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

Pietro Veronesi

University of Chicago, NBER and CEPR

First Version: August 2000

This Version: March 15, 2001

*I thank Paolo Ghirardato, Fulvio Ortu, Nick Polson, Tano Santos, Costis Skiadas, Fernando Zapatero and seminar participants at University of Southern California and University of Minnesota for comments and suggestions. My most sincere thanks to Marciano Siniscalchi for the lengthy telephone 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.

Abstract

This paper reinterprets standard axioms in choice theory to introduce the concepts of *"belief dependent" utility functions* and *aversion to "state-uncertainty."* Within a standard pure-exchange economy in which investors ignore the long run drift of consumption growth ("the state") I 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, stock return predictability and volatility clustering. Since the long-run drift of consumption determines the (average) path of future consumption, in this context "aversion to state uncertainty" has the natural interpretation of "aversion to the dispersion of long-run consumption paths," which differs from the standard notion of (local) risk aversion in its temporal dimension. In a parsimonious parametrization, it is shown that the model calibrated to real consumption generates unconditional moments for asset returns that are in line with the empirical observation. In addition, when estimated using consumption data the fitted model produces posterior distributions on the drift rate of consumption that are relatively dispersed, which further motivates the notion of aversion to "long-run risk" put forward in this paper. Besides showing a good match of unconditional moments, the model also generates a time series of *conditional* volatility in line with the empirical observation.

1 Introduction

This paper introduces the concepts of *"belief dependent" utility functions* and *aversion to "state-uncertainty"* and shows in a standard pure-exchange economy that this type of preferences help explaining the various stylized facts of stock returns, including a high equity risk premium, a low risk-free rate, a high return volatility, stock return predictability 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 known with certainty. The recent literature in asset pricing has studied a number of "state-dependent" utility functions, where we can include the works on habit formation (see e.g. Constantinides (1990), Abel (1990), Campbell and Cochrane (1999)), relative social standing (Bakshi and Chen (1996)) and loss aversion (see e.g. Barberis, Huang and Santos (2000)). In these examples, the "state" is always known with certainty. In this paper I take the approach that the "state" is not observable by the agents but they can infer it from various signals. Hence, at any point in time, agents have a posterior distribution on the possible states. State-dependent utilities are then recovered as a special case in which agents have a degenerate posterior distribution.

More specifically, I first use standard axioms in the decision theory literature to show that "belief-dependent" utility functions can be obtained by re-interpreting the representation results about state-dependent utility function. I show that this notion of belief dependent utility functions naturally induces the concept of "*aversion to state-uncertainty*," that is, the aversion to a more diffuse distribution on the unknown state of nature. I also characterize belief-dependent utility functions with constant coefficient of absolute or relative risk aversion or with belief-dependent coefficient of risk aversion.

As an application of this approach to asset pricing, I study a Lucas (1978) exchange economy where investors are uncertain about the current drift rate of dividends (the "state"). This set-up is of particular interest because it implies a natural interpretation of the concept of aversion to state uncertainty in terms of *aversion to long-run risk.* In fact, since the drift rate of consumption determines the average path of *future* consumption, "aversion to state uncertainty" can be interpreted as "aversion to the dispersion of long-run consumption paths," which differs from the standard notion of (local) risk aversion in its temporal dimension. In other words, while the standard notion of risk aversion applies to the variability of local consumption, aversion to state-uncertainty applies to the dispersion on the whole path of future consumption, which can then be termed "long-run risk." One implication is that it is possible to separate the intertemporal elasticity of substitution from this "aversion to long-run risk" while retaining the time-separability of the utility function. This in turn makes the model highly tractable. Indeed, I obtain simple closed-form formulas for stock returns and interest rates whose characteristics can be fully interpreted.

In a parsimonious parametrization, I show that when calibrated to consumption data the model is able to explain many of the empirical features in the asset pricing literature. Specifically, *aversion to state uncertainty* (or long-run risk) has the effect of increasing the risk premium, lowering the risk free rate and increasing return volatility. In addition, accordingly with previous literature on learning (see e.g. David (1997), Veronesi (1999, 2000)), I also find time-varying expected returns, predictability and stochastic volatility. The set up is able to achieve these results even by assuming that investors have a utility function characterized by *constant relative risk aversion*. Intuitively, aversion to state-uncertainty (long run risk) generates a high equity premium and a high return volatility because it increases the sensitivity of the marginal utility of consumption to news. In addition, it also lowers the interest rate because it increases the demand for bonds from investors who are concerned about the long-run mean of their consumption. Indeed, under the interpretation put forward in the paper, it is not only the "volatility" of consumption that matters in generating risk premia, but the uncertainty on its long run drift as well. Hence, the consumption process can even be rather smooth (as in the data) and yet carry a somewhat high risk due to the uncertainty relating its long-term average dispersion (which depends on the drift). Finally, since I am not constrained in assuming a high coefficient of (local) *risk* aversion to generate a high risk premium, interest rates result low also because I can calibrate the model to reasonable values for the elasticity of intertemporal substitution.

To substantiate the discussion above, I also estimate the model using quarterly consumption data for the post-war period. Interestingly, I find that the time series of fitted posterior distributions on the drift rate of consumption tend to display a relatively large dispersion. This empirical finding further strengthens the notion that investors may be averse to this "long run risk" besides the local risk aversion stemming from the (low) volatility of consumption. Besides matching the unconditional first and second moments of returns and the level and the volatility of the interest rate, I also obtain a time series of fitted return volatility that matches well the realized volatility of stock returns in the post-war period (unfortunately, quarterly data for consumption are not available before 1946). I also show that the model implied price-dividend ratio is reasonable, although it matches the realized one to a less degree.

The paper is related to a number of recent articles: First, recent literature uses utility functions with habit formation to describe a preference over consumption in relative terms, that is, relative to other agents' aggregate consumption (see e.g. Abel (1990) and Campbell and Cochrane (1999)). In other words, for given consumption level today, the investor's utility is higher the farther away this level is from the habit level. This has the effect of making the marginal utility of consumption "state-dependent" where the state is some function of the habit level. This in turn affects asset prices by yielding for example time-varying risk aversion. Second, some of the recent literature uses the concept of "recursive" utility in order to disentangle the risk aversion from the elasticity of intertemporal substitution (see e.g. Epstein and Zin (1989), Weil (1989), Campbell (1996)). High risk aversion generates high risk premia while high elasticity of intertemporal substitution keeps the real rate low. Finally, a number of studies have concentrated on the concept of "Knightian uncertainty" to explain asset prices (see e.g. Epstein and Wang (1995), Maenheud (1999), Hansen et. al (1999), Cagetti et al (2000)).

Belief-dependent utility functions as obtained in this paper have the same intuitive motivations as those in these recent approaches, but it substantially differs from them in many respects. Compared to habit formation, for example, also belief-dependent utility functions yield a marginal utility of consumption that varies over time in response to past innovations in consumption. However the economic interpretation of the two approaches is different. Consider for example the set-up in Campbell and Cochrane (1999): in that model a positive innovation in consumption decreases the individual agent's marginal utility because now his/her consumption level is farther away from a slow moving habit. Instead, in my set-up good news in consumption reduce the current marginal utility because the investor now expects even better times for his/her own future. That is, the relevant comparison is not with other people's consumption but with the agent's own consumption in the future.

Similarly, by incorporating the overall distribution of beliefs in the utility function I effectively endow investors with a preference for resolution of their own uncertainty on the underlying true state of nature. Although this is different from preferences over early/late resolution of uncertainty as understood in the recent literature, it yields nonetheless similar implications. For example, the elasticity of intertemporal substitution in my approach is still equal to the inverse of the coefficient of relative risk aversion. However, as mentioned above, since in my set up investors are averse to "uncertainty" (or long run risk), it is still possible to obtain a high equity premium and high volatility of returns without affecting the intertemporal elasticity of substitution.

Finally, although state-uncertainty is rather different from the famous "Knightian uncertainty" a la' Gilboa and Schmeidler (1993), whereby agents are endowed with families of prior distributions on a given state and then use the max-min rule to take decisions, this paper retains the intuitive appeal that "uncertainty" is bad and that agents prefer certainty to uncertainty. In addition, the present approach does not suffer from the known problem of the "Knightian uncertainty" paradigm about the rational updating of a family of beliefs on the state of nature. In my approach, investors are endowed with only one posterior distribution which is simply updated using Bayes' rule (but see Cagetti et al. (2000) on this point).

An additional advantage of the present approach is that it is extremely manageable: Beliefs enter linearly in the instantaneous utility functions and hence intertemporal preferences preserve additivity both across time and across states (the beliefs). Since in dynamic models belief must evolve according to some linear model (and if the underlying state of nature is constant, they are actually martingales), it is quite simple to compute current expectation of future belief-dependent utilities. This allows me to obtain simple formulas for the stochastic discount factor and hence obtain simple and very interpretable formulas for asset prices and interest rates.

The article proceeds as follows: Next section introduces the concept of belief-dependent utility functions and aversion to state uncertainty. It also discusses some of the properties. Section 3 introduces the asset pricing model and section 4 obtains closed form solutions for asset prices for a class of belief-dependent utility functions. Section 5 specializes the analysis to the case of Constant Relative Risk Aversion and obtains stock returns implications. Section 6 takes the model to the data: After introducing a statistical model for dividends characterized by random jumps in their drift rate, it contains a calibration of the model and an empirical application. Section 7 concludes.

2 Belief-Dependent Utility Functions

In this first part of the paper I introduce the concepts of belief dependent utilities and aversion to state uncertainty. Specifically, I first point out that the axiomatic foundation of state-dependent utility functions naturally implies the foundation of belief-dependent utility functions upon reinterpretation of the states of Nature and the timing of resolution of uncertainty over states. A number of axiomatic approaches have been proposed and I recall one of these in Appendix A. Here, I set out the minimum notation necessary to understand the nature of the representation. The discussion is taken from Myerson (1991). Let $\mathcal{C}$ be a set of prizes and Θ a set of states. A *lottery* f is a function assigning a probability distribution on $\mathcal{C}$ to each state $\theta \in \Theta$. That is, $f : \Theta \longrightarrow \Delta(\mathcal{C})$ where $\Delta(\mathcal{C})$ is the set of probability distributions on $\mathcal{C}$. For every event $S \subseteq \Theta$, let us denote by $\succeq_S$ a *conditional preference relation* on the set of lotteries on $\mathcal{C}$. Assuming that $\succeq_S$satisfies the (standard) axioms listed in Appendix A, we then have that there exists a state-dependent utility function $u : \mathcal{C} \times \Theta \longrightarrow R$ and a subjective conditional probability function $\pi(.|S)$ on Θ such that for all lotteries f and g ,

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

Here, $\pi(.|S)$ is simply a conditional probability distribution on Θ with unit mass on the event S that satisfies Bayes law.

To better interpret the representation of preferences in (1), consider its specialization to a special types of lotteries called the "constant" lotteries, that is such that $f(c|\theta) = 1$ for every

$\theta \in \Theta$. If we denote such a lottery by $[c]$, then the application of the representation result (1) implies that

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

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.

A couple of simple examples will help to clarify this point:

Examples: (i) Consider an agent in a large class of students who just took an exam and got a B. If the "utility" from the B changes depending on the overall average obtained by the (large) class, then the student utility is "state-dependent" (where the state can be considered the "average grade"). However, suppose that the agents ignores what is the exact average class grade but only holds a probability distribution on it. If shifts in this probability distribution changes the "utility" from the grade, now the utility is "belief-dependent" because it depends on the (subjective) belief on the average grade. (ii) Example (i) considered the state to be "external." The same example works on an "internal" basis, that is, when the student judges his/her own performance in an exam compared to his/her previous performance in other exams. Suppose that the student took a few exams in the past weeks but only one grade is already available. If the utility he/she derives from a grade of B depends on his/her beliefs on the average grades on all the other exams, the agent holds a belief-dependent utility function. In other words, state-dependent utility functions are the special case of belief-dependent utilities when the state is known with certainty.

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

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

2.1 Aversion to State-Uncertainty

Characterization (3) naturally leads to a definition of aversion to "state-uncertainty." To emphasize it, let me recall again the implication of this utility representation: Given a constant act $[c]$ yielding a known prize c and a set of states $\theta \in \Theta$, the effective utility for the decision maker is given by a weighted average of the state-dependent utilities, weighted by the agent's beliefs on the state. Since the agent knows "$u(c|\theta)$" for every prize c and every state θ, although (3) holds formally only ex ante, it must hold also ex-post, after observing c.

Given 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. I use this concept to define aversion to state uncertainty:

Definition: *(a)* Let π and c be given. A belief-dependent utility function $U(c, \pi)$ displays *aversion to state-uncertainty given* 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 function displays *aversion to state uncertainty* if (4) holds for all c.

In the context of example *(i)* in the previous subsection, if the student is happy to learn that the average grade of the class was exactly B when he/she was assigning equal probabilities to be either C+, B and A-, then he/she is averse to state-uncertainty. I will give other examples below in the context of a dynamic economy.

2.2 Absolute and Relative Risk Aversion of Belief Dependent Utility Functions

In this section I provide the characterization of belief dependent utilities in the case of constant absolute and relative risk aversion. I introduce some notation first:

Absolute Risk Aversion: The absolute risk aversion of belief dependent utility function $U(c,\pi)$ is given by

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

Relative Risk Aversion: The relative risk aversion of belief dependent utility function $U(x,\pi)$ is given by

$$\gamma(\pi,c) = -\frac{c\partial^2 U(c,\pi)/\partial c^2}{\partial U(c,\pi)/\partial c} \tag{6}$$

These are the analogous notions of absolute and 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), (5) and (6) reflect the *local* curvatures of the utility function that are necessary and sufficient to generate "aversions" to fair bets (either in absolute or in relative terms).

Given their importance in finance applications, I now characterize the belief dependent utility function for the case of constant absolute or relative risk aversion:

Proposition 1: *(a)* $A(c,\pi) = A$ constant if and only if

$$u(c,\pi) = E_t[k_1(\theta)] - E_t[k_2(\theta)]\, e^{-Ac} \tag{7}$$

where $k_2(\theta) > 0$ for all $\theta \in \Theta$ and $E_t[k_i(\theta)] = \sum_{j=1}^n \pi_j k_i(\theta_j)$.

(b) $\gamma(c,\pi) = \gamma$ constant if and only if

$$u(c,\pi) = E_t[k_1(\theta)] + E_t[k_2(\theta)]\, \frac{c^{1-\gamma}}{1-\gamma} \tag{8}$$

with $k_2(\theta) > 0$ for all $\theta \in \Theta$.

Proof: See Appendix. ■

Albeit easy to prove, the implications of this proposition are rather interesting. Consider for instance the case with constant relative risk aversion (8). Although in a static model this representation is basically equivalent to one with no state-dependent utility because one can define $\alpha = E[k_1(\theta)]$ and $\beta = E[k_2(\theta)]$ and proceed as usual, in a dynamic context where agents learn about the true state θ over time, this representation becomes important.

For example, suppose that $k_2(\theta)$ is a convex function of $\theta \in \Theta$. Then, an immediate consequence of the above proposition is that a mean preserving spread on the posterior distribution over θ (that is, an increase in uncertainty for given $E[\theta]$) implies an increase of $E[k_2(\theta)]$ and hence it decreases the "utility" for $\gamma > 1$ (because $c^{1-\gamma}/(1-\gamma) < 0$) and it increases the marginal utility of consumption

$$\frac{\partial U(c,\pi)}{\partial c} = E[k_2(\theta)]\, c^{-\gamma} \tag{9}$$

We will see that his has interesting implications for asset prices, for example.

2.3 Belief-Dependent Relative Risk Aversion

Even though the agent does not have full information about the underlying state of nature, according to Myerson's axiomatic approach he/she may nonetheless establish the level of his/her relative risk aversion conditional on each state being true. That is

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

In this case, the representation of utility function is as follows:

Proposition 2: The belief-dependent utility function with belief-dependent risk aversion is given by

$$U(c,\pi) = E_t[k_1(\theta)] + \sum_{i=1}^{n} \pi_i k_2(\theta_i)\, \frac{c^{1-\gamma(\theta_i)}}{1-\gamma(\theta_i)} \tag{10}$$

Proof: See Appendix. ■

I shall use this representation of belief-dependent utilities to state the main asset pricing result of this paper, contained in section 4. However, I will then restrict the analysis of the properties of stock returns and interest rates as well as the empirical work to the case where investors have constant relative risk aversion, that is, $\gamma(\theta) = \gamma$ for all $\gamma \in \Theta$. This will allow me to depart from standard state-independent preferences in only one dimension and hence it will make it clearer the additional effects that one obtains when "aversion to state-uncertainty" is introduced.

3 A Pure Exchange Economy

Let W be a Wiener process defined on a complete probability space $\left(\Omega, \mathcal{P}^0, \mathcal{F}^0\right)$. I make the following assumptions about the economy

Assumption 1: *Real* dividends evolve according to the stochastic differential equation

$$\frac{dD}{D} = \theta dt + \sigma_D dW \tag{11}$$

where $\theta(t)$ is unobservable and its dynamics is described below and σ_D is a 1×2 constant vector.

