-\phi\tau+\beta(\log(c_{t+\tau})-\log(c_t))} k(\theta_{t+\tau})\right] = \sum_{i=1}^{n} E\left[\sum_{\tau=1}^{\infty} \alpha e^{-\phi\tau+\beta\sum_{j=1}^{\tau} g_{t+j}+\beta\sigma\sum_{j=1}^{\tau}\varepsilon_{t+j}-\rho\theta_{t+\tau}} | \nu_t = \nu^i\right] \pi_t^i$$

Let

$$\widetilde{B}_i = E\left[\sum_{\tau=1}^{\infty} \alpha e^{-\phi\tau+\beta\sum_{j=1}^{\tau} g_{t+j}+\beta\sigma\sum_{j=1}^{\tau}\varepsilon_{t+j}-\rho\theta_{t+\tau}} | \nu_t = \nu^i\right]$$

We can write

$$\begin{aligned}
\widetilde{B}_i &= E\left[\alpha e^{-\phi+\beta g_{t+1}+\beta\sigma\varepsilon_{t+1}} e^{-\rho\theta_{t+1}} | \nu_t = \nu^i\right] \\
&\quad + E\left[\sum_{\tau=2}^{\infty} \alpha e^{-\phi\tau+\beta\sum_{j=1}^{\tau} g_{t+j}+\beta\sigma\sum_{j=1}^{\tau}\varepsilon_{t+j}} e^{-\rho\theta_{t+\tau}} | \nu_t = \nu^i\right] \\
&= \alpha e^{-\phi+\frac{1}{2}\beta^2\sigma^2} \sum_{j=1}^{n} \lambda_{ij} e^{\beta g^j - \rho\theta^j} \\
&\quad + \sum_{j=1}^{n} \lambda_{ij} e^{-\phi+\beta g^j+\frac{1}{2}\beta^2\sigma^2} E\left[\sum_{\tau=2}^{\infty} \alpha e^{-\phi\tau+\beta\sum_{j=2}^{\tau} g_{t+j}+\beta\sigma\sum_{j=2}^{\tau}\varepsilon_{t+j}} e^{-\rho\theta_{t+\tau}} | \nu_{t+1} = \nu^j\right]
\end{aligned}$$

That is

$$\widetilde{B}_i = e^{-\phi+\frac{1}{2}\beta^2\sigma^2} \sum_{j=1}^{n} \lambda_{ij} e^{\beta g^j} \alpha e^{-\rho\theta^j} + \sum_{j=1}^{n} \lambda_{ij} e^{-\phi+\beta g^j+\frac{1}{2}\beta^2\sigma^2} \widetilde{B}_j$$

leading to

$$\left(\mathbf{I} - \overline{\boldsymbol{\psi}}_\beta\right) \widetilde{\mathbf{B}} = \mathbf{D}$$

and thus $\widetilde{\mathbf{B}} = \left(\mathbf{I} - \overline{\boldsymbol{\Psi}}_\beta\right)^{-1} \mathbf{D}$ where $\overline{\boldsymbol{\Psi}}_\beta$ and $\mathbf{D}$ are defined in the text of the proposition. Substituting in the pricing function and using $c_t = D_t$, we obtain

$$P_t = \frac{c_t}{\boldsymbol{\pi}_t' \cdot \mathbf{k}} \boldsymbol{\pi}_t' \cdot \widetilde{\mathbf{B}} = D_t \times \overline{\boldsymbol{\pi}}_t' \cdot \mathbf{B}$$

where $B_i = k_i^{-1} \widetilde{B}_i$ and $\overline{\pi}_t^i = \frac{\pi_t^i k_i}{\boldsymbol{\pi}_t' \cdot \mathbf{k}}$.

