Modeling the Long Run: Valuation in Dynamic Stochastic Economies

Lars Peter Hansen*

June 6, 2008

Abstract

I explore the value implications of economic models by highlighting what features persist in the long-term. I accomplish this by developing a decomposition of valuation dynamics (DVD). I use it to distinguish components of an underlying economic model that influence values over long horizons from components that impact only the short run. I apply my approach to study example economies from the asset pricing literature, and I speculate about the long-term valuation implications of a broader class of economic models. I develop a perturbation method to quantify the role of parameter sensitivity and to impute long-term risk prices. A DVD is enabled by constructing operators indexed by the elapsed time between the date of pricing and the date of the future payoff (*i.e.* the future realization of a consumption claim). Thus formulated, methods from applied mathematics permit me to characterize valuation behavior as the time between price determination and payoff realization becomes large. An outcome of this analysis is the construction of a *multiplicative* martingale component of a process that is used to *represent* valuation in a dynamic economy with stochastic growth. I contrast the differences in the applicability between this *multiplicative* martingale method and an *additive* martingale method familiar from time series analysis that is currently used to identify shocks with long-run economic consequences.

1 Introduction

In this paper I propose to augment the toolkit for modeling economic dynamics and econometric implications with methods that reveal the important economic components of long-

*Prepared for the Fisher-Schultz Lecture at the 2006 European Meetings of the Econometric Society. I appreciate helpful comments by Jaroslav Borovicka, Buz Brock, Valentin Haddad, Mark Hendricks, Jose Scheinkman and Neng Wang. Portions of this work are joint with John Heaton, Nan Li and Jose Scheinkman, and very much influenced by related work I have done with Xiaohong Chen and Tom Sargent.

term valuation in economies with stochastic growth. These tools enable informative decompositions of a model's dynamic implications for valuation. They are the outgrowth of my observation and participation in an empirical literature that aims to understand the low frequency links between financial market indicators and macroeconomic aggregates.

Current dynamic models that relate macroeconomics and asset pricing are constructed from an amalgam of assumptions about preferences, (such as risk aversion or habit persistence, etc) technology (productivity of capital or adjustment costs to investment), and exposure to unforeseen shocks. Some of these components have more transitory effects while others have a lasting impact. In part my aim is to illuminate the roles of these model ingredients by presenting a structure that features long run implications for value. By *value* I mean either market or shadow prices of physical, financial or even hypothetical assets.

These methods are designed to address three questions:

- What are the long-term value implications of nonlinear economic models with stochastic growth?
- To which components of the *uncertainty* are long-run valuations most sensitive?
- What kind of hypothetical changes in preferences and technology have the most potent impact on the long run? What changes are transient?

Although aspects of these questions have been studied using log-linear models and log-linear approximations around a growth trajectory, the methods I describe offer a different vantage point. These methods are designed for the study of valuation in the presence of stochastic inputs that have long-run consequences. While the methods can exploit any linearity, by design they can accommodate nonlinearity as well. In this paper I will develop these tools, as well as describe their usefulness at addressing these three economic questions. I will draw upon some diverse results from stochastic process theory and time series analysis, although I will use these results in novel ways.

There are a variety of reasons to be interested in the first question. When we build dynamic economic models, we typically specify transitional dynamics over a unit of time for discrete-time models or an instant of time for continuous time models. Long-run implications are encoded in such specifications; but they can be hard to decipher, particularly in nonlinear stochastic models. I explore methods that describe long-run limiting behavior, a concept which I will define formally. I see two reasons why this is important. First some economic inputs are more credible when they target low frequency behavior. Second these inputs may be essential for meaningful long-run extrapolation of value. Nonparametric statistical alternatives suffer because of limited empirical evidence on the long-run behavior of macroeconomic aggregates and financial cash flows.