Assumption 2: $\theta(t)$ follows a Markovian process defined on a finite set $\Theta = \{\theta_i\}_{i=1}^n$. We assume that for every θ_i,$\theta_j \in \Theta$ there exists λ_{ij} such that in the infinitesimal interval Δ we have:[1]

$$\Pr\left(\theta(t+\Delta) = \theta_j | \theta(t) = \theta_i\right) = \lambda_{ij}\Delta$$

The infinitesimal generator Λ is such that $[\Lambda]_{ij} = \lambda_{ij}$ for $i \neq j$ and $[\Lambda]_{ii} = -\sum_{j \neq i} \lambda_{ij}$.

Assumption 2 makes the approach very convenient and allow us to obtain closed form solutions for asset prices under rather general conditions on the process $\theta(t)$. Although it assumes that $\theta(t)$ can only take a finite number of values, the assumption leaves unspecified the *number* of states. Hence, we can effectively approximate any continuous-time, continuous-state stationary Markov process by choosing a sufficiently fine grid $\Theta = [\theta_1, \theta_2, ..., \theta_n]$ on the real line and by carefully choosing the transition probabilities λ_{ij}. Section 6 contains one such example, where $\theta(t)$ is assumed to follow a pure jump process with state-dependent jump intensity. Approximations to other mean reverting processes such as Ornstein-Uhlenbeck processes can also be implemented, by choosing appropriate tri-diagonal matrices.[2]

[1] In fact, Lemma 1 below also holds when Θ is countable and not only finite. Since in this case the set Θ could be dense in the real line, it can approximate well a continuous-state model. I keep the set finite for convenience in the asset pricing part of the paper, but I am confident that those results too readily extend to the case where Θ is countable. I nonetheless leave the generalization to future work.

[2] Indeed, it is known that even under a standard state-independent iso-elastic utility function it is not possible to solve for prices in closed form when the underlying drift $\theta(t)$ follows a simple Ornstein-Uhlembeck process (see e.g. Yan (1999) and Brennan and Xia (2001)). Numerical solutions must always be implemented. The

Let investors' subjective probability that the state is θ_i at time t given their past observation of $(D(\tau))_{\tau=1}^{t}$ be denoted by

$$\pi_i(t) = \Pr(\theta(t) = \theta_i | \mathcal{F}(t))$$

Assuming that investors use Bayes law to update their beliefs, we have the following result:

Lemma 1: The probabilities $\pi_i(t)$ evolve according to the system of n stochastic differential equations

$$d\pi_i = [\pi\Lambda]_i \, dt + \pi_i \sigma_i(\pi) \, d\widetilde{W} \tag{12}$$

where

$$\sigma_i(\pi) = \left(\theta_i - \sum_{j=1}^{n} \pi_j \theta_j\right) \frac{1}{\sigma_D} \text{ and } d\widetilde{W} = \frac{1}{\sigma_D}\left(\frac{dD}{D} - E_t\left[\frac{dD}{D}\right]\right) \tag{13}$$

$d\widetilde{W}$ is a Wiener process defined on the probability space $(\Omega, \mathcal{P}, \mathcal{F})$ where $\mathcal{F}$ is the filtration generated by dD/D.

Proof: See Liptser and Shyriaev (1974). ■

I can also rewrite the original dividend process under the new filtration

$$\frac{dD}{D} = \mu_D(\pi) \, dt + \sigma_D d\widetilde{W} \tag{14}$$

where $\mu_D(\pi) = \sum_{i=1}^{n} \pi_i \theta_i$. Since the probabilities $\pi(t)$ are "known" to the investor at time t and the information filtration generated by $\{D(t)\}$ is equivalent to the one generated by $\left\{\widetilde{W}(t)\right\}$, one can take the processes (12) and (14) as primitive of the analysis. However, the interpretation of the state-variables $\pi(t) = (\pi_1(t), ..., \pi_n(t))$ as the time t posterior distribution on the underlying unknown drift $\theta(t)$ is key to understand the next assumption about investors' preferences.

Assumption 3: As in section 2, the representative investor utility function is belief-dependent and it has the following form[3]

$$U(c, t, \pi) = \sum_{j=1}^{n} \pi_j(t) \, u(c, t | \theta_j) \tag{15}$$

approach taken here offers a viable alternative, where the approximation is on the state-space and simple closed form solutions can then be obtained.

[3]We could use the more general form as in (10) but it turns out it has no consequences on the price of assets. In order to limit the amount of notation, we set $k_1(\theta) = 0$.

where

$$u(c,t|\theta)=e^{-\phi t}k(\theta)\,\frac{c^{1-\gamma(\theta)}}{1-\gamma(\theta)} \tag{16}$$

This form nests the two possibilities of constant relative risk aversion (just set $\gamma(\theta)=\gamma$ for all $\theta\in\Theta$) and belief dependent relative risk aversion. In addition, setting $k(\theta)=1$ for all θ yields the standard iso-elastic utility function. This assumption also further clarifies the concept of "belief-dependent utility function" if we notice that investors know the level of consumption today $c(t)$ but they do not know under what "state" (θ) was $c(t)$ generated. In this case, the heuristic idea of the model is that if the same level of consumption $c(t)$ is generated under different "regimes" (say, recessions or booms) then the investors will have different levels of (cardinal) utility. However, investors do not know whether the economy is in a boom or a recession itself and so his/her utility is belief-dependent.

3.1 Long-Term Risk Aversion and Habit Formation

One question that arises at this point is why should the utility function of an investor depend on the (unknown) drift rate of dividends. For example, if the investor does not invest in the stock, why should his/her utility be affected by his belief on the drift rate of dividends? To better understand the set up proposed here, consider first the maximization problem of an investor at time t :

$$\max_{\{x(\tau),c(\tau)\}} E_t\left[\int_t^{\infty} U(c(s),\pi(s),s)\,ds\right] \tag{17}$$

subject to the dynamic budget constraint

$$dZ=Z\left(x\frac{dP+Ddt}{P}+(1-x)\,rdt\right)-cdt \tag{18}$$

where $P(t)$ denotes the price of the asset and $x(t)$ the fraction of wealth invested in stock.

Notice that problem (17) subject to (18) given the dividend process (14) and the "state-variables" (12) defined with respect to the filtered Brownian motion $\widetilde{W}$ obtained in (13) is a rather standard problem. Inded, since markets are complete, it is known that optimal consumption is given by

$$c(t)=\mathcal{I}_U(\zeta\chi(t),\pi(t),t)$$

where $\mathcal{I}_U(., \pi(t), t)$ is the inverse of the marginal utility function $U_c(., \pi, t)$, ζ is a constant and $\chi(t)$ is the state-price density defined by the interest rate and the return process in the usual way (see e.g. Karatzas and Shreve (1998), Ch. 3.9). Hence, one can conjecture that the optimal consumption will be an Ito's process of the form

$$\frac{dc}{c} = \mu_c(t)\,dt + \sigma_c(t)\,d\widetilde{W}$$

where $\mu_c(t)$ and $\sigma_c(t)$ are some ($\mathcal{F}$-adapted) processes. At this point, we can use the definition of $\widetilde{W}$ in (13), that is $d\widetilde{W} = \sigma_D^{-1}(dD/D - E[dD/D]) = \sigma_D^{-1}(\theta - \mu_D(\pi))\,dt + dW$, and substitute it back in the optimal consumption process to obtain

$$\frac{dc}{c} = (\alpha(t)\cdot\theta + \beta(t))\,dt + \sigma_c(t)\,dW \tag{19}$$

where $\alpha(t) = \sigma_c(t)\,\sigma_D^{-1}$ and $\beta(t) = \mu_c(t) - \mu_D(\pi)\,\sigma_c(t)\,\sigma_D^{-1}$. Equation (19) shows that the drift rate of optimal consumption depends on the unobservable state variable θ, over which the investor has only a dispersed distribution $\pi(t) = (\pi_1(t), ..., \pi_n(t))$.[4] The notion of state uncertainty now becomes more apparent: A rational agent who maximizes utility over future consumption paths must recognize that the optimal drift rate of his/her consumption is determined by the value of the (unobserved) drift of dividends (which affect the sequence of innovations $d\widetilde{W}$). It is for this reason that his/her utility may be belief-dependent: Changes on the current belief on the drift rate of the (optimal) consumption process may change the utility from consuming a given amount of consumption good.

To further the intuition it could be helpful to strike a parallel with habit formation models. Consider the case where $\gamma(\theta) = \gamma$. Then, the instantaneous marginal utility of consumption

[4]Indeed, in equilibrium we must have that the optimal consumption follows the process

$$\frac{dc}{c} = \sum_{i=1}^{n} \pi_i\theta_i + \sigma_D d\widetilde{W}$$

which can then be rewritten as

$$\frac{dc}{c} = \theta + \sigma_D dW$$

under the original Brownian motion W. Hence, a rational (representative) investors who can expect that $c = D$ (because the good is perisheable) would also expect that the drift rate of his/her consumption be fully determined by θ alone.

is given by

$$U_c(c,t,\pi) = e^{-\phi t} E_t[k(\theta)] c(t)^{-\gamma} \tag{20}$$

That is, for given consumption level $c(t)$ the marginal utility of consumption is time varying where the time variation depends on the expectation $E_t[k(\theta)] = \sum_{i=1}^{n} \pi_i(t) k(\theta_i)$, which in turn depends on past realizations of consumption (= dividends). This is a common feature of external habit formation models: For example, the utility function postulated by Campbell and Cochrane (1999) leads to a marginal utility function given by $U_c(c_t, s_t) = c_t^{-\gamma} s_t^{-\gamma}$ where $s_t = \log((c_t - X_t)/c_t)$ (the surplus) follows a slow mean reverting process that depends on past realizations of consumption (here X_t denotes the "habit" level). In their model a positive innovation in current consumption c_t leads to an *increase* in s_t because consumption in now further away from the habit, and hence a reduction in the marginal utility of consumption. Since in the model outlined in Assumptions 1-2 a positive innovation in consumption increases its expected drift rate $\mu_C(\pi) = E_t[\theta]$, by postulating that $k(\theta)$ is decreasing in θ we obtain that a positive innovation in consumption decreases $E_t[k(\theta)]$. Hence, the model yields the same intuitive implication as Campbell and Cochrane (1999) and, in general, external habit formation models (see also Abel (1990)). However, the interpretation is different. While in habit formation models good news in consumption decrease the marginal utility because they increase the distance of current consumption from a slow moving "habit," in my set-up positive innovations in consumption reduce the marginal utility because the investor now expect even better times for his/her future.

In addition, given assumption 3 with $\gamma(\theta) = \gamma > 1$, we also have

$$U(c,t,\pi) = \sum_{j=1}^{n} \pi_j(t) k(\theta_j) e^{-\phi t} \frac{c^{1-\gamma}}{1-\gamma}$$

Hence, if $k(\theta)$ is also assumed convex, then a mean preserving spread on the distribution $\pi(t)$ decreases the belief dependent utility $U(c,t,\pi)$ (recall that $c^{1-\gamma}/(1-\gamma) < 0$) thereby yielding the aversion to state uncertainty characteristics. From (20) aversion to state-uncertainty also leads to an increase in the marginal utility of consumption: in other words, during periods of higher uncertainty about the drift rate of future dividends investors value of one unit of current consumption more. Since I argued above that uncertainty on θ implies uncertainty on the drift rate of the investor's optimal consumption path, one can term the aversion to state uncertainty more intuitively as "*aversion to long-term risk*," that is, to the dispersion of the *average* path of optimal consumption.

To summarize, I make the following assumption:

Assumption 4: The function $k(\theta)$ is positive, monotonically decreasing and convex and such $\sum_{i=1}^{n} \overline{\pi}_i k(\theta_i) = 1$ where $\overline{\pi}$ is the unconditional distribution on Θ implied by the matrix Λ. If it does not exist (e.g. $\Lambda = 0$) then $\overline{\pi} = \pi(0)$, investors' prior at time $t = 0$.

Before turning to the asset pricing implications, it is important to remark that this type of utility function takes the posterior probabilities as given: The signals that investors receive are outside their own control. In particular, the choice of the consumption plans does not affect per se the evolution of the investors' posterior distribution on the state of the economy.

4 Equilibrium Asset Prices

In this section I characterize asset prices under a rational expectation equilibrium defined as follows:

Definition: A *Rational Expectations Equilibrium* is given by a set of prices $P(t)$, consumption $c(t)$, asset allocation $x(t)$ such that investors solve the intertemporal maximization problem (17) subject to (18) and the market clearing conditions are satisfied, i.e. $D(t) = c(t)$ and $x(t) = 1$.

For comparison with later results, it is useful to first find the value of stocks under the stronger assumption of belief independent utility functions.

Proposition 3: Let the utility be belief-independent, that is, $U(c,t,\pi) = U(c,t) = e^{-\phi t}\dfrac{c^{1-\gamma}}{1-\gamma}$. Then: *(a)* The stock price is given by

$$P_B(t) = D(t)\left(\sum_{i=1}^{n} \pi_i(t)\,\overline{B}_i\right) \tag{21}$$

with

$$\overline{B}_i = \sum_{j=1}^{n} e_j \left(\phi\mathbf{I} - \overline{\Lambda}'\right)^{-1} e_i' \tag{22}$$

where $\overline{\Lambda} = \Lambda + (1-\gamma)\, diag\,(\theta_1, ..., \theta_n) - \frac{1}{2}\gamma\,(1-\gamma)\,\sigma_c^2$ and e_i is the i-th row of the identity matrix.

(b) The real rate of interest is

$$r_B\,(t) = \phi + \gamma \sum_{i=1}^{n} \pi_i\,(t)\,\theta_i - \frac{1}{2}\gamma\,(1+\gamma)\,\sigma_c^2 \tag{23}$$

Proof: Special case of Proposition 4. See also Veronesi (2000). ■

A few comments are in order: First, for $\gamma > 1$ we typically have that $\overline{B}_k$ are decreasing with k, i.e. a higher growth rate of the economy is associated with a lower price-dividend ratio.[5] This effect stems from the interplay of income and substitution effect for the power utility case. A low growth rate of dividends implies that future consumption is lower than today and hence the desire to smooth consumption make investor increase their savings and hence demand for assets. This pushes the real rate down and the price of stocks high.

To state my main result, I need a last piece of notation that will be used in the propositions below. For all $i = 1, ..., n$ let

$$\overline{\Lambda}_i = \Lambda + diag\left(\widehat{\theta}_1^i, ..., \widehat{\theta}_n^i\right) \tag{24}$$

with

$$\widehat{\theta}_j^i = (1-\gamma_i)\,\theta_j - \frac{1}{2}\,(1-\gamma_i)\,\gamma_i \sigma_c^2 \tag{25}$$

and where $\gamma_i = \gamma\,(\theta_i)$. The following is my main result.[6]

Proposition 4: For each $k = 1, ..., n$ and for given $\pi\,(t)\,, c\,(t)$ let

$$\pi_i^*\,(c) = \frac{k_i \pi_i\,(t)\,c\,(t)^{-\gamma_i}}{\sum_{j=1}^{n} k_j \pi_j\,(t)\,c\,(t)^{-\gamma_j}} \tag{26}$$

$$B_i = \frac{1}{k_i}\sum_{j=1}^{n} k_j e_j \left(\phi \mathbf{I} - \overline{\Lambda}_j'\right)^{-1} e_i' \tag{27}$$

[5] 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.

[6] Along the lines of the proof of Proposition 4, it is equally possible to obtain closed form solutions for *bond prices* of any maturity. In the interest of space, I omit the formulas in this case and leave the investigation of bond returns and the term-structure of interest rates under belief-dependent utilities to future work.

with $k_i = k(\theta_i)$. Then:

(a) the price of the asset is

$$P(t) = D(t) \sum_{i=1}^{n} \pi_i^* (c) B_i \tag{28}$$

(b) The real rate of interest is

$$r(t) = \phi + \sum_{i=1}^{n} \pi_i^*(c) \gamma_i \theta_i - \frac{1}{2} \sum_{i=1}^{n} \pi_i^*(c) \gamma_i (\gamma_i + 1) \sigma_c^2 - \sum_{j=1}^{n} \pi_j^*(c) C_j^*(c) \tag{29}$$

where

$$C_j^*(c) = \sum_{i=1}^{n} \lambda_{ji} \frac{k_i}{k_j} c^{\gamma_j - \gamma_i}$$

Proof: See Appendix ■

These asset pricing formulas have a number of properties that I discuss in the next few pages. In the sequel, I will refer to the constants B_i's appearing in (27) as *conditional price-dividend ratios*, because each of them is the price-dividend ratio that would occur at time t if there was perfect certainty on the underlying state. Indeed, from (26) we see immediately that if $\pi_i(t) = 1$ for some i, then $\pi_i^*(c) = 1$ and hence $P(t)/D(t) = B_i$. In addition, a "*" on an expectation operator $E_t^*(.)$ implies that the expectation is computed using the distribution (26) rather than $\pi(t)$.

I point out immediately two properties stemming from proposition 4: First of all, suppose that the state $\theta(t)$ was constant and known equal to θ_k (i.e. $\pi_k = 1$ and $\lambda_{kj} = 0$ for all j), then the formulas (28) and (29) reduce to

$$\begin{aligned} P(t) &= \frac{D(t)}{\phi - (1 - \gamma_k)\theta_k + \frac{1}{2}(1 - \gamma_k)\gamma_k \sigma_c^2} \\ &= \frac{D(t)}{r + \gamma \sigma_c^2 - \theta_k} \\ r &= \phi + \gamma_k \theta_k - \frac{1}{2} \gamma_k (1 + \gamma_k) \sigma_c^2 \end{aligned}$$

Hence, perfect certainty reduces the model to the usual Lucas (1978) model, where the risk premium is simply given by $\gamma \sigma_c^2$. This emphasizes that the main implications of the belief-dependent utility function specification stem from the very uncertainty of investors on the true drift rate of dividends.