Similarly, we can compute the price of any bond:

$$Q_{t,t+\tau} = E_t\left[\frac{m_{t+\tau}}{m_t}\right] = \frac{1}{c_t^{\beta-1} \boldsymbol{\pi}_t' \cdot \mathbf{k}} E_t\left[e^{-\phi\tau} c_{t+\tau}^{\beta-1} \boldsymbol{\pi}_{t+\tau}' \cdot \mathbf{k}\right] = \frac{1}{\boldsymbol{\pi}_t' \cdot \mathbf{k}} E_t\left[e^{-\phi\tau+(\beta-1)(\log(c_{t+\tau})-\log(c_t))-\rho\theta_{t+\tau}} \alpha\right]$$

Notice

$$E_t\left[e^{-\phi\tau+(\beta-1)(\log(c_{t+\tau})-\log(c_t))-\rho\theta_{t+\tau}} \alpha\right] = \sum_{i=1}^{n} E_t\left[e^{-\phi\tau+(\beta-1)(\log(c_{t+\tau})-\log(c_t))-\rho\theta_{t+\tau}} \alpha | \boldsymbol{\nu}_t = \boldsymbol{\nu}^i\right] \pi_t^i$$

Let

$$
\begin{aligned}
Q^i_\tau &= E_t\left[e^{-\phi\tau+(\beta-1)(\log(c_{t+\tau})-\log(c_t))-\rho\theta_{t+\tau}}\alpha|\boldsymbol{\nu}_t=\boldsymbol{\nu}^i\right] \\
&= E\left[e^{-\phi\tau+(\beta-1)\sum_{j=1}^{\tau}g_{t+j}+\sum_{j=1}^{\tau}(\beta-1)\sigma\varepsilon_{t+j}-\rho\theta_{t+\tau}}\alpha|v_t=v_i\right] \\
&= E\left[e^{-\phi+(\beta-1)g_{t+1}+(\beta-1)\sigma\varepsilon_{t+1}}e^{-\phi(\tau-1)+(\beta-1)\sum_{j=2}^{\tau}g_{t+j}+\sum_{j=2}^{\tau}(\beta-1)\sigma\varepsilon_{t+j}-\rho\theta_{t+\tau}}\alpha|v_t=v_i\right] \\
&= \sum_{j=1}^{n}\lambda_{ij}E\left[e^{-\phi+(\beta-1)g_{t+1}+(\beta-1)\sigma\varepsilon_{t+1}}e^{-\phi(\tau-1)+(\beta-1)\sum_{j=2}^{\tau}g_{t+j}+\sum_{j=2}^{\tau}(\beta-1)\sigma\varepsilon_{t+j}-\rho\theta_{t+\tau}}\alpha|v_{t+1}=v_j\right] \\
&= \sum_{j=1}^{n}\lambda_{ij}e^{-\phi+(\beta-1)g_j+\frac{1}{2}(\beta-1)^2\sigma^2}Q^j_{\tau-1}
\end{aligned}
$$

We obtain the difference equation

$$\mathbf{Q}_\tau=\overline{\boldsymbol{\Psi}}_{\beta-1}\mathbf{Q}_{\tau-1}$$

where

$$\overline{\boldsymbol{\Psi}}_{\beta-1}=\boldsymbol{\Psi}\times diag\left(e^{-\phi+(\beta-1)g_j+\frac{1}{2}(\beta-1)^2\sigma^2}\right)$$

Since by definition $Q^i_0=E_t\left[e^{-\rho\theta_t}\alpha|\boldsymbol{\nu}_t=\boldsymbol{\nu}^i\right]=k_i$ we have the final condition $\mathbf{Q}_0=\mathbf{k}$, which then yields

$$\mathbf{Q}_\tau=\left(\overline{\boldsymbol{\Psi}}_{\beta-1}\right)^\tau\mathbf{k}$$

The price of the bond is then

$$Q_{t,t+\tau}=\frac{1}{\boldsymbol{\pi}'_t\cdot\mathbf{k}}E_t\left[e^{-\phi\tau+(\beta-1)(c_{t+\tau}-c_t)-\rho\theta_{t+\tau}}\alpha\right]=\frac{\boldsymbol{\pi}'_t\cdot\mathbf{Q}_\tau}{\boldsymbol{\pi}'_t\cdot\mathbf{k}}$$

TABLE I
Parameter Estimates

Panel A: Smooth Autoregressive Model

μ	σ	a	σ_g	Likelihood
0.0012	0.0057	0.7888	0.0018	726.7021
0.0005	0.0005	0.0788	0.0005	