Recent empirical research in macro-finance has highlighted economic modeling successes at low frequencies. After all, models are approximations; and applied economics necessarily employs models that are misspecified along some dimensions. In this context, then, I hope these methods for extracting long-term implications from a dynamic stochastic model will be welcome additional research tools. Specifically, I will show how to deconstruct a dynamic stochastic equilibrium implied by a model, revealing what features dominate valuation over long time horizons. Conversely, I will formalize the notion of transient contributions to valuation. These tools will help to formalize long-term approximation and to understand better what proposed model fixups do to long-term implications.

This leads me to the second question. Many researchers study valuation under uncertainty by risk prices, and through them, the equilibrium risk-return tradeoff. In equilibrium, expected returns change in response to shifts in the exposure to various components of macroeconomic risk. The tradeoff is typically depicted over a single period in a discrete-time model or over an instant of time in a continuous time model. I will extend the log-linear analysis in Hansen et al. (2008) by deriving the long-run counterpart to this familiar exercise. Specifically, I will perform a sensitivity analysis that recovers prices of exposure to the component parts of long-run (growth-rate) risk. I will define formally risk prices in nonlinear models as they depend on the investment horizon, and in particular characterized their limiting behavior. These limits are basic inputs into the study of the term structure of risk prices.

These same methods facilitate long-run welfare comparisons in explicitly dynamic and stochastic environments.

Finally, consider the third question. Many components of a dynamic stochastic equilibrium model can contribute to value in the long run. Changing some of these components will have a more potent impact than others. To determine this, we could perform value calculations for an entire family of models indexed by the model ingredients. When this is not practical, an alternative is to explore local changes in the economic environment. We may assess, for example, how modification in the intertemporal preferences of investors alter long term risk prices and interest rates. The resulting derivatives quantify these and other impacts and can inform statistical investigations.

1.1 Game plan

My game plan for the technical development in this paper is as follows:

i) *Underlying Markov structure* (section 2): I pose a Markov process in continuous time. The continuous-time specification simplifies some of our characterizations, but it is not essential to our analysis. I build processes that grow over time by accumulating the impact of the Markov state and shock history. The result will be functionals, additive or multiplicative. Additive functionals are typically logarithms of macro or financial variables and multiplicative functionals are levels of these same time series.

ii) *Decomposition of Additive functionals* (section 3): An additive functional accumulates the impact of a Markov state over time via summation or integration. I produce a familiar decomposition of an additive functional Y into three components:

$$\begin{array}{ccccccc} Y_t & = & \nu t & + & \hat{Y}_t & -g(X_t) + g(X_0) & \\ & & \uparrow & & \uparrow & \uparrow & . \\ & & \textbf{linear trend} & & \textbf{martingale} & \textbf{stationary} & \end{array} \qquad (1)$$

This decomposition nests decompositions from the macroeconomic time series literature and the stochastic process literature on central limit approximation. This decomposition identifies permanent shocks as increments to the martingale component. Such shocks dominate the stochastic component of growth over long-horizons and reflect exposure to risk that have long-term consequences for valuation.

iii) *Multiplicative processes and valuation* (section 4): I build multiplicative functionals by exponentiating additive ones. Thus I work with levels instead of logarithms as in the case of additive functionals. Alternative multiplicative functionals can capture stochastic discounting or stochastic growth. The stochastic discount factor processes are deduced by economic models and designed to capture both pure discount effects and risk adjustments. The multiplicative construction reflects the effect of compounding over intervals of time. Growth fluctuations are modeled by accumulating local stochastic growth exponentially over intervals of time. I study valuation in conjunction with growth by constructing families of operators indexed by the valuation horizon. The operators will map the transient components to payoffs, cash flows or Markov claims to a numeraire consumption good. As special cases I will study growth abstracting from valuation and valuation abstracting from growth. I use multiplicative functionals constructed from the underlying Markov process to represent the previously described family of operators.