Second, if we let $k(\theta) = k = 1$ and $\gamma_i = \gamma_j = \gamma$ we obtain $\pi_i^* = \pi_i$ and hence we obtain back the price P_B and the interest rate r_B as in Proposition 3. As already commented, all the implications of this asset pricing formula have been detailed in Veronesi (2000).

I now turn to the implications of Proposition 4 in the two cases of constant and belief-dependent coefficient of relative risk aversion.

5 Asset Prices and Returns under CRRA

In this paper I only concentrate on the case where the relative risk aversion is constant. As we will see below, this gives already many interesting results that are then easier to interpret if I depart from the standard state-independent utility framework only in one dimension. I leave the study of returns under belief-dependent risk aversion to future work.

5.1 Asset Prices

In the case of constant relative risk aversion, we have $\gamma_i = \gamma_j = \gamma$ and hence we obtain the following corollary to Proposition 4.

Corollary 1: Let $\gamma_i = \gamma_j = \gamma$ so that now

$$\pi_i^*(t) = \frac{k_i \pi_i(t)}{\sum_{j=1}^n k_j \pi_j(t)} \tag{30}$$

$$B_i = \frac{1}{k_i} \sum_{j=1}^n k_j e_j \left(\phi \mathbf{I} - \overline{\Lambda}'\right)^{-1} e_i' \tag{31}$$

Then: (a) The stock price is given by:

$$P(t) = D(t) \sum_{i=1}^n \pi_i^*(t) B_i \tag{32}$$

(b) The interest rate is given by

$$r(t) = \phi + \gamma \sum_{i=1}^n \pi_i^*(t)\,\theta_i - \frac{1}{2}\gamma(\gamma+1)\sigma_c^2 - \sum_{j=1}^n \pi_j^*(t)\, C_j^* \tag{33}$$

$$= r_B(t) + \gamma V_U - \sum_{j=1}^n \pi_j^*(t)\, C_j^* \tag{34}$$

where $r_B(t)$ is given in (23) and $C_j^* = \sum_{i=1}^{n} \frac{k_i}{k_j} \lambda_{ji}$ and

$$V_U = \frac{\sum_{i=1}^{n} k_i \pi_i \left(\theta_i - \sum_{j=1}^{n} \pi_j \theta_j \right)}{\sum_{i=1}^{n} \pi_i k_i} \quad (35)$$

$$= E_t^* [\theta] - E_t [\theta] \quad (36)$$

The effect of a belief-dependent utility function shows itself in two terms in the pricing function (32) compared to the benchmark case (21): First, it affects the "conditional price-dividend ratios" B_i as it can be seen by comparing (31) with (22). Intuitively, from (20) the current marginal utility of consumption is now belief-dependent. Hence, even when we condition on a particular state (i.e. we set $\pi_i = 1$ 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. Second, when $\pi(t)$ is a non-degenerate distribution, the conditional price-dividend ratios B_i are weighted by the probability distribution $\pi_i^*(t)$ rather than the original $\pi(t)$. Intuitively, the probabilities $\pi_i^*(t)$ 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 $\pi_i^*(t)$ becomes relatively bigger than $\pi_j^*(t)$: Stock prices now reflect more those states characterized by higher marginal utility of consumption.

Turning to part (b) of the corollary, we find that also the interest rate is affected by the state-dependent utility function formulation. Specifically, the term

$$V_U = E_t^* [\theta] - E_t [\theta]$$

directly enters into the formula, where a "*" denotes an expectation taken with respect to the distribution $\pi^*(t)$. This term will play an important role also in the expected return formula for stocks. Intuitively, V_U measures two important components of the belief-dependent utility function formulation: First, the behavior of the function $k(\theta)$ (i.e. the degree by which various states are pulled apart) and second the extent of the "uncertainty" contained in the posterior distribution $\pi(t)$. In fact, we can readily see that $V_U = 0$ if either $k(\theta) = k$ for all θ or if $\pi_i(t) = 1$ for some i. In order to characterize V_U further, I need to use the assumption that $k(\theta)$ is decreasing and convex, as in assumption 4. We then have the following:

Lemma 2: Let Assumption 4 hold.

(a) If $\pi(t)$ is a non-degenerate distribution, then $V_U < 0$;

(b) Let $\overline{k}(\theta) = k(\theta) \times (\theta - \theta_1)$. If $\overline{k}''(\theta) < 0$, then a mean preserving spread on the distribution $\pi(t)$ decreases V_U.

Proof: See Appendix. ■

We then have the following implications:

Corollary 3: Let Assumption 4 hold.

(a) For non-degenerate probability distributions, the real rate of interest is always lower than the benchmark interest rate. That is, $V_U < 0$ and $\sum_{j=1}^{n} \pi_j^*(t)\, C_j^* > 0$.

(b) Let $\overline{k}(\theta) = k(\theta)(\theta - \theta_1)$. If $\overline{k}''(\theta) < 0$ and there are no regime shifts ($\lambda_{ji} = 0$ for all i, j), then higher uncertainty decreases the real rate of interest. That is, if $\widetilde{\pi}$ is a mean-preserving spread of π, then $\widetilde{r}(t) < r(t)$.

Proof: Immediate from Lemma 3. ■

.

Part (a) of the corollary shows that belief-dependent utility generates lower interest rates compared to the usual Lucas economy implication. This effect goes in the direction of weakening the risk-free rate puzzle (the interest rate being too high for high levels of risk aversion). The reason why higher uncertainty generates lower interest rates is intuitive: With a belief-dependent utility function higher uncertainty increases the volatility of future marginal utility of consumption, thereby increasing the demand for bonds to at least reduce the variability of future consumption. The second term $\sum_{j=1}^{n} \pi_j^*(t)\, C_j^*$ also goes in the proper direction to lower interest rates, and the effect stems now from the term $C_j^* = \sum_{i=1}^{n} \frac{k_i}{k_j} \lambda_{ji}$. Since λ_{ji} gives the probability of moving from state θ_j to a different state θ_i characterized by the marginal utility weight $k(\theta_i)$, this forecasted probability to move to different marginal utility level make again the investor increase his/her demand for bonds to decrease the fluctuations in consumption. Indeed, if we set $k(\theta) = k$ for all θ, then we have $C_j^* = 0$ for all j, eliminating this effect.

Part (b) of Corollary 3 shows that if $k(\theta)$ has sufficient curvature, then increases in uncertainty lower the real rate of interest when the underlying unobservable drift $\theta(t)$ is in fact constant. The latter assumption is very loose, and the same can happen when $\theta(t)$ moves over time as in Assumption 2 but whether this occurs or not should be studied on a case-by-case basis.

5.2 Stock Returns

Define first the following quantity, that will be important below:

$$V_B = \frac{\sum_{i=1}^{n} k_i B_i \pi_i \left(\theta_i - \sum_{j=1}^{n} \pi_j \theta_j\right)}{\sum_{i=1}^{n} \pi_i k_i B_i}$$

The subscript "B" in V_B is mnemonic for "benchmark": This quantity was introduced in Veronesi (2000) which is the benchmark case with state independent utility, with the only important caveat that in that case $k_i = k = 1$, a constant. Once again, as V_U introduced in (35) in the previous section, the quantity V_B is sensitive to the dispersion of beliefs $\pi(t)$, being equal to zero when $\pi_i(t) = 1$ for some i. Additionally, it is sensitive to the dispersion of the conditional price-dividend ratios B_i weighted now by k_i. For example, if $k(\theta) = k = 1$ and $\gamma = 1$ (i.e. log utility), then $B_i = B_j$ and $V_B = 0$. For later reference, notice that as in Lemma 3, if $k_i B_i$ is decreasing in i, then $V_B < 0$. This is typically the case in the benchmark case when $\gamma > 1$ and $k(\theta) = k$, as I already mentioned in the previous section (see Proposition 3). In the case of belief-dependent utility, whether $V_B > 0$ or $V_B < 0$ depends on the parametrization of the model.

Let us denote the excess return as

$$dR = \frac{dP + Ddt}{P} - rdt$$

I then obtain the following result:

Proposition 5: (a) The excess return dR follows the process

$$dR = \mu_R dt + \sigma_R d\widetilde{W}$$

with

$$\mu_R = \gamma\left(\sigma_D^2 + V_B - V_U\right) - V_U\left(1 + h_D^2\left(V_B - V_U\right)\right) \tag{37}$$

$$\sigma_R = \sigma_D + h_D (V_B - V_U) \tag{38}$$

where $h_D = 1/\sigma_D$.

(b) The risk-free rate evolves according to the process

$$dr = \mu_r (\pi) dt + \sigma_r (\pi) d\widetilde{W} \tag{39}$$

where μ_r is given in the appendix and

$$\sigma_r (\pi) = r h_D (V_r + V_U) - \left(\phi + \frac{1}{2}\gamma (\gamma + 1) \sigma_c^2 \right) h_D V_r \tag{40}$$

where

$$V_r = \frac{\sum_i C_i^r \pi_i \left(\theta_i - \sum_j \pi_j \theta_j \right)}{\sum_{i=1}^n \pi_i (t) C_i^r}$$

and $C_i^r = \gamma k_i \theta_i - C_i^*$.

Proof: See Appendix. ■

From the above formulas, it is clear that returns are characterize mainly by the quantity $V_B - V_U$. Before interpreting the results, recall that under Assumption 4, Lemma 2 shows that $V_U < 0$. However, the same assumption may imply that $k_i B_i$ is also decreasing, which would lead to $V_B < 0$. In this case it is hard to gauge whether $(V_B - V_U)$ is positive or negative. However, the next result shows that if we have that the conditional price-dividend ratio B_i is increasing in θ_i, then we can sign unambiguously $(V_B - V_U)$. As we shall see in Section 6, an increasing conditional price-dividend ratio is to be expected for most parameter values.

Lemma 3: Let $k(\theta)$ be decreasing in θ and B_i increasing in θ_i. Then $V_B - V_U > 0$.

Proof. See the appendix. ■

Equation (37) shows that the expected excess returns with belief-dependent utility are affected by the two measures V_U and V_B, which for given posterior distribution $\pi(t) = (\pi_1(t), ..., \pi_n(t))$, measure the extent of aversion to state uncertainty (V_U) and its effect on the conditional price-dividend ratios (V_B). In case of perfect certainty, then both $V_U = V_B = 0$ and we have the usual result that the instantaneous expected returns are equal to $\gamma\sigma_D^2$. Uncertainty with a high degree of uncertainty aversion makes $V_U < 0$ and $V_B - V_U > 0$, which increase the expected return substantially. From (38), the same effect increases the volatility of stock returns. Recall from the previous section (equation (34)) that the same assumption entails a lower interest rate with respect to the standard case. Finally, from (40) we see that the volatility of the interest rate is affected by state-uncertainty. The quantity V_r has the same form as the other V_U and V_B. However, it is difficult to say whether the volatility of interest rates decreases or not with uncertainty aversion, because the form of C_i^* is rather complex.

Rather than commenting further these formulas, I now turn the empirical implications, calibration and estimation for a tightly parametrized version of the model discussed so far.

6 Empirical Implications

In this section I specialize the model for dividends (consumption) to a very parsimonious set-up which also yields simple formulas for the conditional price-dividend ratios B_i. Together with next subsection where an empirical analysis of the model for consumption is carried out, this section also demonstrates the flexibility of the approach of discretizing the state-space, as in Assumption 2 in Section 3. Indeed, this assumption not only makes it possible to obtain closed form solutions for asset prices, but it also makes it relatively easy to implement the model from an econometric standpoint.

Suppose for example that the drift rate of dividends (consumption) follows the pure jump process

$$d\theta_t = (J_t - \theta_t)\, dQ_t^{p(\theta_t)}$$

where $dQ_t^{p(\theta_t)}$ denotes the increment of a Poisson process with intensity $p(\theta_t)$ and J_t is a random variable with any density $f(\theta)$.[7] We can transform this process into a process satisfying Assumption 2 as follows: Let us first select a fine grid $\Theta = [\theta_1, .., \theta_n]$ with boundaries θ_1 and θ_n chosen such that $P(J_t < \theta_1) = P(J_t > \theta_n) \approx 0$. Given the grid Θ, let us also define $p_i = p(\theta_i)$ and $f_i = F(\theta_i + h/2) - F(\theta_i - h/2)$, where h is the interval size of the grid and $F(.)$ is the cumulative distribution on $f(.)$.. Finally, one can then obtain the $n \times n$ infinitesimal matrix Λ as $\lambda_{ij} = p_i f_j$ for $j \neq i$ and $\lambda_{ii} = p_i f_i - p_i$. In this case we have the following corollary:

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

Corollary 5: Let $\lambda_{ij} = p_i f_j$ for $j \neq i$ and $\lambda_{ii} = p_i f_i - p_i$. Define the two constants

$$H_1 = \sum_{i=1}^{n} \frac{f_i k_i}{\phi + \widehat{\theta}_i + p_i} \text{ and } H_2 = \sum_{i=1}^{n} \frac{f_i p_i}{\phi + \widehat{\theta}_i + p_i}$$

where $\widehat{\theta}_i = (1-\gamma)\,\theta_i - \frac{1}{2}(1-\gamma)\,\gamma\sigma_c^2$. Then, for all $i = 1, ..., n$:

$$B_i = \frac{1 - H_2 + \frac{p_i}{k_i} H_1}{\left(\phi + p_i + (1-\gamma)\,\theta_i - \frac{1}{2}(1-\gamma)\,\gamma\sigma_c^2\right)(1 - H_2)}$$

Proof: See Appendix. ■

Notice that if $k_i = 1$ for all i (i.e., no belief dependent utility), then the price dividend ratio conditional on state i is decreasing with θ_i unless also p_i is increasing with i. This is due to the well-known wealth effect for power utility functions. The function k_i that enters into the utility function $U(c, \pi) = \sum_{i=1}^{n} \pi_i k_i c^{1-\gamma} / (1-\gamma)$ has the effect of increasing the conditional price-dividend ratio B_i as θ_i increases as long as $p_i > 0$.[8]

Before describing further the properties of the conditional price-dividend ratios B_i, I need to specify the form of the function $k_i = k(\theta_i)$ which determines the time variation in marginal utility as the distribution $\pi(t)$ change and the form of the distribution $f = (f_1, ..., f_n)$ from which new drifts are chosen if a jump arrives. In addition, we must also bound above and below the possible drifts rates θ to ensure that all the expectations that lead to the pricing formula (32) actually exist.

I start from the form of $f = (f_1, ..., f_n)$ and choose f as the discretization of a normal distribution on the fine grid $\Theta = [\theta_1, ..., \theta_n]$ of possible drifts. That is, let

$$f_i = \int_{\theta_i - \frac{h}{2}}^{\theta_i + \frac{h}{2}} \frac{1}{\sqrt{2\pi}\sigma_\theta} e^{-\frac{1}{2\sigma_\theta^2}(x-\mu)^2} dx \tag{41}$$

In the calibration I will set $\mu = 0.0195$ which is the unconditional annual consumption growth in the data I use below, while $\sigma_\theta = .01$, that is, new drifts would be chosen in the interval $[-.0005, .0395]$ with 95% probability. I also set the boundaries of θ_1 and θ_n so that f_1 and

[8] Indeed, even in the more general case (27) we can see that if $\lambda_{ij} = 0$ for all ij, then the form of the conditional price-dividend ratios B_i collapses to the one with state independent utilities. However, this does not imply that uncertainty has no effect on returns, volatility or interest rates themselves. As it is clear from (37)-(38), we still have $V_B < 0$ and $V_U < 0$ generating some action on returns and interest rates.

$f_n < 1e^{-5}$. In the results below I shall assume $p_i = p$, a constant, which I vary between $p = .1$ (a shift every ten years) and $p = .2$ (a shift every five year) to check how results depend on the probability of shifting away from a particular state. I also set $\sigma_c = 0.0165$, which is the annual volatility of consumption growth. Finally, to approximate well the interval between θ_1 and θ_n without losing in numerical efficiency, I choose $n = 200$.

Turning to the function $k(\theta)$, I choose a simple parametrization of the function $k(\theta)$, specifically:

$$k(\theta) = \overline{k} \times (\theta_1 + 1 + \theta)^{-\rho}$$

where $\overline{k}$ is set so that the unconditional expected $k(\theta)$ equals one, that is $\overline{k} = 1/\left[\sum_{j=1}^{n} f_j (\theta_1 + 1 + \theta)^{-\rho}\right]$. This specific form ensures that $k(\theta) > 0$ for every $\theta \in \Theta$ as required by the characterization in Section 2.2. The parameter ρ defines the curvature of $k(\theta)$. $\rho = 0$ corresponds to the state-independent utility function investigated in Veronesi (2000) which will help making comparisons. Figure 1 plots the conditional price-dividend ratios B_i for the parameters $p = .1$, $\phi = .025$, $\sigma_\theta = .01$, $\mu = .0195$, $\sigma_c = .0165$ and $\gamma = 2$ and for three values of $\rho = 0, 10, 20$. We can see that as the aversion to state uncertainty increases, the conditional price-dividend ratio moves from being negatively sloped with respect to θ_i to being positively sloped. Hence, increases in consumption growth would reduce the price-dividend ratio in the base case $\rho = 0$ but it would increase it for $\rho = 20$.