Panel B: Autoregressive with Symmetric Jumps

μ	σ	a	σ_g	p	σ_ζ	Likelihood
0.0006	0.0057	0.8834	0.0005	0.0767	0.0056	728.245
0.0003	0.0004	0.0511	0.0001	0.0586	0.0021	

Panel C: Autoregressive with Asymmetric Jumps

μ	σ	a	σ_g	p	σ_ζ	η_ζ	Likelihood
0.0006	0.0056	0.8906	0.0004	0.1084	277260	67748000	728.6236
0.0003	0.0005	0.0534	0.0001	0.1213	86942	38968000	

This table reports the Maximum Likelihood estimates of the model for consumption growth, $\Delta \log(c_{t+1}) = g_t + \sigma\ \varepsilon_{t+1}$ where g_t follows the process

$$g_{t+1} = \begin{cases} \mu + a g_t + \sigma_g \varepsilon_{g,t+1} & \text{with prob.} \quad 1-p \\ \zeta_{t+1} & \text{with prob.} \quad p \end{cases}$$

and

$$\zeta_{t+1} \sim F_\zeta \text{ with } E_F\left[\zeta_{t+1}\right] = \frac{\mu}{1-a}$$

with $E[\varepsilon_t, \varepsilon_{g,t}] = 0$. In Panel A, $p = 0$. In Panel B, the distribution F_ζ is a normal distribution with variance σ_ζ^2, while in Panel C the distribution F_ζ is the Weibul distribution

$$F_\zeta(x) = I_{\{x \geq \mu_\zeta\}} \eta_\zeta \left(\frac{x-\mu_\zeta}{\sigma_\zeta}\right)^{\eta_\zeta - 1} e^{-\left(\frac{x-\mu_\zeta}{\sigma_\zeta}\right)^{\eta_\zeta}} \quad \text{for } \eta_\zeta, \sigma_\zeta > 0$$

Estimates are obtained by discretizing the interval $[-0.0461, 0.0536]$ in $n = 100$ intervals and applying the ML estimation method as in Kitagawa (1987) and Hamilton (1989). Standard errors are (Newey-West) corrected for heteroskedasticity and autocorrelation.

TABLE II
Calibration

Panel A: Unconditional Moments (1952 - 2001)

	$E[R]$	$\sigma(R)$	r_f	$\sigma(r_f)$	SR	$\overline{P/D}$
Data	0.0719	0.1651	0.0153	0.0105	0.4355	32.04

Panel B: Smooth Autoregressive Model

γ	ρ	ϕ	$E[R]$	$\sigma(R)$	r_f	$\sigma(r_f)$	SR	$\overline{P/D}$
1.5	0	0.01	0.0006	0.0091	0.0435	0.0027	0.0659	47.72
1.5	0	0.02	0.0006	0.0092	0.0536	0.0027	0.0652	32.23
1.5	40	0.01	0.0140	0.0941	0.0346	0.0167	0.1488	48.19
1.5	40	0.02	0.0139	0.0934	0.0447	0.0167	0.1488	32.55
1.5	80	0.01	0.0415	0.1798	0.0196	0.0359	0.2308	49.13
1.5	80	0.02	0.0412	0.1783	0.0296	0.0360	0.2311	33.18
3.0	0	0.01	-0.0002	0.0022	0.0769	0.0055	-0.0909	18.73
3.0	0	0.02	-0.0002	0.0021	0.0871	0.0055	-0.0952	15.74
3.0	40	0.01	0.0134	0.0803	0.0662	0.0141	0.1669	18.94
3.0	40	0.02	0.0133	0.0797	0.0764	0.0141	0.1669	15.91
3.0	80	0.01	0.0407	0.1635	0.0493	0.0334	0.2489	19.33
3.0	80	0.02	0.0404	0.1622	0.0595	0.0335	0.2491	16.24