iv) *Long-run approximation* (sections 5 and 6): I measure long-run growth and the associated value decay through the construction of *principal eigenvalues* and *principal eigenfunctions*. I use an extended version of Perron-Frobenius theory to establish a multiplicative analog to decomposition (1):

$$M_t = \underset{\textbf{exponential trend}}{\underset{\uparrow}{\exp(\rho t)}} \quad \underset{\textbf{martingale}}{\underset{\uparrow}{\hat{M}_t}} \quad \underset{\textbf{transient}}{\underset{\uparrow}{\left[\frac{\hat{e}(X_t)}{\hat{e}(X_0)}\right]}} . \tag{2}$$

where M is the exponential of an additive functional and is chosen to represent valuation in the presence of stochastic growth. This gives a decomposition of the valuation dynamics (DVD). Although superficially similar, this factorization is distinct from (1) because the exponential of a martingale is not itself a martingale. I will use this factorization to obtain a convenient characterization of long-run behavior in valuation and to formally define permanent and transitory model components. In my applications, M's will be constructed from explicit economic models and hypothetical changes in stochastic growth trajectories. Prior to developing a DVD from a general perspective in section 6, I explore the special case of a finite-state Markov chain in section 5.

v) *Sensitivity and long-run pricing* (section 7): Of special interest is how the long-run attributes of valuation change when we alter the growth processes or when we alter the stochastic discount factor used to represent valuation. I show formally how to conduct a sensitivity analysis with two applications in mind. We consider changes in the risk exposure of hypothetical growth processes which give rise to long-run risk prices. I also explore how long run values and rates of return are predicted to change as the attributes of the economic environment are modified.

vi) *Applications to the asset-pricing literature* (section 8): I apply the methods to study some existing models of asset pricing and to compare their long-run implications. While the methods are much more generally applicable, I feature some specifications for which quasi-analytical characterizations are possible. I show when specific features of asset pricing models have transient implications.

The ideas developed in this paper have important antecedents from a variety of literatures. Rather than provide a comprehensive literature review at the outset, I will point out important prior contributions as I develop results.

2 Probabilistic specification

While there are variety of ways to introduce nonlinearity into time series models, for tractability we concentrate on Markovian models. For convenience, we will feature continuous time models with their sharp distinctions between small shocks modeled as Brownian increments and large shocks modeled as Poisson jumps. Let X denote the underlying Markov process summarizing the state of an economy. We will use this process as a building block in our construction of economic relations.

2.1 Underlying Markov process

I consider a Markov process X defined on a state space $\mathcal{E}$. Suppose that this process can be decomposed into two components: $X^c + X^d$. The process X is right continuous with left limits. With this in mind I define:

$$X_{t-} = \lim_{u \downarrow 0} X_{t-u}.$$

I depict local evolution of X^c as:

$$dX_t^c = \mu(X_{t-})dt + \sigma(X_{t-})dW_t$$

where W is a possibly multivariate standard Brownian motion. The process X^d is a jump process. This process is modeled using a finite conditional measure $\eta(dx^*|x)$ where $\int \eta(dx^*|X_{t-})$ is the jump intensity. That is for small ϵ, $\epsilon \int \eta(dx^*|X_{t-})$ is the approximate probability that there will be a jump. The conditional measure $\eta(dx^*|x)$ scaled by the jump intensity is the probability distribution for the jump conditioned on a jump occurring. Thus the entire Markov process is parameterized by (μ, σ, η).

I will often think of the process X as stationary, but strictly speaking this is not necessary. As we will see next, nonstationary processes will be constructed from X.

2.2 Convenient functions of the Markov process

Consider the frictionless asset pricing paradigm. Asset prices are depicted using a stochastic discount factor process S. Such a process cannot be freely specified. Instead restrictions are implied by the ability of investors to trade at intermediate dates. The use of a Markov assumption in conjunction with valuation leads us naturally to the study of multiplicative functionals or their additive counterparts formed by taking logarithms. I will also use multiplicative functionals to depict growth components of cash flows or consumption processes.

An additive functional Y is constructed from the underlying Markov process such that that $Y_{t+\tau} - Y_t = \phi_\tau(X_u \;\text{ for }\; t < u \leq t+\tau)$ for any $t \geq 0$ and any $\tau \geq 0$. For convenience, it is initialized at $Y_0 = 0$. Notice that what I call an additive functional is actually a stochastic process defined for all $t \geq 0$. Even if the underlying Markov process is stationary, an additive functional will typically not be. Instead it will have increments that are stationary and hence the Y process can display arithmetic growth (or decay) even when the underlying process X does not. An additive functional can be normally distributed, but I will also interested other specifications. Conveniently, the sum of two additive functionals is additive.