Table 1 below shows the effect of belief-dependent utilities on expected returns, the interest rate level and their volatilities.[9] In the calibration, investors' posterior distribution $\pi(t)$ is also assumed normal as in (41) with center $\mu = 0.0195$, but I vary its dispersion, call it σ_π, between $\sigma_\pi = .01$ and $\sigma_\pi = .005$ to show the effect of higher or lower uncertainty. Notice in the table that negative entries for the "volatility" simply imply a negative correlation with consumption growth.

First of all, consider the results for $\rho = 0$: This corresponds to set-up also studied in Veronesi (2000) and may set the benchmark. In all cases, we see that as the coefficient of risk aversion γ increases, the equity premium first increases but then decreases turning negative

[9] The reported numbers are obtained using the formulas for expected returns, volatilities and interest rates obtained in closed form in the previous section. Monte Carlo simulations confirmed the results in the table.

for high levels of γ. Similarly, the level of volatility is first decreasing and then increasing in γ, when equilibrium returns become negatively correlated with consumption growth. Both these effects are due to the well-known wealth effect of the power utility function: When investors do not know the growth rate of dividends, they learn about it thereby inducing a positive (instantaneous) autocovariance in their expected growth rate of consumption. As it is well known (see e.g. Campbell (2000)), as we increase the coefficient of risk aversion, the increase in the desire to smooth consumption make investors dump stocks when they obtain good news on dividends and buy stock when they obtain bad news about dividends. This generate a weak or even negative covariance between returns and consumption growth, thereby yielding the low or even negative risk-premia.

In the following, I will fix the coefficient of risk aversion $\gamma = 2$, so that the elasticity of intertemporal substitution is about .5, as many studies show. We can now see that as we increase ρ, i.e. the curvature in $k(\theta)$ that measures the "aversion to state-uncertainty," the equity premium increases, the volatility of stock returns increases, the risk-free rate decreases but its volatility level is U-shaped. So, for example, in case I where $p = .2$ and $\sigma_\pi = .01$ we find that for $\rho = 0$ we have $E[R] = .03\%$, $\sigma_R = 3\%$, $r = 3.46\%$ and $\sigma_r = .003$, while the numbers become $E[R] = 8.1\%$, $\sigma_R = 24\%$, $r = 5.3\%$ and $\sigma_r = -.05$ for $\rho = 50$. We can also see that decreasing the frequency of shifts to $p = .1$ (case II) we obtain a lower equity premium of 6.2%, a lower volatility 18.7%, the same interest rate but with higher interest rate volatility $\sigma_r = -.019\%$ (but lower in absolute value). Decreasing the dispersion σ_π of the distribution $\pi(t)$ has also similar effects (case III and IV): Both the equity premium and the volatility decline, the interest rate decreases but its volatility increases (but decreases in absolute value).

6.1 Fitting Real Consumption Data

In this section I estimate the simple model in the previous section using quarterly data on real consumption growth from 1946-1999. 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. Estimates are obtained by Maximum Likelihood by applying standard regime switching methods to the "discretized" model discussed above.[10,11] The results are reported in Table 2. To gain in estimation efficiency, the mean μ of the jump distribution $f \sim \mathcal{N}\left(\mu, \sigma_\theta^2\right)$ has been set equal to the long-run mean consumption growth.[12]

Table 2 shows that the (annualized) switching probability $p = 0.1816$, implying a "jump" every five years approximately. This is quite in line with the average length of a boom and recession. The annualized volatility of consumption growth has been estimated at $\sigma_c = 0.0134$, slightly below to the sample standard deviation of consumption growth, which is equal to 0.0165. Of course, this is to be expected when there are jumps in the drift as part of the variability of consumption is now captured by the jumps themselves. Finally, we see from Table 2 that if a "jump" occurs, the new drift is chosen according with the distribution $f \sim \mathcal{N}\left(\mu, \sigma_\theta^2\right)$ with $\sigma_\theta = 0.0533$. This is a rather large value which implies that given the average growth rate of consumption equal to $\mu = 0.0195$ the new drift is in the range $[-.0871, 0.1261]$ with probability 95%. Figure 2 plots the density distribution. Although such a wide range in the stationary distribution of the drift parameters is probably due to the first few years in the sample, they still indicate that a somewhat large "uncertainty" may ensue due to "jumps" in the drift rate of consumption. Notice also that the parameter σ_θ is significant at the 5% level, even after correcting the standard errors for heteroskedasticity and autocorrelation. Finally, it

[10] Compared to standard regime switching models as in Hamilton (1989), discretizing the state-space with a large number of states makes it more "likely" the need to deal with the singularities. The Likelihood function tends to $+\infty$ as $p \longrightarrow 1$ or $\sigma_\theta \longrightarrow 0$. As recommended in the literature (see Hamilton (1994)), these problems are typically avoided by restricting the parameters to lie in a compact set that does not include the singular point.

[11] I should notice that the model described in this paper is set in continuous time, while the econometric methodology is set in discrete time. I acknoledge that a bias may result in the estimates, but keep nontheless the present approach for simplicity. In addition, the main point is to show how uncertainty changes over time and how this affects asset prices because of "uncertainty aversion." I am confident that using continuous time estimation methods to correct for the bias would result in second order implications for the empirical results described below.

[12] I also estimated the full model including the mean μ of the distribution f. The estimates of the other parameters were very similar to the ones obtained in Table 2, while $\mu = -0.0025$ (annualized) was found to be not significant (t-stat around 2). The same implications about stock returns and prices were obtained as in Table 3 by assuming the higher discount rate $\phi = 0.06$.

is known indeed that consumption growth was much more variable and volatile for the pre-war sample, which unfortunately is not available at the quarterly frequency. However, if agents used past data to determine the jump distribution f, they would have probably estimated a distribution even more dispersed than the one I obtained.

Figure 3 reports the evolution of the posterior distribution $\pi(t) = (\pi_1(t), ..., \pi_n(t))$ from 1946-1999.[13] It is rather interesting to note the wide fluctuations in the dispersion of the posterior distribution over time. The intuition is rather simple: When a jump occurs (or agents believe it occurred) it takes some time to learn the new drift. 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 $E[\theta] = \sum_{i=1}^{n} \pi_i(t)\theta_i$ (panel A) and the root mean square error of agents distribution $RMSE = \sqrt{V(\theta)}$ where $V(\theta) = E(\theta^2) - \left(E(\theta)^2\right)$ (Panel (B)). The figure clearly shows a high variability of expected drift in the first two years of the sample accompanied by a high root MSE. The "uncertainty" decreased in the 50s and 60s but recovered in the 70s and especially the beginning of the 80s. Higher uncertainty was again realized in the 90s. It is to this uncertainty that the agents in the economy are averse to.

Given the estimated parameters for the underlying consumption process, we need to assess the size of the equity risk premium, volatility, risk-free rate and its volatility. By choosing the parameter values $\rho = 10$, $\gamma = 1.8$ and $\phi = 0.04$ I obtained the unconditional moments for stock returns and interest rate as reported in Table 3. The expected excess return is about 5.26%, slightly below the equity premium in actual data of about 6.8%, while the fitted value of the volatility is about 17.7%, which is slightly above the actual value of 15.96%. The average level of the interest rate is about 1.2%, which is rather close to the standard 2% level considered "reasonable," although the volatility of (changes) in interest rates is about 0.0557, which is somewhat above the realized one of 0.0226.

[13] Given the discretization methodology used to estimate the model, the posterior distribution can be computed easily by Bayes law (see e.g. Hamilton (1989)):

$$\pi_i(t+1) = \frac{e^{-\frac{1}{2\sigma^2}(\Delta c(t+1)-\theta_i)^2}\left[\pi(t)\Lambda\right]_i}{\sum_{j=1}^{n} e^{-\frac{1}{2\sigma^2}\left(\Delta c(t+1)-\theta_j\right)^2}\left[\pi(t)\Lambda\right]_j}$$

where $\Delta c(t) = \ln(c(t+1)) - \ln(c(t))$ is the real consumption growth and Λ is now the transition probability matrix for the discrete time case.

To emphasize the model ability to match also conditional moments, Figures 5 plots the actual ex-post stock return volatility from 1946 to 1999 and the one implied by the model from consumption data. The ex-post volatility is computed as the standard deviation of daily stock returns in the relevant quarter while the model volatility is the one implied by formula (38) for the posterior probability $\pi(t)$ in the same quarter. Clearly, the model overpredicted the volatility at the beginning of the sample and underpredicted the volatility at the end of the sample. However, over the 50 years of sample, the volatility moved in a similar fashion as the one in the data. Figure 6 plots the time series of price-dividend ratios implied by the model. Although the match is not as good as in the case of volatility, I notice that the model generated price-dividend ratio has the same dispersion (at least until the mid 90s) of the realized one as it fluctuate between 15 and 42. One of the problems with this price-dividend ratio as is evident from the plot is that it is too variable compared to the realized one. It turns out that this is due to the rather high value of the jump probability p (= .1816). As p moves closer to 0.1, the match of the price-dividend ratio improves substantially.

7 Conclusions

In this article I re-interpreted standard axioms in the theory of choice delivering a state-dependent utility in order to give a foundation to a "*belief*-dependent" utility function, that is, a utility function representing preferences such that changes in subjective beliefs on an underlying state generate changes in preference orderings over prizes. I show that this interpretation naturally leads to a notion of "aversion to state-uncertainty" and then characterized the (belief-dependent) utility functions displaying constant relative and absolute risk aversion. Since posterior probabilities enter linearly in the utility function, this approach is particularly useful to study dynamic systems where agents update their subjective probability distribution on a an external state (which they cannot control).

In an exchange economy where the "state" is the underlying drift rate of the economy, I argue that aversion to state uncertainty can be naturally interpreted as aversion to long run risk. This is because the drift rate of consumption is responsible for the long-term average consumption path. As a consequence, an "uncertainty averse" investor is averse to the dispersion of long-term average consumption paths. In other words, while "risk aversion" applies to the local volatility of consumption, "uncertainty aversion" applies also to the dispersion of drifts, which by their own nature have a longer term connotation.

In this set up, I obtain closed form solutions for asset prices. Specifically, I obtained interpretable formulas for stock returns and interest rates, showing that higher uncertainty tends to raise the equity premium and the average volatility, decrease the interest rate and generate volatility clustering. The reason for this behavior is that uncertainty changes the marginal utility of investors (for given coefficient of relative risk aversion), making stock returns react more to news on consumption. This increases the equity premium and the demand for the risk-free asset (which decreases the interest rate). I finally show that a simple and parsimonious parametrization of the model makes it possible to match the unconditional first and second moments of stock returns and the risk-free rate. In addition, 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. 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 data.

A final remark is in order: The asset pricing model presented in the previous pages assumes that the "state" is the drift rate of consumption for the economy. The model can be readily extended to account for many other sources of uncertainty, such as inflation rates or political outcomes. The key ingredient to solve the model is that probabilities update in a linear fashion if the underlying environment is Markovian. Hence, computing expectations is relatively simple. Other potentially interesting applications of the set-up proposed in this article are left for future research.

8 References

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

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 (2000) "Prospect theory and asset prices," forthcoming in the *Quarterly Journal of Economics.*

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

Cagetti Marco, Lars P. Hansen, Thomas Sargent and Noha Williamn (2000), "Robustness and pricing with uncertain growth," Working Paper, University of Chicago.

Campbell, John Y. (1996) "Understanding Risk and Returns," *Journal of Political Economy*, 104, 2 298-345.

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.

Constantinides, George (1991) "Habit formation: A resolution of the equity premium puzzle," *Journal of Political Economy*, 98, 842-862.

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

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.

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).

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

Kreps, David M. and E.L. Porteus (1978) "Temporal resolution of uncertainty and dynamic choice theory," *Econometrica*, 46, 185-200.

Liptser, R.S., A.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.

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

Timmerman, 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.*

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

Yan, Hong (1999) "Uncertainty growth prospects, learning and asset prices," University of California at Berkeley.

9 Appendix A: Axiomatic Foundation of State-dependent Utility

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 $\pi : \Xi \rightarrow \Delta(\Theta)$ 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

$$f \succeq_S g \Longleftrightarrow \sum_{\theta\in S} \pi(\theta|S) \sum_{c\in\mathcal{C}} f(c|\theta)\, u(c|\theta) > \sum_{\theta\in S} \pi(\theta|S) \sum_{c\in\mathcal{C}} g(c|\theta)\, u(c|\theta) \tag{42}$$

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 (42) 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 $B : S \to R$ 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 (42) and point (b) for the uniqueness of the probability and utility representation).

10 Appendix B: Proofs of Propositions

Proof of Proposition 1: (a) Using the definition we have

$$\frac{\partial^2 U(c,\pi)}{\partial c^2} = -A \frac{\partial U(c,\pi)}{\partial c} \tag{43}$$

Since $U(c,\pi) = \sum_{\theta \in \Theta} U(c|\theta) \pi(\theta)$ we have

$$\begin{aligned} \frac{\partial U(c,\pi)}{\partial c} &= \sum_{\theta} \pi(\theta) \frac{\partial u(c|\theta)}{\partial c} \\ \frac{\partial U(c,\pi)}{\partial c} &= \sum_{\theta} \pi(\theta) \frac{\partial^2 u(c|\theta)}{\partial c^2} \end{aligned}$$

Hence, we can rewrite (43)

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

This is true for all $\pi(\theta)$ if and only if

$$\frac{\partial^2 u(c|\theta)}{\partial c^2} = -A \frac{\partial u(c|\theta)}{\partial c}$$

This is simple second order differential equation, whose solution is

$$u(c|\theta) = k_1(\theta) + k_2(\theta) e^{-Ac}$$

(b) Using the definition we have

$$-c\frac{\partial^2 U(c,\pi)}{\partial c^2} = \gamma \frac{\partial U(c,\pi)}{\partial c} \tag{44}$$

Hence, (44) 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: From the proof of proposition 1 (b) it is immediate to see that we could let γ be function of θ, i.e. $\gamma = \gamma(\theta)$, obtaining the representation

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

■

Proof of Proposition 4. To prove this proposition, I need the following lemma first:

Lemma 2: Let $n_i = c^{\beta}\pi_i$ and define the matrix $\overline{\Lambda}_{\beta} =$
$\Lambda + diag\left(\theta_1^{\beta}, ..., \theta_n^{\beta}\right)$ with $\theta_i^{\beta} = \beta\theta_i + \frac{1}{2}\beta\left(\beta - 1\right)\sigma_c^2$. Then for $u > t$ we have

$$E\left[n_i\left(u\right)|\mathcal{F}\left(t\right)\right] = \sum_{k=1}^{N} n_k\left(t\right)\sum_{j=1}^{N} w\left(\beta\right)_{jk}^{-1} w_{ij}\left(\beta\right) e^{\omega_j(\beta)(u-t)}$$

where $\omega_j\left(\beta\right)$ are the eigenvalues of $\overline{\Lambda}_{\beta}'$ and $w_{ij}\left(\beta\right)$ are associated eigenvectors and $w\left(\beta\right)_{ij}^{-1} = [W^{-1}]_{ij}$.

Proof of Lemma 2: By Ito's lemma

$$\begin{aligned} dn_i &= c^{\beta}d\pi_i + \beta c^{\beta-1}\pi_i dc + \frac{1}{2}\beta\left(\beta-1\right)\pi_i c^{\beta-2}dc^2 + \beta c^{\beta-1}d\pi_i dc \\ &= c^{\beta}\left[\pi\Lambda\right]_i dt + c^{\beta}\pi_i\sigma_i d\widetilde{W} + \beta c^{\beta-1}\pi_i c\mu_C\left(\pi\right)dt + \beta c^{\beta-1}\pi_i c\sigma_C d\widetilde{W} \\ &\quad + \frac{1}{2}\beta\left(\beta-1\right)\pi_i c^{\beta-2}c^2\sigma_C^2 + \beta c^{\beta-1}\pi_i c\sigma_i\left(\pi\right)\sigma_C dt \\ &= \left\{\left[n\Lambda\right]_i + \beta n_i\mu_C\left(\pi\right) + \frac{1}{2}\beta\left(\beta-1\right)n\sigma_C^2 + \beta n_i\sigma_i\left(\pi\right)\sigma_C\right\}dt \\ &\quad + n_i\left(\sigma_i\left(\pi\right) + \sigma_C\right)d\widetilde{W} \end{aligned}$$

where $n = (n_1, ..., n_n)$. Notice that by Lemma 1:

$$\sigma_i\left(\pi\right)\sigma_C = \left(\theta_i - \sum_{j=1}^{n}\pi_j\theta_j\right)\frac{1}{\sigma_D} \times \sigma_C = \theta_i - \sum_{j=1}^{n}\pi_j\theta_j$$

Hence

$$dn_i = \left\{\left[n\Lambda\right]_i + \frac{1}{2}\beta_i\left(\beta_i - 1\right)n_i\sigma_c^2 + \beta_i n_i\theta_i\right\}dt + n_i\left(\sigma_i\left(\pi\right) + \sigma_c\right)d\widetilde{W}$$

We can write this in vector form: let $n = (n_1, ..., n_N)$ be a $1 \times N$ vector, we then have

$$dn = n\overline{\Lambda}dt + n\Sigma\left(\pi\right)d\widetilde{W} \tag{45}$$

where

$$\overline{\Lambda}_{\beta} = \Lambda + diag\left(\theta_1^{\beta}, ..., \theta_n^{\beta}\right)$$

$$\theta_i^{\beta} = \beta\theta_i + \frac{1}{2}\beta\left(\beta - 1\right)\sigma_c^2$$

and $\Sigma\left(\pi\right)$ is some bounded $n \times 1$ vector.