Panel C: Autoregressive with Symmetric Jumps

γ	ρ	ϕ	$E[R]$	$\sigma(R)$	r_f	$\sigma(r_f)$	SR	$\overline{P/D}$
1.5	0	0.01	0.0001	0.0080	0.0428	0.0029	0.0125	47.96
1.5	0	0.02	0.0001	0.0080	0.0529	0.0029	0.0125	32.34
1.5	40	0.01	0.0117	0.0954	0.0359	0.0148	0.1226	48.36
1.5	40	0.02	0.0116	0.0946	0.0460	0.0148	0.1226	32.61
1.5	80	0.01	0.0409	0.1874	0.0202	0.0341	0.2182	49.41
1.5	80	0.02	0.0406	0.1855	0.0303	0.0342	0.2189	33.31
3.0	0	0.01	-0.0004	0.0089	0.0756	0.0059	-0.0449	18.93
3.0	0	0.02	-0.0004	0.0087	0.0858	0.0059	-0.0460	15.88
3.0	40	0.01	0.0110	0.0780	0.0668	0.0120	0.1410	19.12
3.0	40	0.02	0.0109	0.0774	0.0770	0.0120	0.1408	16.04
3.0	80	0.01	0.0397	0.1664	0.0491	0.0314	0.2386	19.56
3.0	80	0.02	0.0394	0.1648	0.0592	0.0315	0.2391	16.40

TABLE II (cntd.)

Panel D: Autoregressive with Asymmetric Jumps

γ	ρ	ϕ	$E[R]$	$\sigma(R)$	r_f	$\sigma(r_f)$	SR	$\overline{P/D}$
1.5	0	0.01	0.0004	0.0083	0.0420	0.0034	0.0482	48.48
1.5	0	0.02	0.0004	0.0083	0.0522	0.0034	0.0482	32.58
1.5	40	0.01	0.0163	0.1122	0.0332	0.0209	0.1453	48.96
1.5	40	0.02	0.0162	0.1113	0.0433	0.0210	0.1456	32.90
1.5	80	0.01	0.0654	0.2213	0.0061	0.0494	0.2955	50.45
1.5	80	0.02	0.0648	0.2193	0.0161	0.0495	0.2955	33.89
3.0	0	0.01	-0.0001	0.0111	0.0740	0.0069	-0.0090	19.28
3.0	0	0.02	-0.0001	0.0109	0.0842	0.0069	-0.0092	16.12
3.0	40	0.01	0.0157	0.0946	0.0628	0.0172	0.1660	19.51
3.0	40	0.02	0.0155	0.0939	0.0730	0.0172	0.1651	16.31
3.0	80	0.01	0.0641	0.2007	0.0328	0.0454	0.3194	20.15
3.0	80	0.02	0.0636	0.1989	0.0429	0.0456	0.3198	16.84

Panel A reports the ex-post mean excess stock returns, its volatility, the level and volatility of the real-interest rate, the average Sharpe ratio and the average price-dividend ratio for the sample 1952-2001. Panel B-D report the results of the simulation of 4,000 quarters of artificial data for the same moments implied by the autoregressive model for g_t, with no jumps, with a symmetric jump distribution and with an asymmetric jump distribution, according to the estimates in Panels A - C in Table I.

TABLE III
Time-Varying Volatility

Panel A: Data (1952 - 2001)

	ω	β	α	
GARCH(1,1)	0.0013	0.6292	0.2606	
	5.00E-04	0.0857	0.0745	
	ω	β	α	c
EGARCH(1,1)	-0.3307	0.9123	0.164	0.2511
	0.0111	0.0033	0.0015	0.1063

Panel B: Smooth Autoregressive Model

	ω	β	α	
GARCH(1,1)	1.40E-03	0.7941	1.0259e-007	
	1.40E-05	7.90E-05	5.00E-06	
	ω	β	α	c
EGARCH(1,1)	-1.2791	0.4894	0.0004	41.21
	0.0301	0.0114	0.00023	.9672

Panel C: Autoregressive with Symmetric Jumps

	ω	β	α	
GARCH(1,1)	0.003	0.2766	0.3245	
	2.00E-04	0.0377	0.0294	
	ω	β	α	c
EGARCH(1,1)	-1.4407	0.509	0.2717	0.0121
	0.0475	0.025	0.0257	0.0398

Panel D: Autoregressive with Asymmetric Jumps

	ω	β	α	
GARCH(1,1)	0.0044	0.2993	0.2623	
	0.0005	0.0632	0.0250	
	ω	β	α	c
EGARCH(1,1)	-1.1257	0.5778	0.1771	0.4151
	0.0038	0.0005	0.0007	0.1831

Panel A reports the estimates of a GARCH(1,1) and a EGARCH(1,1) model, fitted to quarterly returns in the samepl 1952 - 2001. The GARCH(1,1) and EGARCH(1,1) model assume that returns are conditionally normally distributed

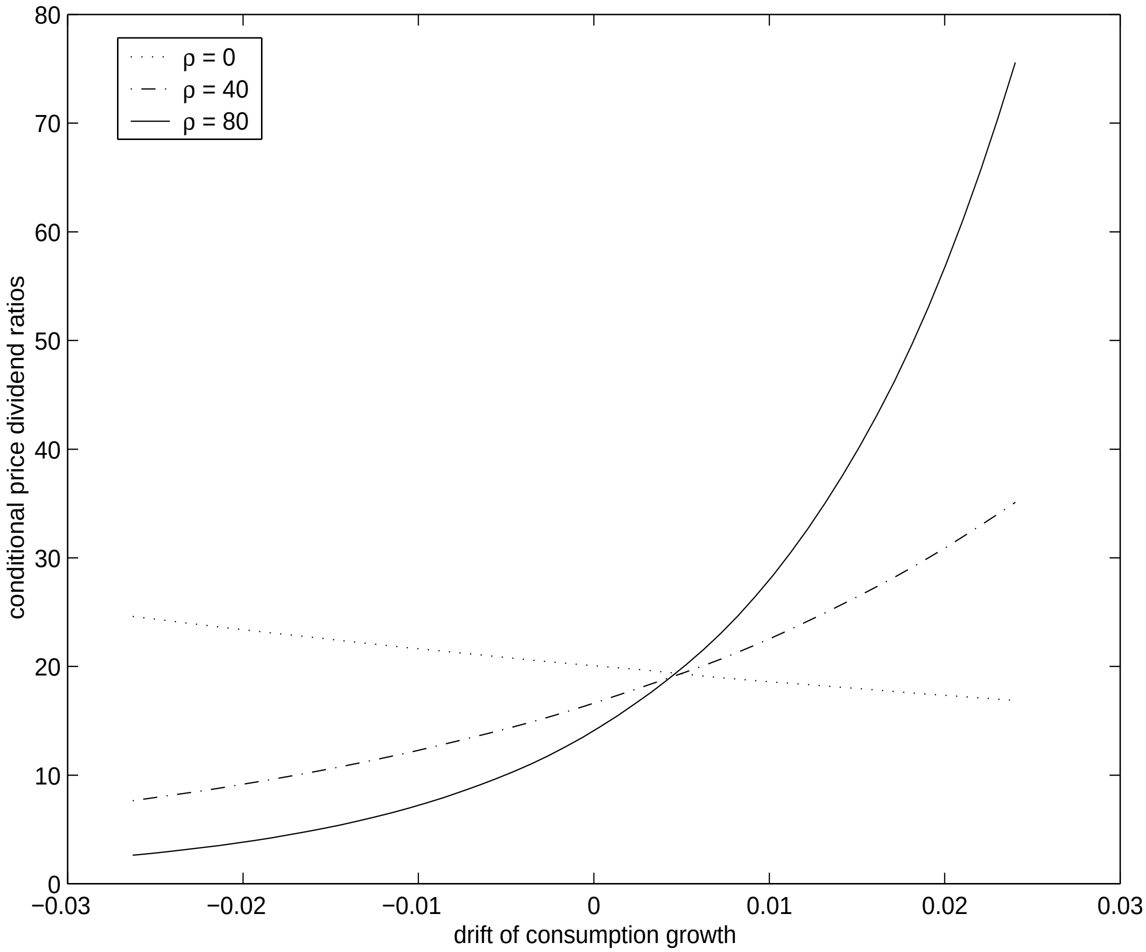
$R_{t+1} = \mu + \sigma_t \varepsilon_{t+1}$, $\varepsilon_{t+1} \sim \mathcal{N}(0,1)$, with volatilty given by, respectively:

$$\text{GARCH(1,1)} \quad : \quad \sigma_{t+1}^2 = \omega + \beta \sigma_t^2 + \alpha R_{t+1}^2 \tag{1}$$

$$\text{EGARCH(1,1)} \quad : \quad \log(\sigma_{t+1}) = \omega + \beta \log(\sigma_t) + \alpha \left[|\varepsilon_{t+1}| - c\varepsilon_{t+1} \right] \tag{2}$$