Let

$$\widetilde{n}\left(u\right) = E\left[\widetilde{n}\left(u\right)|\mathcal{F}\left(t\right)\right]$$

We can write (45) in integral form

$$n(u) = n(t) + \int_t^u n(s)\,\overline{\Lambda}_\beta ds + \int_t^u n(s)\,\Sigma(\pi(s))\,d\widetilde{W}(s)$$

Taking expectations on both sides and using the fact that the stochastic integral has zero expectations, we then have

$$\widetilde{n}(u) = n(t) + \int_t^u \widetilde{n}(s)\,\overline{\Lambda}_\beta ds$$

This can we rewritten as

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

The solution to this system of ordinary differential equations with initial condition $\widetilde{n}(t) = n(t)$ is

$$\widetilde{n}_i(u) = \sum_{k=1}^{N} n_k(t) \sum_{j=1}^{N} w(\beta)_{jk}^{-1}\, w_{ij}(\beta)\, e^{\omega_j(\beta)(u-t)}$$

where $\omega_j(\beta)$ are the eigenvalues of $\overline{\Lambda}'_\beta$ and $w_{ij}(\beta)$ are associated eigenvectors and $w(\beta)_{ij}^{-1} = [W^{-1}]_{ij}$. This concludes the proof of the lemma.■

Proof of proposition 4: 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. For notational convenience, let me write $\gamma_i = \gamma(\theta_i)$ and $k_i = k(\theta_i)$. Hence, the price of an asset must satisfy

$$\begin{aligned} P(t) &= E_t\left[\int_t^\infty \frac{U_c[c(s), s, \pi(s)]}{U_c[c(t), t, \pi(t)]} D(s)\,ds\right] \\ &= \frac{1}{U_c[C(t), t, \pi(t)]} E_t\left[\int_t^\infty e^{-\phi s} \sum_i k_i c(s)^{-\gamma_i} \pi_i(s)\, D(s)\,ds\right] \\ &= \frac{1}{U_c[C(t), t, \pi(t)]} E_t\left[\int_t^\infty e^{-\phi s} \sum_i k_i c(s)^{1-\gamma_i} \pi_i(s)\,ds\right] \end{aligned}$$

The consumption process follow

$$\frac{dc}{c} = \mu_C(\pi)\,dt + \sigma_C d\widetilde{W}$$

where $\mu_C(\pi) = \sum_{i=1}^n \pi_i \theta_i$ and $\sigma_C = \sigma_D$. Hence, we can use the result in Lemma 2 to obtain the value of $P(t)$

$$
\begin{aligned}
P(t) &= \frac{1}{U_c[c(t), t, \pi(t)]} E_t\left[\int_t^\infty e^{-\phi s} \sum_i k_i c(s)^{1-\gamma_i} \pi_i(s)\, ds\right] \\
&= \frac{1}{e^{-\phi t}\sum_{i=1}^n k_i \pi_i(t) C(t)^{-\gamma_i}} \left\{\int_t^\infty e^{-\phi s} \sum_i k_i E_t\left[c(s)^{1-\gamma_i} \pi_i(s)\right] ds\right\} \\
&= \frac{1}{e^{-\phi t}\sum_{i=1}^n k_i \pi_i(t) C(t)^{-\gamma_i}} \left\{\int_t^\infty e^{-\phi s} \sum_i k_i \widetilde{n}_i(s)\, ds\right\}
\end{aligned}
$$

We now apply the result of Lemma 2 and substitute for each

$$
\widetilde{n}_i(s) = \sum_{k=1}^N n_k(t) \sum_{j=1}^N w(\beta_i)^{-1}_{jk}\, w_{ij}(\beta_i)\, e^{\omega_j(\beta_i)(s-t)}
$$

where for all $i = 1, ..., n$ we have $\beta_i = 1 - \gamma_i$ and $w_{ij}(\beta_i)$ and $\omega_j(\beta_i)$ are the eigenvectors and eigenvalues of the matrix

$$
\overline{\Lambda}_i = \Lambda + diag\left(\widehat{\theta}_1^i, ..., \widehat{\theta}_n^i\right)
$$

with $\widehat{\theta}_j^i = (1-\gamma_i)\theta_j - \frac{1}{2}(1-\gamma_i)\gamma_i\sigma_c^2$. This yields

$$
\begin{aligned}
P(t) &= \frac{1}{e^{-\phi t}\sum_{i=1}^n k_i \pi_i(t) c(t)^{-\gamma_i}} \left\{\int_t^\infty e^{-\phi s} \sum_{i=1}^N k_i \sum_{k=1}^N n_k(t) \sum_{j=1}^N w(\beta_i)^{-1}_{jk}\, w_{ij}(\beta_i)\, e^{\omega_j(\beta_i)(s-t)} ds\right\} \\
&= \frac{1}{\sum_{i=1}^n k_i \pi_i(t) c(t)^{-\gamma_i}} \left\{\int_t^\infty \sum_{i=1}^N k_i \sum_{k=1}^N n_k(t) \sum_{j=1}^N w(\beta_i)^{-1}_{jk}\, w(\beta_i)_{ij}\, e^{(\omega_j(\beta_i)-\phi)(s-t)} ds\right\} \\
&= \frac{1}{\sum_{i=1}^n k_i \pi_i(t) c(t)^{-\gamma_i}} \left\{\sum_{i=1}^N k_i \sum_{k=1}^N n_k(t) \sum_{j=1}^N w(\beta_i)^{-1}_{jk}\, w(\beta_i)_{ij} \int_t^\infty e^{(\omega_j(\beta_i)-\phi)(s-t)} ds\right\} \\
&= \frac{1}{\sum_{i=1}^n k_i \pi_i(t) c(t)^{-\gamma_i}} \left\{\sum_{k=1}^N n_k(t) \sum_{i=1}^N \sum_{j=1}^N k_i w(\beta_i)^{-1}_{jk}\, w(\beta_i)_{ij} \frac{1}{\phi - \omega_j(\beta_i)}\right\} \\
&= \frac{1}{\sum_{i=1}^n k_i \pi_i(t) c(t)^{-\gamma_i}} \sum_{k=1}^N n_k(t) \widehat{B}_k \\
&= \frac{\sum_{k=1}^n \pi_k(t) c(t)^{1-\gamma_k} \widehat{B}_k}{\sum_{i=1}^n k_i \pi_i(t) c(t)^{-\gamma_i}}
\end{aligned}
$$

where

$$
\widehat{B}_k = \sum_{i=1}^N \sum_{j=1}^N k(\theta_i)\, w(\beta_i)^{-1}_{jk}\, w(\beta_i)_{ij} \frac{1}{\phi - \omega_j}
$$

Let

$$\pi_k^*(t,c) = \frac{k_k \pi_k(t)\, c(t)^{-\gamma_k}}{\sum_{i=1}^n k_i \pi_i(t)\, c(t)^{-\gamma_i}}$$

$$B_k = \frac{\widehat{B}_k}{k_k} = \sum_{i=1}^N \sum_{j=1}^N k_i w(\beta_i)_{jk}^{-1} w(\beta_i)_{ij} \frac{1}{\phi - \omega_j}$$

Then we can rewrite

$$P(t) = D(t) \sum_{k=1}^n \pi_k^*(t,c) B_k$$

We finally prove that

$$\widehat{B}_k = \sum_{i=1}^n k_i e_i \left(\phi \mathbf{I} - \overline{\Lambda}'_{\beta_i}\right)^{-1} e_k$$

That is, that

$$\sum_{j=1}^n w(\beta_i)_{jk}^{-1} w(\beta_i)_{ij} \frac{1}{\phi - \omega_j(\beta_i)} = e_i \left(\phi \mathbf{I} - \overline{\Lambda}'_{\beta_i}\right)^{-1} e_k$$

Let Ω_{β_i} be the he diagonal matrix with the eigenvalues $\omega_j(\beta_i)$ of $\overline{\Lambda}'_{\beta_i}$ on the principal diagonal. Then we know that

$$\left(\phi \mathbf{I} - \overline{\Lambda}'_{\beta_i}\right)^{-1} = W_{\beta_i} \left(\mathbf{I}\phi - \Omega_{\beta_i}\right)^{-1} W_{\beta_i}^{-1}$$

where W_{β_i} is the matrix with the eigenvectors of $\overline{\Lambda}'_{\beta_i}$ as columns. Let $D(\beta_i) = \left(\mathbf{I}\phi - \Omega_{\beta_i}\right)^{-1}$: Since this is a diagonal matrix, we finally obtain

$$e_i \left(\phi \mathbf{I} - \overline{\Lambda}'_{\beta_i}\right)^{-1} e_k = \sum_j \sum_\ell w(\beta_i)_{ij} D(\beta_i)_{j\ell} w(\beta_i)_{\ell k}^{-1} = \sum_{j=0}^n \frac{w(\beta_i)_{ij} w(\beta_i)_{jk}^{-1}}{\phi - \omega_j(\beta_i)}$$

as was to be shown.

(b) We know that the real rate of interest rate is given by

$$r(t) = -E_t \left[\frac{dm}{m}\right]$$

where $m(t)$ is the real pricing kernel given by

$$m(t) = \partial U / \partial c = e^{-\phi t} \sum_{i=1}^n k_i \pi_i(t)\, c(t)^{-\gamma_i}$$

We then have

$$dm(t) = -\phi m(t) + e^{-\phi t} \sum_{i=1}^n k_i c^{-\gamma_i} d\pi_i + e^{-\phi t} \sum_{i=1}^n k_i \pi_i (-\gamma_i)\, c(t)^{-\gamma_i - 1} dc$$

$$
\begin{aligned}
& +e^{-\phi t}\frac{1}{2}\sum_{i=1}^{n}k_i\pi_i\left(-\gamma_i\right)\left(-\gamma_i-1\right)c\left(t\right)^{-\gamma_i-2}dc^2+e^{-\phi t}\sum_{i=1}^{n}k_i\left(-\gamma_i\right)c^{-\gamma_i-1}d\pi_i dc \\
= \quad & -\phi m\left(t\right)+e^{-\phi t}\sum_{i=1}^{n}k_i c^{-\gamma_i}\left[\pi\Lambda\right]_i dt+e^{-\phi t}\sum_{i=1}^{n}k_i c^{-\gamma_i}\pi_i\left(\theta_i-\sum_{j=1}^{n}\pi_j\theta_j\right)\frac{1}{\sigma_c}d\widetilde{W} \\
& -e^{-\phi t}\sum_{i=1}^{n}k_i\pi_i\gamma_i c\left(t\right)^{-\gamma_i}\mu_c\left(\pi\right)-e^{-\phi t}\sum_{i=1}^{n}k_i\pi_i\gamma_i c\left(t\right)^{-\gamma_i}\sigma_c d\widetilde{W} \\
& +e^{-\phi t}\frac{1}{2}\sum_{i=1}^{n}k_i\pi_i\gamma_i\left(\gamma_i+1\right)c\left(t\right)^{-\gamma_i}\sigma_c^2-e^{-\phi t}\sum_{i=1}^{n}k_i\gamma_i c^{-\gamma_i}\pi_i\left(\theta_i-\sum_{j=1}^{n}\pi_j\theta_j\right) \\
= \quad & -\phi m\left(t\right)+e^{-\phi t}\sum_{i=1}^{n}k_i c^{-\gamma_i}\left[\pi\Lambda\right]_i dt+\frac{1}{2}e^{-\phi t}\sum_{i=1}^{n}k_i\pi_i\gamma_i\left(\gamma_i+1\right)c\left(t\right)^{-\gamma_i}\sigma_c^2 dt \\
& -e^{-\phi t}\sum_{i=1}^{n}k_i\gamma_i c^{-\gamma_i}\pi_i\theta_i dt+e^{-\phi t}\sum_{i=1}^{n}k_i c^{-\gamma_i}\pi_i\left(\theta_i-\sum_{j=1}^{n}\pi_j\theta_j\right)\frac{1}{\sigma_c}d\widetilde{W} \\
& -e^{-\phi t}\sum_{i=1}^{n}k_i\pi_i\gamma_i c\left(t\right)^{-\gamma_i}\sigma_c d\widetilde{W}
\end{aligned}
$$

Hence

$$
\begin{aligned}
r\left(t\right)=-E_t\left[\frac{dm}{m}\right] \quad = \quad & \phi+\frac{1}{\sum_{i=1}^{n}k_i\pi_i\left(t\right)c\left(t\right)^{-\gamma_i}}\left(+\sum_{i=1}^{n}k_i\gamma_i c^{-\gamma_i}\pi_i\theta_i\right. \\
& \left.-\frac{1}{2}\sum_{i=1}^{n}k_i\pi_i\gamma_i\left(\gamma_i+1\right)c\left(t\right)^{-\gamma_i}\sigma_c^2-\sum_{i=1}^{n}k_i c^{-\gamma_i}\left[\pi\Lambda\right]_i\right)
\end{aligned}
$$

By redefining variables we obtain expression (29). ■

Proof of Corollary 1: (a) is immediate and (b) stems from the following manipulation:

$$
\begin{aligned}
r\left(t\right) \quad = \quad & \phi+\gamma\sum_{i=1}^{n}\pi_i^*\left(t\right)\theta_i-\frac{1}{2}\gamma\left(\gamma+1\right)\sigma_c^2-\sum_{j=1}^{n}\pi_j^*C_j^* \\
= \quad & \phi+\gamma\sum_{i=1}^{n}\pi_i\left(t\right)\theta_i+\gamma\left(\sum_{i=1}^{n}\pi_i^*\left(t\right)\theta_i-\sum_{i=1}^{n}\pi_i\left(t\right)\theta_i\right)-\frac{1}{2}\gamma\left(\gamma+1\right)\sigma_c^2-\sum_{j=1}^{n}\pi_j^*C_j^* \\
= \quad & r_B\left(t\right)+\gamma V_U-\sum_{j=1}^{n}\pi_j^*C_j^*
\end{aligned}
$$

■

Proof of Lemma 2: The proof is analogous (albeit with different interpretation) to Lemma 3 in Veronesi (2000). ■

Proof of Proposition 5: The price of the asset in the constant relative risk aversion case is

$$P(t) = D(t)\frac{\sum_{j=1}^{n} \pi_j(t) k_j B_j}{\sum_{i=1}^{n} k(\theta_i) \pi_i(t)}$$

Let $X(t) = \sum_i k_i \pi_i(t)$. Hence,

$$dX = \sum_i k_i d\pi_i = \sum_i k_i [\pi\Lambda]_i dt + \sum_i k_i \pi_i \left(\theta_i - \sum_{j=1}^{n} \pi_j \theta_j\right) \frac{1}{\sigma_c} d\widetilde{W}$$

or

$$\begin{aligned}
\frac{dX}{X} &= \frac{\sum_i k_i [\pi\Lambda]_i}{\sum_i k(\theta_i) \pi_i(t)} dt + \frac{\sum_i k_i \pi_i \left(\theta_i - \sum_{j=1}^{n} \pi_j \theta_j\right)}{\sum_i k(\theta_i) \pi_i(t)} \frac{1}{\sigma_c} d\widetilde{W} \\
&= \mu_X dt + \sigma_X d\widetilde{W}
\end{aligned}$$

It is convenient to define

$$\widetilde{P}(t) = D(t) \sum_{i=1}^{n} \pi_i k_i B_i$$

We then have

$$\begin{aligned}
d\widetilde{P} &= \sum_{i=1}^{n} \pi_i k_i B_i dD + D \sum_{i=1}^{n} k_i B_i d\pi_i + \sum_{i=1}^{n} k_i B_i d\pi_i dD \\
&= \sum_{i=1}^{n} \pi_i k_i B_i \mu_D dt + \sum_{i=1}^{n} \pi_i k_i B_i \sigma_D d\widetilde{W} + D \sum_{i=1}^{n} k_i B_i [\pi\Lambda]_i dt \\
&\quad + D \sum_{i=1}^{n} k_i B_i \pi_i \left(\theta_i - \sum_{j=1}^{n} \pi_j \theta_j\right) \frac{1}{\sigma_c} d\widetilde{W} + D \sum_{i=1}^{n} k_i B_i \pi_i \left(\theta_i - \sum_{j=1}^{n} \pi_j \theta_j\right) dt \\
&= \widetilde{P} \left(\mu_D + \frac{\sum_{i=1}^{n} k_i B_i [\pi\Lambda]_i + \sum_{i=1}^{n} k_i B_i \pi_i \left(\theta_i - \sum_{j=1}^{n} \pi_j \theta_j\right)}{\sum_{i=1}^{n} \pi_i k_i B_i}\right) dt \\
&\quad + \widetilde{P} \left(\sigma_D + \frac{\sum_{i=1}^{n} k_i B_i \pi_i \left(\theta_i - \sum_{j=1}^{n} \pi_j \theta_j\right)}{\sum_{i=1}^{n} \pi_i k_i B_i} \frac{1}{\sigma_c}\right) d\widetilde{W} \\
&= \widetilde{P} \widetilde{\mu}_P dt + \widetilde{P} \widetilde{\sigma}_P d\widetilde{W}
\end{aligned}$$