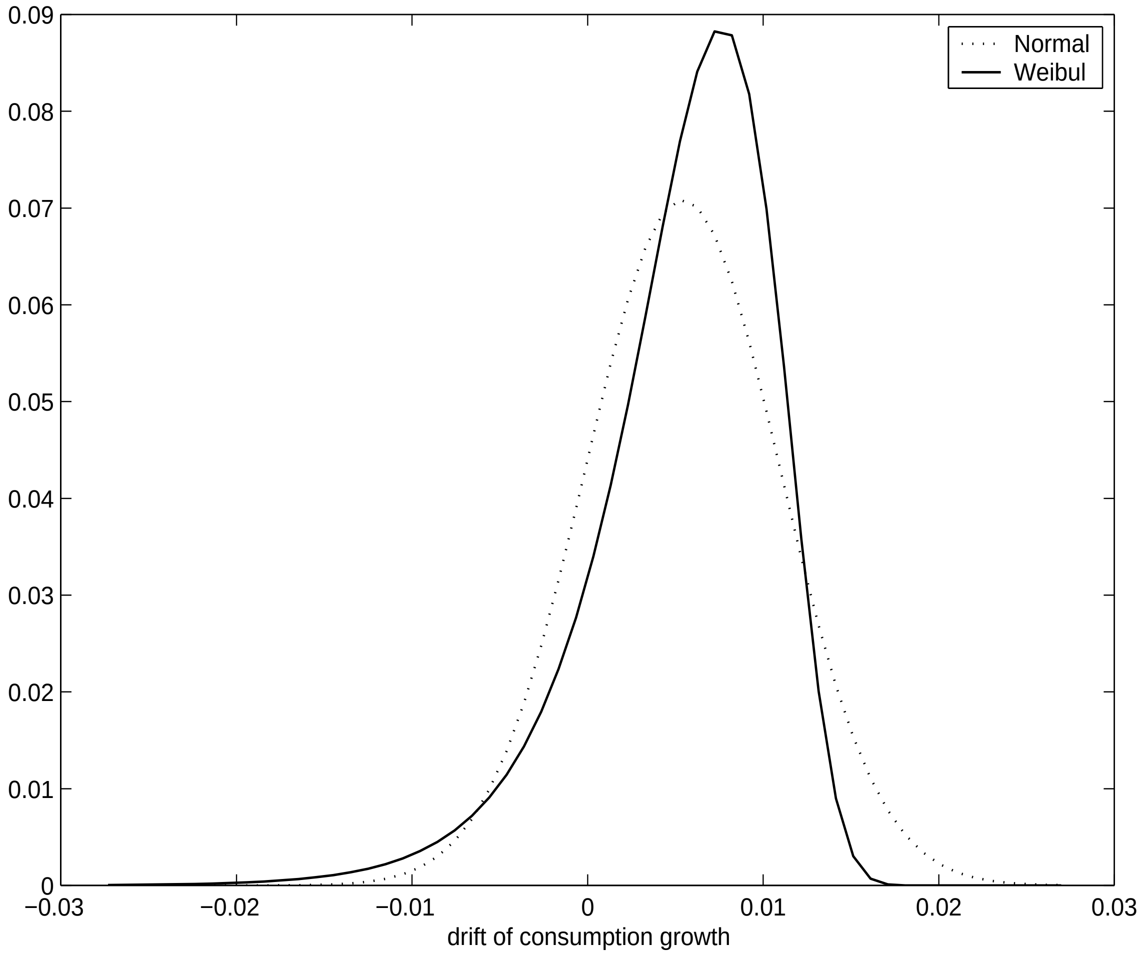
Panels B - D report the estimates of the same model for 4,000 quarters of simulated returns data. The simulations were performed for the three models for consumption growth, autoregressive with no jumps, with symmetric jump distribution and for asymmetric (negatively skewed) jump distribution, as estimated in Table I from consumption data.