Notice that since $P = \widetilde{P}/X$ we have

$$\begin{aligned}
dP &= \frac{1}{X} d\widetilde{P} - \frac{\widetilde{P}}{X} \frac{dX}{X} + \frac{\widetilde{P}}{X} \frac{dX^2}{X^2} - \frac{1}{X} d\widetilde{P} \frac{dX}{X} \\
&= P \widetilde{\mu}_P dt + P \widetilde{\sigma}_P d\widetilde{W} - P \mu_X dt - P \sigma_X d\widetilde{W} + P \sigma_X^2 dt - P \widetilde{\sigma}_P \sigma_X dt
\end{aligned}$$

So that, we obtain

$$\frac{dP}{P} = \left(\widetilde{\mu}_P - \mu_X + \sigma_X^2 - \widetilde{\sigma}_P \sigma_X\right) dt + \left(\widetilde{\sigma}_P - \sigma_X\right) d\widetilde{W}$$

The equilibrium condition requires that

$$dR = \frac{dP + Ddt}{P} - rdt$$

is such that

$$E_t\left[dR\right] = -Cov\left(dR, \frac{dm}{m}\right)$$

In this case

$$m\left(t\right) = \partial U/\partial c = e^{-\phi t} \sum_{i=1}^{n} k\left(\theta_i\right) \pi_i\left(t\right) c\left(t\right)^{-\gamma}$$

and hence

$$\begin{aligned} dm &= m\left(-\phi + \frac{\sum_i k_i \left[\pi\Lambda\right]_i}{\sum_j k_j \pi_j} + \gamma\left(\gamma+1\right)\sigma_c^2 - \gamma\frac{\sum_{i=1}^n k_i \pi_i \theta_i dt}{\sum_{i=1}^n k_i \pi_i dt}\right) dt \\ &\quad + m\left(-\gamma\sigma_c + \frac{\sum_{i=1}^n k_i \pi_i \left(\theta_i - \sum_{j=1}^n \pi_j \theta_j\right)}{\sum_{i=1}^n k_i \pi_i\left(t\right)} \frac{1}{\sigma_c}\right) d\widetilde{W} \\ &= m\mu_m dt + m\sigma_m d\widetilde{W} \end{aligned}$$

Hence

$$\begin{aligned} E\left[dR\right] &= -\left(\widetilde{\sigma}_P - \sigma_X\right)\sigma_m dt \\ &= -\left(\sigma_D + \frac{\sum_{i=1}^n k_i B_i \pi_i \left(\theta_i - \sum_{j=1}^n \pi_j \theta_j\right)}{\sum_{i=1}^n \pi_i k_i B_i} \frac{1}{\sigma_c} - \frac{\sum_i k_i \pi_i \left(\theta_i - \sum_{j=1}^n \pi_j \theta_j\right)}{\sum_i k_i \pi_i} \frac{1}{\sigma_c}\right) \\ &\quad \times \left(-\gamma\sigma_c + \frac{\sum_{i=1}^n k_i \pi_i \left(\theta_i - \sum_{j=1}^n \pi_j \theta_j\right)}{\sum_{i=1}^n k_i \pi_i} \frac{1}{\sigma_c}\right) \\ &= \gamma\left(\sigma_c^2 + \frac{\sum_{i=1}^n k_i B_i \pi_i \left(\theta_i - \sum_{j=1}^n \pi_j \theta_j\right)}{\sum_{i=1}^n \pi_i k_i B_i}\right) \\ &\quad -\gamma\frac{\sum_i k_i \pi_i \left(\theta_i - \sum_{j=1}^n \pi_j \theta_j\right)}{\sum_i k_i \pi_i} - \frac{\sum_{i=1}^n k_i \pi_i \left(\theta_i - \sum_{j=1}^n \pi_j \theta_j\right)}{\sum_{i=1}^n k_i \pi_i} \\ &\quad - \frac{\sum_{i=1}^n k_i \pi_i \left(\theta_i - \sum_{j=1}^n \pi_j \theta_j\right)}{\sum_{i=1}^n k_i \pi_i} \frac{\sum_{i=1}^n k_i B_i^* \pi_i \left(\theta_i - \sum_{j=1}^n \pi_j \theta_j\right)}{\sum_{i=1}^n \pi_i k_i B_i^*} \frac{1}{\sigma_c^2} \end{aligned}$$

$$
\begin{aligned}
&+\left(\frac{\sum_{i=1}^{n} k_i\pi_i\left(\theta_i-\sum_{j=1}^{n}\pi_j\theta_j\right)}{\sum_{i=1}^{n} k_i\pi_i}\frac{1}{\sigma_c}\right)^2 \\
= \quad &\gamma\left(\sigma_c^2+\frac{\sum_{i=1}^{n} k_iB_i\pi_i\left(\theta_i-\sum_{j=1}^{n}\pi_j\theta_j\right)}{\sum_{i=1}^{n}\pi_ik_iB_i}\right) \\
&-(\gamma+1)\frac{\sum_i k_i\pi_i\left(\theta_i-\sum_{j=1}^{n}\pi_j\theta_j\right)}{\sum_i k_i\pi_i}+\left(\frac{\sum_{i=1}^{n} k_i\pi_i\left(\theta_i-\sum_{j=1}^{n}\pi_j\theta_j\right)}{\sum_{i=1}^{n} k_i\pi_i}\frac{1}{\sigma_c}\right)^2 \\
&-\frac{\sum_{i=1}^{n} k_i\pi_i\left(\theta_i-\sum_{j=1}^{n}\pi_j\theta_j\right)}{\sum_{i=1}^{n} k_i\pi_i}\frac{\sum_{i=1}^{n} k_iB_i\pi_i\left(\theta_i-\sum_{j=1}^{n}\pi_j\theta_j\right)}{\sum_{i=1}^{n}\pi_ik_iB_i}\frac{1}{\sigma_c^2} \\
= \quad &\gamma\left(\sigma_c^2+V_B\right)-(\gamma+1)V_U+1/\sigma_c^2V_U^2-1/\sigma_c^2V_BV_U
\end{aligned}
$$

Or in a more concise way

$$
E\left[dR\right]=\gamma\left(\sigma_c^2+V_B\right)-(\gamma+1)V_U+\frac{1}{\sigma_c^2}V_U\left(V_U-V_B\right)
$$

where

$$
V_B=\frac{\sum_{i=1}^{n} k_iB_i\pi_i\left(\theta_i-\sum_{j=1}^{n}\pi_j\theta_j\right)}{\sum_{i=1}^{n}\pi_ik_iB_i} \text{ and } V_U=\frac{\sum_{i=1}^{n} k_i\pi_i\left(\theta_i-\sum_{j=1}^{n}\pi_j\theta_j\right)}{\sum_{i=1}^{n} k_i\pi_i}
$$

(b) The interest rate can be written

$$
r(t)=\text{const.}+\frac{\sum_{i=1}^{n}\pi_iC_i^r}{X}=\text{const.}+\frac{\widetilde{r}}{X}
$$

where $X=\sum_{i=1}^{n}k_i\pi_i$ and $C_i^r=\gamma k_i\theta_i-C_i^*$. We have

$$
\begin{aligned}
d\widetilde{r} &= \sum_i C_i^r\left[\pi\Lambda\right]_i dt+\sum_i C_i^r\pi_i\left(\theta_i-\sum_j\pi_j\theta_j\right)\frac{1}{\sigma_c}d\widetilde{W} \\
&= \widetilde{\mu}_r dt+\widetilde{\sigma}_r d\widetilde{W} \\
\frac{dX}{X} &= \frac{\sum_i k_i\left[\pi\Lambda\right]_i}{\sum_{i=1}^{n}k_i\pi_i}dt+\frac{\sum_i k_i\pi_i\left(\theta_i-\sum_j\pi_j\theta_j\right)}{\sum_{i=1}^{n}k_i\pi_i}\frac{1}{\sigma_c}d\widetilde{W} \\
&= \mu_X dt+\sigma_X d\widetilde{W}
\end{aligned}
$$

Hence,

$$
\begin{aligned}
dr &= \frac{1}{X}d\widetilde{r}-\frac{\widetilde{r}}{X}\frac{dX}{X}+\frac{\widetilde{r}}{X}\left(\frac{dX}{X}\right)^2-\frac{1}{X}d\widetilde{r}\frac{dX}{X} \\
&= \widetilde{\mu}_r^* dt+\widetilde{\sigma}_r^* d\widetilde{W}-r\mu_X dt-r\sigma_X d\widetilde{W}+r\sigma_X^2 dt-\widetilde{\sigma}_r^*\sigma'_X dt
\end{aligned}
$$

where

$$\widetilde{\mu}_r^* = \frac{\sum_i C_i^r \left[\pi\Lambda\right]_i}{\sum_{i=1}^n k_i \pi_i} \text{ and } \widetilde{\sigma}_r^* = \frac{\sum_i C_i^r \pi_i \left(\theta_i - \sum_j \pi_j \theta_j\right)}{\sum_{i=1}^n k_i \pi_i} \frac{1}{\sigma_c}$$

Notice that

$$\begin{aligned} r &= \phi + \gamma \frac{\sum_{i=1}^n \pi k_i \theta_i}{X} - \frac{1}{2}\gamma\left(\gamma+1\right)\sigma_c^2 - \frac{\sum_{j=1}^n \pi_j k_j C_j^*}{X} \\ &= \phi + \frac{\sum_{i=1}^n \pi_i C_i^r}{X} - \frac{1}{2}\gamma\left(\gamma+1\right)\sigma_c^2 \end{aligned}$$

Hence

$$X = \frac{\sum_{i=1}^n \pi_i C_i^r}{r - \phi + \frac{1}{2}\gamma\left(\gamma+1\right)\sigma_c^2}$$

and therefore

$$\begin{aligned} \widetilde{\sigma}_r^* &= \frac{\sum_i C_i^r \pi_i \left(\theta_i - \sum_j \pi_j \theta_j\right)}{\sum_{i=1}^n k\left(\theta_i\right) \pi_i} \frac{1}{\sigma_c} \\ &= \left(r - \phi + \frac{1}{2}\gamma\left(\gamma+1\right)\sigma_c^2\right) \frac{\sum_i C_i^r \pi_i \left(\theta_i - \sum_j \pi_j \theta_j\right)}{\sum_{i=1}^n \pi_i\left(t\right) C_i^r} \frac{1}{\sigma_c} \\ &= \left(r - \phi + \frac{1}{2}\gamma\left(\gamma+1\right)\sigma_c^2\right) \frac{1}{\sigma_c} V_r \end{aligned}$$

So that

$$\sigma_r = r h_D\left(V_r + V_U\right) - \left(\phi + \frac{1}{2}\gamma\left(\gamma+1\right)\sigma_c^2\right) h_D V_r$$

■

Proof of Corollary 5: Define $k_i = k\left(\theta_i\right)$ and $k = k_1, .., k_n$. From the formula for B_k in (27) and $\overline{\Lambda}_i'$ being independent of i in (24) we find

$$B_j = \frac{1}{k_j} \sum_{i=1}^n k_i e_i' \left(\phi\mathbf{I} - \overline{\Lambda}'\right)^{-1} e_j$$

or

$$B_j k_j = k' \left(\phi\mathbf{I} - \overline{\Lambda}'\right)^{-1} e_j = e_j' \left(\phi\mathbf{I} - \overline{\Lambda}\right)^{-1} k$$

In vector notation

$$diag\left(k\right) \times B = \left(\phi\mathbf{I} - \overline{\Lambda}\right)^{-1} k$$

which yields

$$\left(\phi\mathbf{I} - \overline{\Lambda}\right) diag\left(k\right) \times B = k$$

Recall now $\overline{\Lambda} = \Lambda + diag\left(\widehat{\theta}_1, ..., \widehat{\theta}_n\right)$ with $\widehat{\theta}_j = (1-\gamma)\,\theta_j - \frac{1}{2}(1-\gamma)\,\gamma\sigma_c^2$ Hence, we have

$$\phi D(k)\,B - \Lambda D(k)\,B - D\left(\widehat{\theta}\right) D(k)\,B = k$$

or

$$\phi k_i B_i + \widehat{\theta}_i k_i B_i = k_i + \sum_{j=1}^{n} \lambda_{ij} k_j B_j$$

Using the assumption on λ_{ij} we obtain

$$\begin{aligned}\left(\phi + \widehat{\theta}_i\right) k_i B_i &= k_i + \sum_{j \neq i} p_i f_j k_j B_j + (p_i f_i - p_i)\, k_i B_i \\ &= k_i + \sum_{j=1}^{n} p_i f_j k_j B_j - p_i k_i B_i\end{aligned}$$

Taking the last term to the right hand side, we have

$$\left(\phi + \widehat{\theta}_i + p_i\right) k_i B_i = k_i + p_i \sum_{j=1}^{n} f_j k_j B_j$$

or

$$k_i B_i = \frac{k_i}{\phi + \widehat{\theta}_i + p_i} + \frac{p_i}{\phi + \widehat{\theta}_i + p_i}\left(\sum_{j=1}^{n} f_j k_j B_j\right) \tag{46}$$

Multiply each side by f_i and sum across i to obtain

$$\begin{aligned}\sum_{i=1}^{n} f_i k_i B_i &= \sum_{i=1}^{n} \frac{f_i k_i}{\phi + \widehat{\theta}_i + p_i} + \sum_{i=1}^{n} \frac{f_i p_i}{\phi + \widehat{\theta}_i + p_i}\left(\sum_{j=1}^{n} f_j k_j B_j\right) \\ &= H_1 + H_2\left(\sum_{j=1}^{n} f_j k_j B_j\right)\end{aligned}$$

where H_1 and H_2 are defined accordingly. Hence, assuming $1 - H_2 \neq 0$ we obtain

$$\sum_{i=1}^{n} f_i k_i B_i = \frac{H_1}{1 - H_2}$$

We can plug this back into (46) to obtain

$$\begin{aligned}B_i &= \frac{1}{\phi + p_i + (1-\gamma)\,\theta_i - \frac{1}{2}(1-\gamma)\,\gamma\sigma_c^2} \\ &\quad + \frac{p_i}{k(\theta_i)\left(\phi + p_i + (1-\gamma)\,\theta_i - \frac{1}{2}(1-\gamma)\,\gamma\sigma_c^2\right)}\left(\frac{H_1}{1 - H_2}\right) \\ &= \frac{1 - H_2 + \frac{p_i}{k(\theta_i)} H_1}{\left(\phi + p_i + (1-\gamma)\,\theta_i - \frac{1}{2}(1-\gamma)\,\gamma\sigma_c^2\right)(1 - H_2)}\end{aligned}$$

■

Proof of Lemma 3: Fix the distribution $(\pi_1, ..., \pi_n)$. By definition, we have

$$
\begin{aligned}
V_U &= \frac{\sum_{i=1}^n \pi_i k_i \left(\theta_i - \sum_{j=1}^n \pi_j \theta_j\right)}{\sum_{j=1}^n \pi_j k_j} = \sum_{i=1}^n \pi_i^k \theta_i - \sum_{i=1}^n \pi_i \theta_i \\
V_B &= \frac{\sum_{i=1}^n \pi_i k_i B_i \left(\theta_i - \sum_{j=1}^n \pi_j \theta_j\right)}{\sum_{j=1}^n \pi_j k_j B_i} = \sum_{i=1}^n \pi_i^{kB} \theta_i - \sum_{i=1}^n \pi_i \theta_i
\end{aligned}
$$

where

$$
\pi_i^k = \frac{\pi_i k_i}{\sum_{j=1}^n \pi_j k_j} \text{ and } \pi_i^{kB} = \frac{\pi_i k_i B_i}{\sum_{j=1}^n \pi_j k_j B_j}
$$

Hence, we have $V_B - V_U = \sum_{i=1}^n \pi_i^{kB} \theta_i - \sum_{i=1}^n \pi_i^k \theta_i$. Notice that we can rewrite

$$
\pi_i^{kB} = B_i \pi_i^k \times \left(\frac{\sum_{j=1}^n \pi_j k_j}{\sum_{j=1}^n \pi_j k_j B_j} \right)
$$

Since B_i is increasing in i, and $\sum_{i=1}^n \pi_i^k = \sum_{i=1}^n \pi_i^{kB} = 1$, it must be the case that there exists ι such that $\pi_i^{kB} > \pi_i^k$ if and only if $i > \iota$. Hence, π_i^{kB} is giving higher weight to high θ_i's, which implies $\sum_{i=1}^n \pi_i^{kB} \theta_i > \sum_{i=1}^n \pi_i^k \theta_i$, yielding the claim.■