Figure 1: Conditional price-dividend ratios



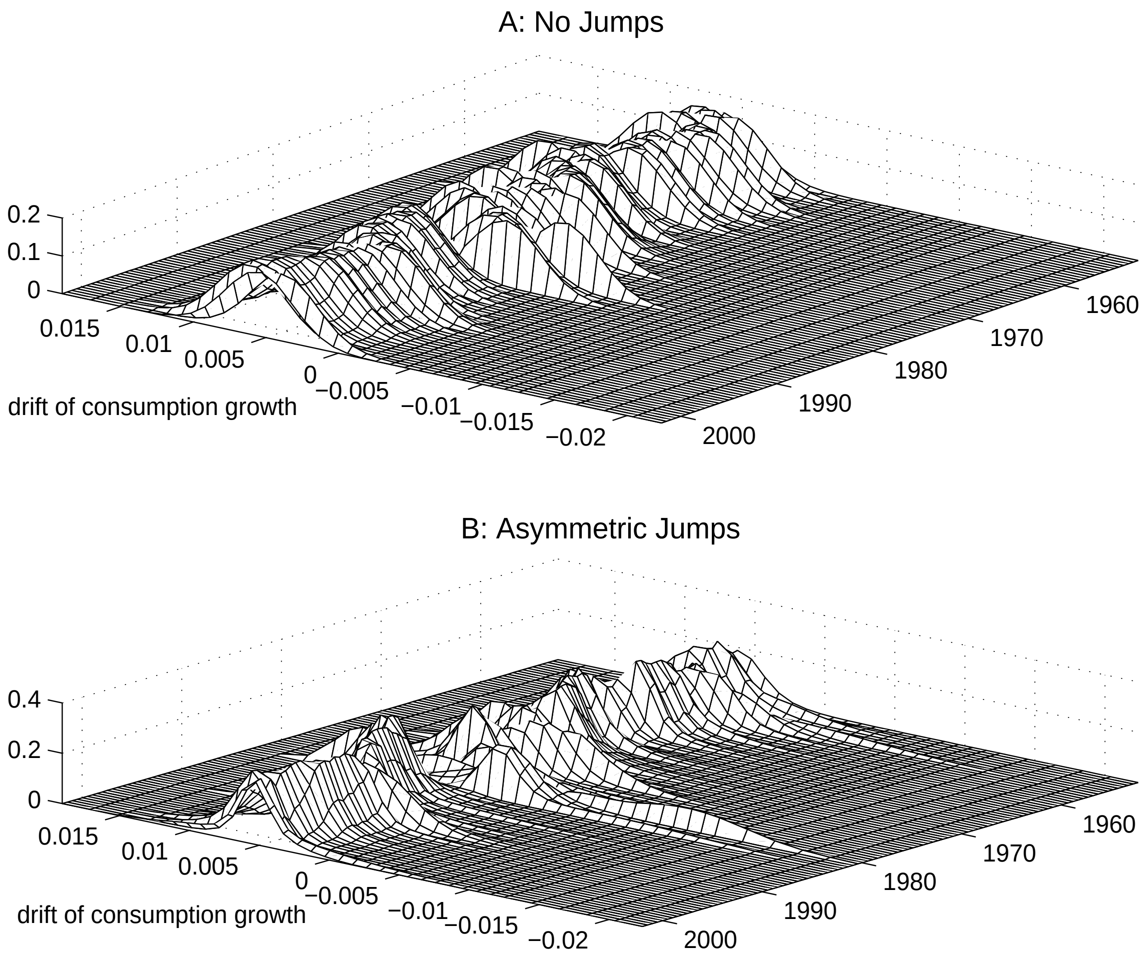
Conditional price-dividend ratios for three values of coefficient of aversion to state uncertainty ρ. The case $\rho = 0$ corresponds to the power utility case. The remaining coefficients are $\gamma = 3$ and $\phi = .01$.