Table 1: Calibration Exercise

		Case I $p = 0.2$, $\sigma_\pi = 0.01$						Case II $p = 0.1$, $\sigma_\pi = 0.01$			
γ	ρ	$E[dR]$	σ_R	r	σ_r	γ	ρ	$E[dR]$	σ_R	r	σ_r
0.5	0	0.00025	0.0306	0.0346	0.003	0.5	0	0.00035	0.0429	0.0346	0.003
0.5	1	0.0005	0.0362	0.0346	0.002	0.5	1	0.0007	0.0481	0.0346	0.0026
0.5	10	0.006	0.087	0.0341	-0.0077	0.5	10	0.0066	0.0956	0.0341	-0.0016
0.5	50	0.0981	0.3136	0.0321	-0.0668	0.5	50	0.0961	0.3074	0.0321	-0.0281
0.5	100	0.3699	0.5979	0.0296	-0.3072	0.5	100	0.3546	0.5732	0.0296	-0.1445
0.5	200	1.3891	1.1477	0.0247	-12.2504	0.5	200	1.3155	1.0869	0.0247	-6.1089
1	0	0.00027	0.0165	0.0442	0.0061	1	0	0.00027	0.0165	0.0442	0.0061
1	1	0.0005	0.0219	0.0441	0.005	1	1	0.0005	0.0214	0.0441	0.0056
1	10	0.0054	0.0704	0.0432	-0.0047	1	10	0.005	0.0651	0.0432	0.0014
1	50	0.0922	0.2872	0.0392	-0.0638	1	50	0.0835	0.2601	0.0392	-0.0251
1	100	0.3505	0.559	0.0342	-0.3043	1	100	0.3164	0.5048	0.0342	-0.1416
1	200	1.3221	1.085	0.0244	-12.2479	1	200	1.1919	0.9781	0.0244	-6.1064
2	0	-0.0003	-0.0084	0.0632	0.0121	2	0	-0.0009	-0.0259	0.0632	0.0121
2	1	-0.0001	-0.0034	0.063	0.0111	2	1	-0.0008	-0.0217	0.063	0.0117
2	10	0.0039	0.0414	0.0612	0.0014	2	10	0.0015	0.0163	0.0612	0.0075
2	50	0.0814	0.2413	0.0531	-0.0579	2	50	0.0627	0.1857	0.0531	-0.0192
2	100	0.3166	0.4921	0.043	-0.2986	2	100	0.2563	0.3984	0.043	-0.1359
2	200	1.2076	0.9778	0.0235	-12.2431	2	200	1.002	0.8113	0.0235	-6.1016
4	0	-0.0033	-0.0495	0.1003	0.0242	4	0	-0.0059	-0.0897	0.1003	0.0242
4	1	-0.0033	-0.0452	0.0999	0.0232	4	1	-0.0062	-0.0864	0.0999	0.0238
4	10	-0.0008	-0.0062	0.0963	0.0134	4	10	-0.0071	-0.0563	0.0963	0.0195
4	50	0.0622	0.1679	0.0802	-0.0464	4	50	0.0288	0.0779	0.0802	-0.0076
4	100	0.2614	0.3865	0.06	-0.2877	4	100	0.1664	0.246	0.06	-0.125
4	200	1.0289	0.8114	0.0209	-12.2345	4	200	0.7317	0.5771	0.0209	-6.093
6	0	-0.0083	-0.084	0.1363	0.0363	6	0	-0.0142	-0.1432	0.1363	0.0363
6	1	-0.0084	-0.0801	0.1357	0.0353	6	1	-0.0148	-0.1405	0.1357	0.0359
6	10	-0.0073	-0.0455	0.1303	0.0253	6	10	-0.0186	-0.1164	0.1303	0.0314
6	50	0.0439	0.1089	0.1061	-0.0351	6	50	-0.0042	-0.0105	0.1061	0.0036
6	100	0.2146	0.3026	0.0759	-0.2774	6	100	0.0839	0.1183	0.0759	-0.1147
6	200	0.8869	0.6817	0.0173	-12.2273	6	200	0.4883	0.3753	0.0173	-6.0857

Table 1 (cntd.): Calibration Exercise

		Case III $p=0.2$, $\sigma_\pi=0.005$						Case IV $p=0.1$, $\sigma_\pi=0.005$			
γ	ρ	$E[dR]$	σ_R	r	σ_r	γ	ρ	$E[dR]$	σ_R	r	σ_r
0.5	0	0.0002	0.02	0.0346	0.0008	0.5	0	0.0002	0.0231	0.0346	0.0008
0.5	1	0.0002	0.0214	0.0346	0.0005	0.5	1	0.0002	0.0244	0.0346	0.0006
0.5	10	0.0008	0.0342	0.0337	-0.0019	0.5	10	0.0009	0.0363	0.0341	-0.0004
0.5	50	0.0077	0.0912	0.0138	-0.0167	0.5	50	0.0076	0.0899	0.0239	-0.007
0.5	100	0.0265	0.1648	-0.0599	-0.0668	0.5	100	0.0258	0.1608	-0.0133	-0.0311
0.5	200	0.1005	0.3204	-0.6858	-0.9684	0.5	200	0.0998	0.3182	-0.3268	-0.48
1	0	0.0003	0.0165	0.0442	0.0015	1	0	0.0003	0.0165	0.0442	0.0015
1	1	0.0003	0.0179	0.0442	0.0013	1	1	0.0003	0.0177	0.0442	0.0014
1	10	0.001	0.03	0.0431	-0.0012	1	10	0.0009	0.0287	0.0436	0.0004
1	50	0.0079	0.0848	0.0228	-0.0159	1	50	0.0073	0.0785	0.0329	-0.0063
1	100	0.0265	0.1569	-0.0516	-0.0661	1	100	0.0248	0.1467	-0.0049	-0.0304
1	200	0.1011	0.3139	-0.6788	-0.9677	1	200	0.0986	0.3062	-0.3198	-0.4793
2	0	0.0003	0.0103	0.0632	0.003	2	0	0.0002	0.0059	0.0632	0.003
2	1	0.0004	0.0115	0.0631	0.0028	2	1	0.0002	0.007	0.0631	0.0029
2	10	0.0011	0.0228	0.0619	0.0003	2	10	0.0008	0.0166	0.0623	0.0019
2	50	0.0081	0.0739	0.0405	-0.0144	2	50	0.0066	0.0609	0.0506	-0.0048
2	100	0.0266	0.1434	-0.0352	-0.0646	2	100	0.0231	0.1248	0.0115	-0.029
2	200	0.1024	0.3026	-0.6648	-0.9663	2	200	0.0973	0.2875	-0.3059	-0.478
4	0	0	0.0002	0.1003	0.0061	4	0	-0.0006	-0.0089	0.1003	0.0061
4	1	0.0001	0.0013	0.1002	0.0058	4	1	-0.0005	-0.008	0.1002	0.006
4	10	0.0009	0.0112	0.0984	0.0033	4	10	0	-0.0003	0.0989	0.0049
4	50	0.0081	0.0569	0.0751	-0.0116	4	50	0.0053	0.037	0.0852	-0.0019
4	100	0.0267	0.1225	-0.0031	-0.0619	4	100	0.0209	0.0959	0.0436	-0.0262
4	200	0.106	0.2854	-0.6378	-0.9639	4	200	0.098	0.2637	-0.2789	-0.4755
6	0	-0.0008	-0.0076	0.1363	0.0091	6	0	-0.0019	-0.0192	0.1363	0.0091
6	1	-0.0007	-0.0067	0.1361	0.0088	6	1	-0.0019	-0.0185	0.1361	0.009
6	10	0.0002	0.0022	0.1339	0.0063	6	10	-0.0014	-0.0119	0.1344	0.0079
6	50	0.0077	0.044	0.1085	-0.0087	6	50	0.0037	0.0213	0.1186	0.0009
6	100	0.0269	0.107	0.0279	-0.0593	6	100	0.0196	0.0779	0.0745	-0.0236
6	200	0.1104	0.273	-0.6119	-0.9617	6	200	0.1022	0.2527	-0.2529	-0.4733

This table reports the calibration exercise for the belief-dependent model parametrized as in section 6. The parameter γ denotes the coefficient of relative risk aversion while ρ denotes the coefficient of aversion to state uncertainty. The parameter p denotes the constant probability of shifting out of any state while σ_π denotes the dispersion of the posterior belief distribution $\pi = (\pi_1, ..., \pi_n)$ defined on the parameter space Θ. Negative entries for the volatility imply a negative covariance with consumption growth. The mean of the probability distributions π and f is $\mu = .0195$, where $f = (f_1, .., f_n)$ is the discretized normal distribution from which new drifts are chosen upon a shift occurs. The dispersion of f is $\sigma_\theta = .01$ (new drifts would be in the interval $[-.0005, .0395]$ with 95% probability) while the volatility of consumption growth is set equal to $\sigma_c = .0165$. Finally, the utility intertemporal discount rate is $\phi = .025$.

Table 2: ML Estimates of Jump Model

Parmeter	p	σ_c	σ_θ	μ
Estimate	0.1816	0.0134	0.0533	0.0195
t-stat.	1.368	9.267	2.056	-

This table reports the Maximum Likelihood annaulized estimates of the statistical model $\Delta \log(c_{t+1}) = \theta_t + \sigma_c \varepsilon_{t+1}$,

$$\theta_{t+1} = \begin{cases} \theta_t & \text{with prob.} \quad 1-p \\ \xi_{t+1} & \text{with prob.} \quad p \end{cases}$$

where $\xi_t \sim \mathcal{N}\left(\mu, \sigma_\theta^2\right)$. Estimates are obtained by discretizing the interval $[-0.4, 0.4]$ in $n = 200$ intervals and applying standard ML estimation methods for regime-switching models. Standard errors are (Newey-West) corrected for heteroskedasticity and autocorrelation. The parameter μ has been fixed to the long-run average of consumption growth to gain in estimation efficiency.

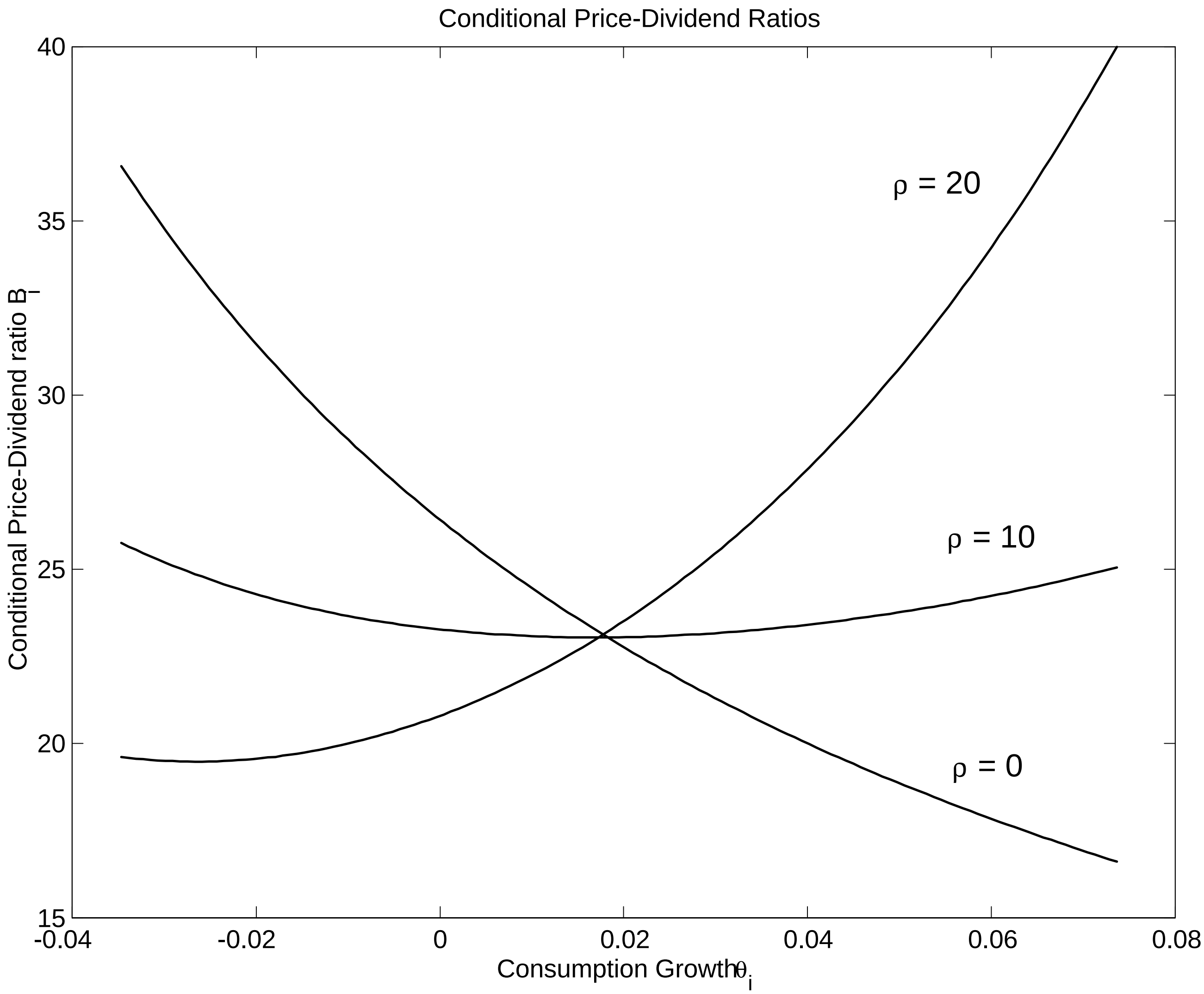
Table 3: Data and Model

	Data	Model
$E[R]$	0.0687	0.0526
σ_R	0.1596	0.1770
r	0.02^a	0.0122
σ_r	0.0226	0.0557

This table reports the ex-post mean excess stock returns, its volatility, and the volatility of the real-interest rate (assuming a constant risk premium). The table also reports the moments implied by the model fitted in Table 2, with utility parameters $\rho = 10$, $\phi = .04$ and $\gamma = 1.8$. The posterior distribution used to compute the unconditional expected returns is $\pi \sim N\left(\mu, \overline{\sigma}_\pi^2\right)$ where μ is the long-term average of consumption growth and $\overline{\sigma}_\pi = 0.0165$ was the average Root Mean Square Error obtained by simulating the process for beliefs for 5000 periods. All the moments have been the computed using the formulas in the text.

a: The value of the real rate has not been estimated. A value approximately of 2% is considered appropriate in the literature.

Figure 1: Conditional Price-Dividend Ratios



This figure plots the conditional price-dividend ratios obtained from the formula

$$B_i = \frac{1}{k(\theta_i)} \sum_{j=1}^{n} k(\theta_j)\, e_j \left(\phi I - \overline{\Lambda}\right) e_i'$$

where $k(\theta_i) = \overline{k} \times (\theta_1 + 1 + \theta_i)^{-\rho}$.

Figure 2: The Jump Distribution

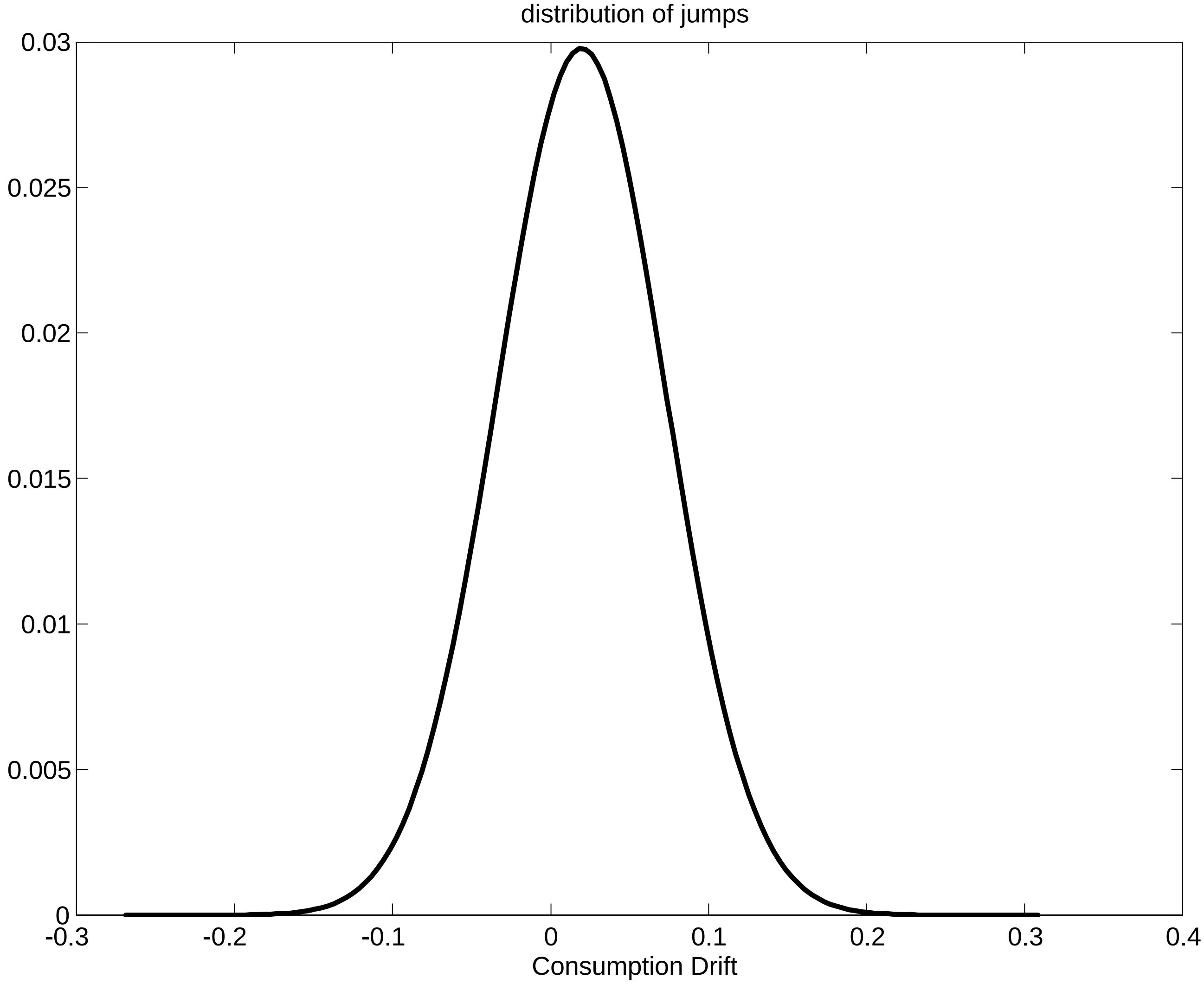
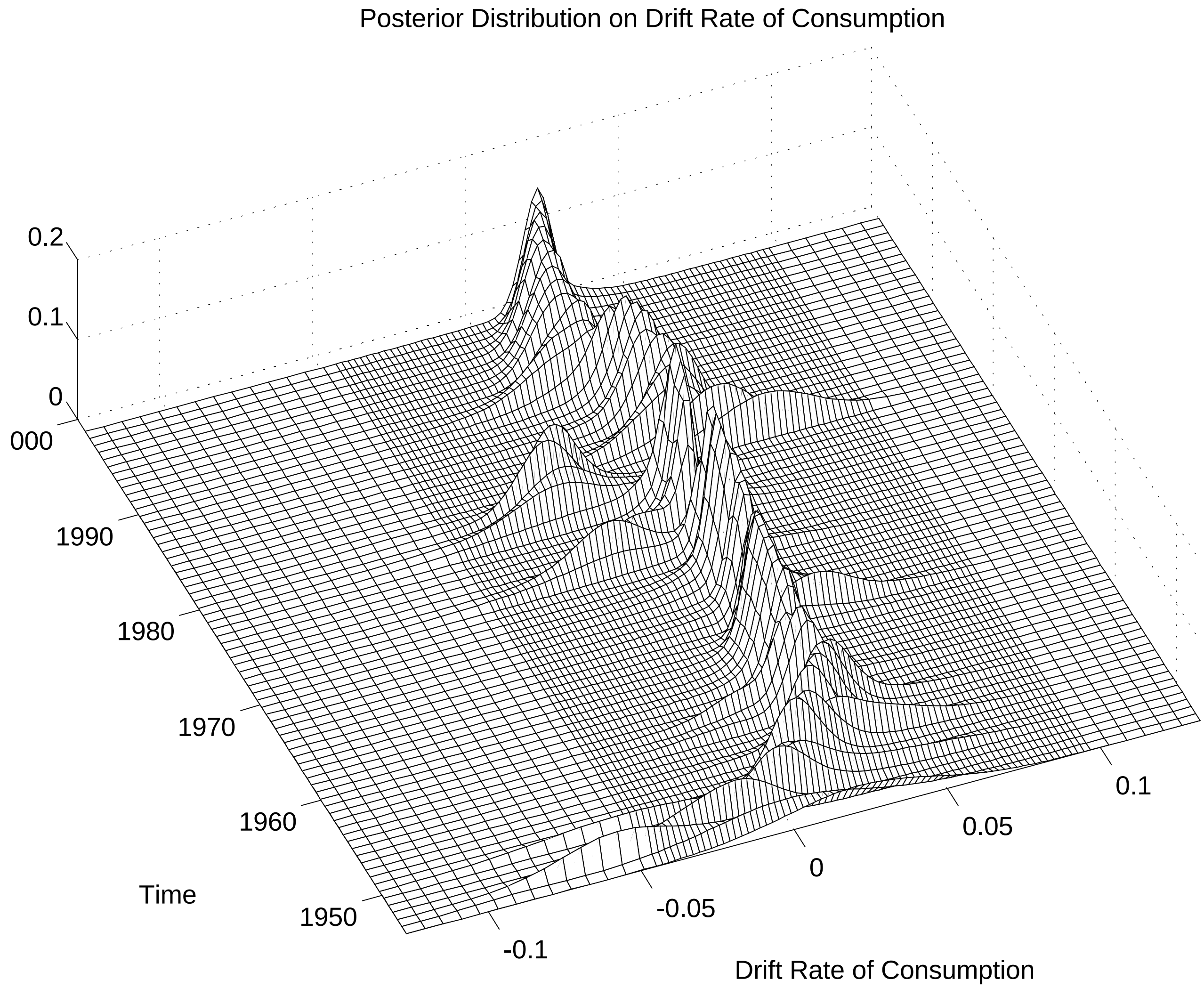


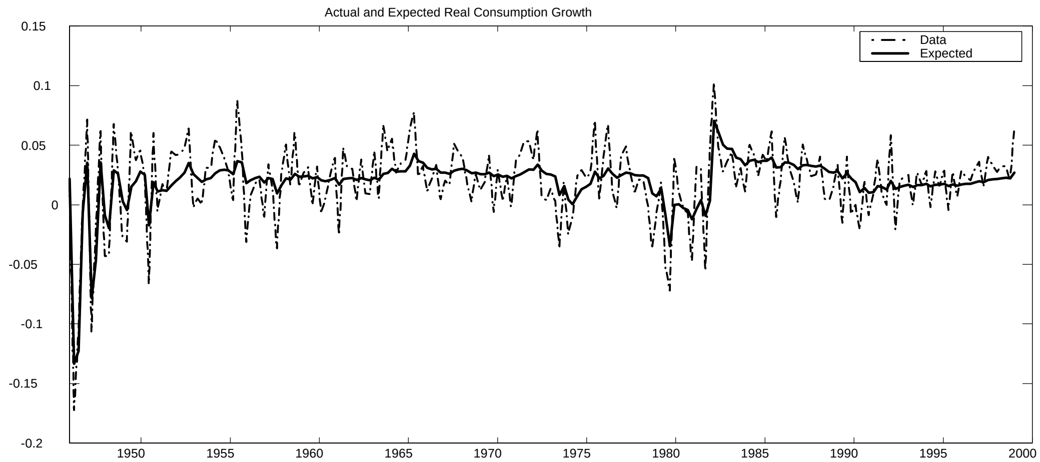
Figure 3: Time Series of Posterior Distribution on Consumption Drifts



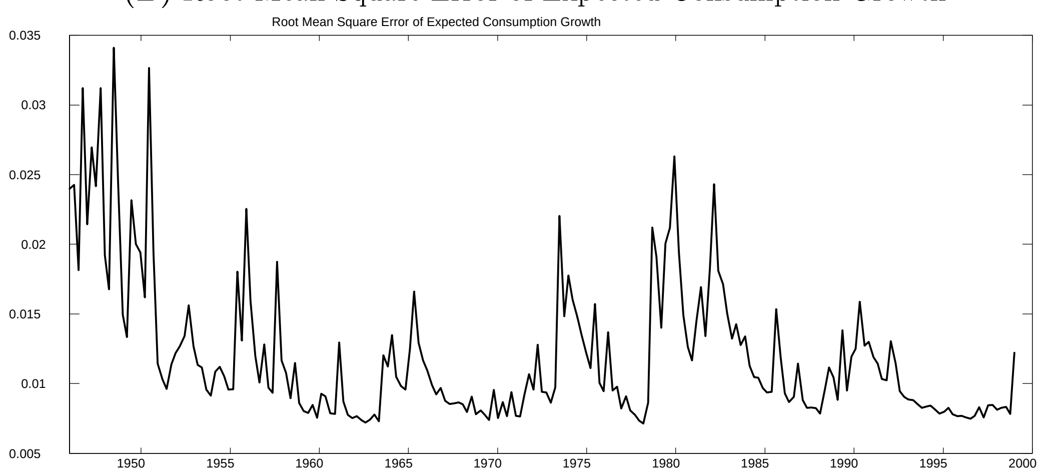
This figure plots the time series of the posterior distribution $\pi(t)$ on the drift rate of consumption θ_t computed using real consumption growth data from 1946 to 1999 and the MLE estimates in Table 2. Updating occurs by standard Bayes rule according to the formula $\pi_i(t+1) = \left\{ e^{-\frac{1}{2\sigma^2}(\Delta c(t+1)-\theta_i)^2} [\pi(t)\Lambda]_i \right\} / \left\{ \sum_{j=1}^{n} e^{-\frac{1}{2\sigma^2}(\Delta c(t+1)-\theta_j)^2} [\pi(t)\Lambda]_j \right\}$ where the transition matrix is given by $\Lambda = pf + (1-p)\mathbf{I}$ and $f = (f_1, ..., f_n)$ is the discrete version of a normal distribution. The number of points in the discretized grid is $n = 200$.

Figure 4: Expected Growth Rate of Consumption and its Root Mean Square Error

(A) Actual and Expected Real Consumption Growht

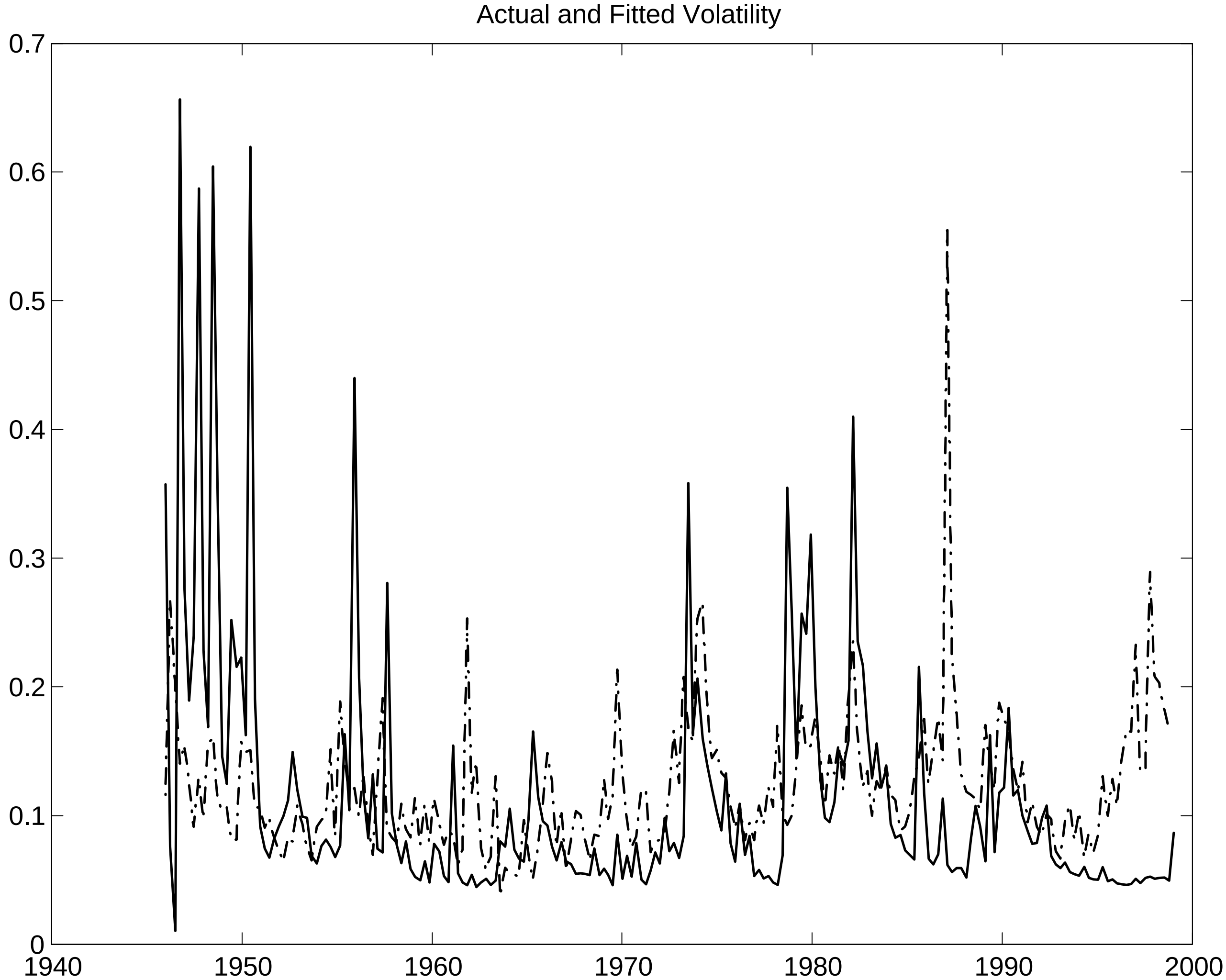


(B) Root Mean Square Error of Expected Consumption Growth



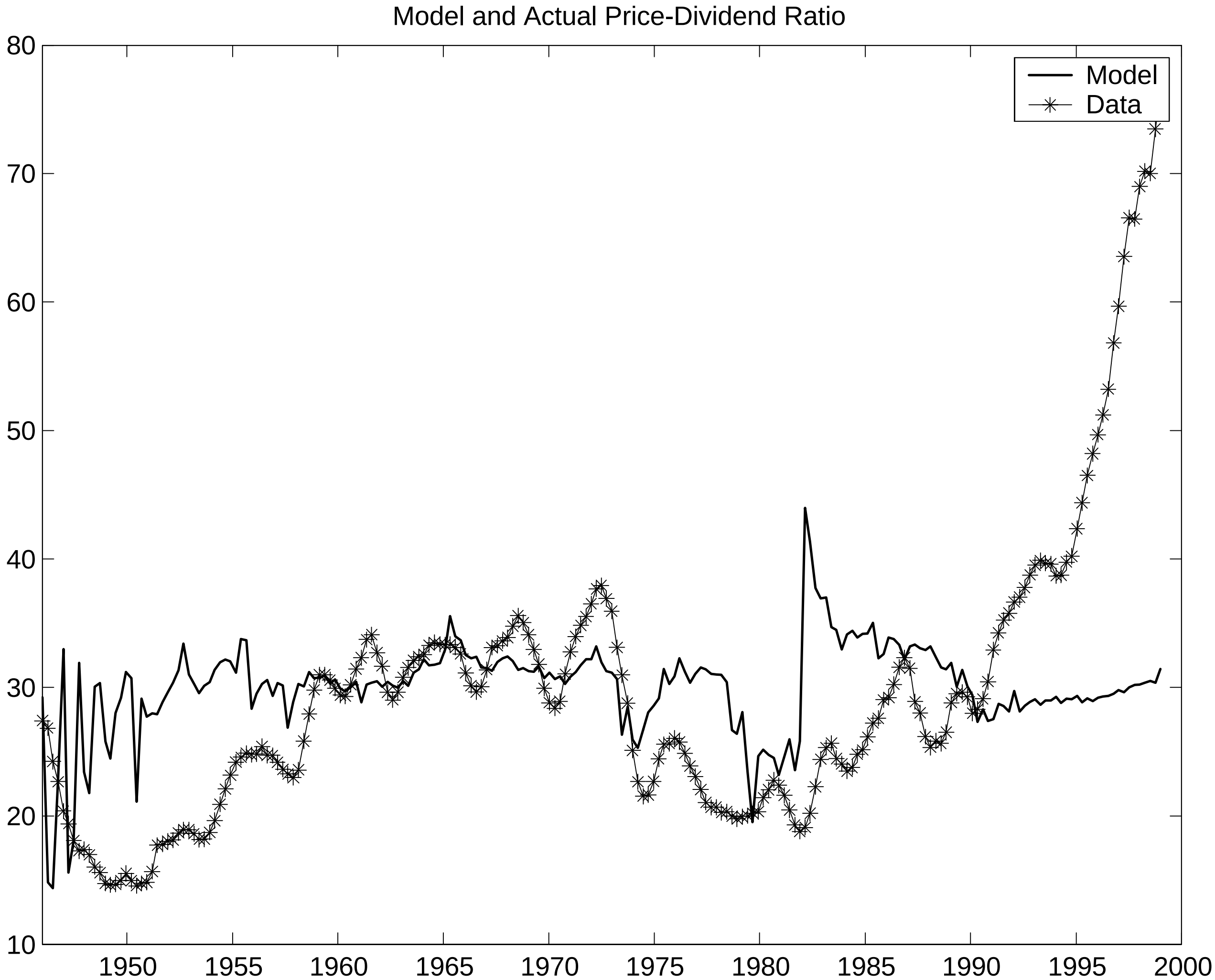
Panel (A) plots the time series of real consumption growth from 1946 to 1999 and the expected consumption growth obtained from the fitted posterior probabilities $\pi(t)$ as described in text and in Figure 2. Panel (B) plots the time series of the Root Mean Square Error of the expected consumption growth.

Figure 5: Actual and Model-Implied Volatility



This figure reports the plot of the conditional volatility implied by the model (solid line) and the ex-post integrated volatility computed from CRSP-data on stock returns (dash-dotted line). The conditional volatility implied by the model is computed applying the formula provided in the text and at every t it only depends on the posterior distribution $\pi(t)$ in Figure 2.

Figure 6: Actual and Model-Implied Price-Dividend Ratios



This figure plots the time series of the price-dividend ratio obtained from CRSP-data and the model-implied price dividend ratio. The latter is only a function of past consumption and it is given by the formula $P(t)/D(t) = \sum_{i=1}^{n} \pi_i^*(t) B_i$.