Figure 2: Jump distributions



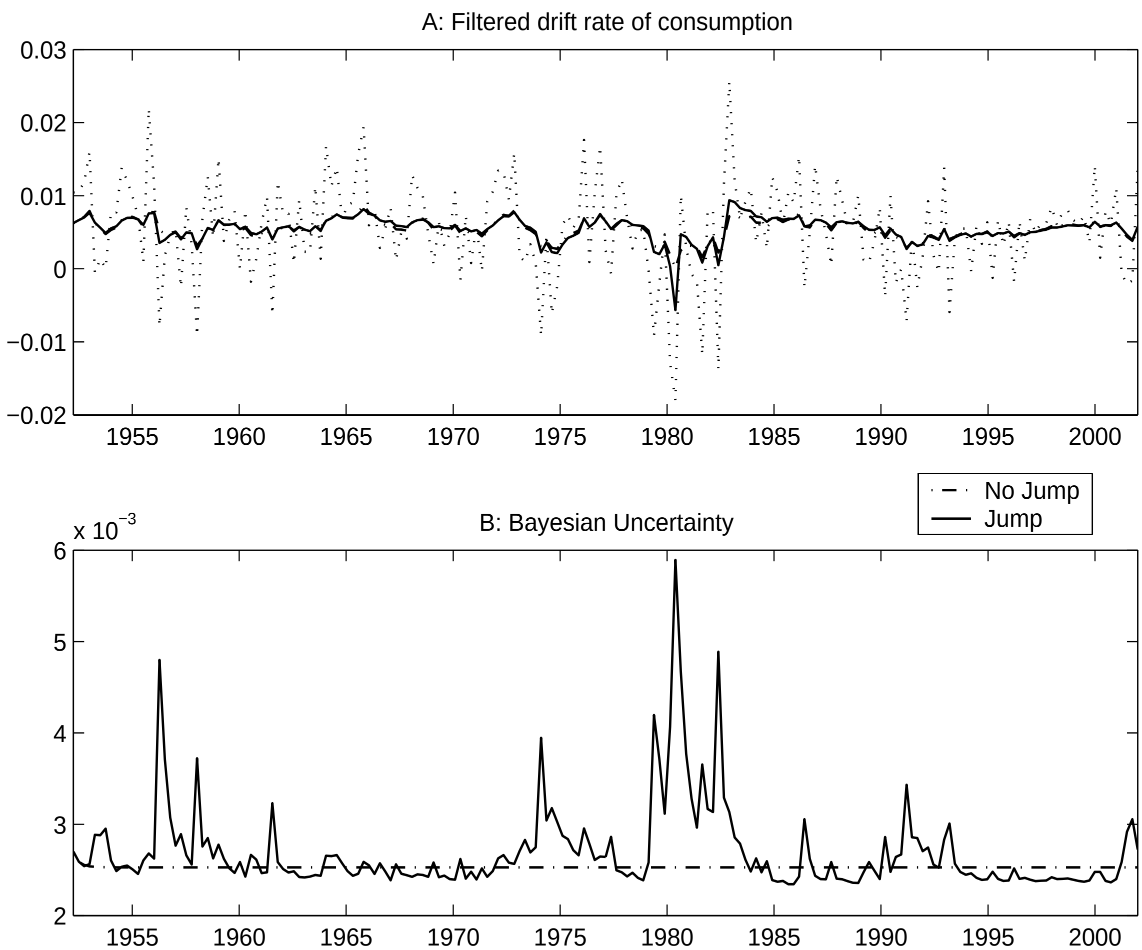
Two jump distributions estimated from consumption data: Symmetric (Normal) and Asymmetric (Weibul).

Figure 3: Posterior distributions on consumption drift (1952 - 2001)



Filtered posterior distribution $\boldsymbol{\pi}_t$ on the consumption drift $\mathcal{G} = \left(g^1, ..., g^n\right)$. The two time-series of posterior distributions in Panel A and B are obtained using the parameter estimates for the autoregressive model with no jumps (Panel A) and with asymmetric jump distribution (Panel B), via Bayes formula and the parameter estimates in Table I.

Figure 4: Expected drift and time-varying Bayesian uncertainty



Panel (A) plots the time series of real consumption growth from 1952 to 2001 and the expected drift of consumption obtained from the fitted posterior probabilities π_t as described in text and in Figure 3. Panel (B) plots the time series of the Root Mean Square Error RMSE=$\sqrt{E_t\left[g^2\right]-E_t\left[g\right]^2}$ of the expected consumption growth.