I consider a family of such functionals parameterized by (β, ξ, λ) where:

i) $\beta : \mathcal{E} \to \mathbb{R}$ and $\int_0^t \beta(X_u)du < \infty$ for every positive t;

ii) $\xi : \mathcal{E} \to \mathbb{R}^m$ and $\int_0^t |\xi(X_u)|^2 du < \infty$ for every positive t;

iii) $\lambda : \mathcal{E} \times \mathcal{E} \to \mathbb{R}$, $\lambda(x, x) = 0$.

$$Y_t = \int_0^t \beta(X_u)du + \int_0^t \xi(X_{u-}) \cdot dW_u + \sum_{0 \le u \le t} \lambda(X_u, X_{u-}) \tag{3}$$

The additive functional Y in (3) has three components, each of which accumulates linearly over time. The first component is a simple integral, $\int_0^t \beta(X_u)du$, and it as a consequence it is locally predictable. The second component is a stochastic integral, $\int_0^t \xi(X_u) \cdot dW_u$, and it reflects how "small shocks" alter the functional Y. These small shocks are modeled as Brownian increments. This component is a local martingale, and I will feature cases in which it is a martingale. Recall that the best forecast of the future value of a martingale is the current value of the martingale. The third component shows how jumps in the underlying process X induce jumps in the additive functional. If X jumps at date t, Y also jumps at date t by the amount $\lambda(X_t, X_{t-})$. The term $\sum_{0 \le u \le t} \lambda(X_u, X_{u-})$ thus reflect the impact of "large shocks". This component is not necessarily a martingale because the jumps may have a predictable component. The integral

$$\tilde{\beta}(x) = \int_{\mathcal{E}} \lambda(x^*, x)\eta(dx^*|x). \tag{4}$$

captures this predictability locally. Integrating $\tilde{\beta}$ over time and subtracting it from the jump component of Y gives an additive local martingale:

$$\sum_{0 \le u \le t} \lambda(X_u, X_{u-}) - \int_0^t \tilde{\beta}(X_u)du.$$

I will be primarily interested in specifications of λ for which this constructed process is a martingale. In summary, an additive functionals grows or decays stochastically in a linear way. Its dynamic evolution can reflect the impact of small shocks represented as a state-dependent weighting of a Brownian increment and the impact of large shocks represented by a possibly nonlinear response to jumps in the underlying process X.

The logarithms of economic aggregates can be conveniently represented as additive functionals as can the logarithms of stochastic discount factors used to represent economic values.[1] I next consider the level counterparts to such functionals.

While a multiplicative functional can be defined more generally, I will consider ones that are constructed as exponentials of additive functionals: $M = \exp(Y)$. Thus the ratio $M_{t+\tau}/M_t$ is constructed as a function of X_u for $t < u \leq t + \tau$.[2] Multiplicative functionals are necessarily initialized at unity.

Even when X is stationary, a multiplicative process can grow (or decay) stochastically in an exponential fashion. Although its logarithm will have stationary increments, these increments are not restricted to have a zero mean.

3 Log-linearity and long-run restrictions

A standard tool for analyzing dynamic economic models is to characterize stochastic steady state relations. These steady states are obtained by deducing a scaling process or processes that capture growth components common to many time series. Similarly, the econometric literature on cointegration is typically grounded in log-linear implications that restrict variables to grow together. Error-correction specifications seek to allow for flexible transient dynamics while enforcing long-run implications. Economics is used to inform us as to which time series move together. See Engle and Granger (1987).[3] Relatedly, Blanchard and Quah (1989) and many others use long-run implications to identify shocks. Supply or technology shocks broadly conceived are the only ones that influence output in the long run. These methods aim to measure the potency of shocks while permitting short-run dynamics.

Prior to studying of multiplicative functionals, consider the decomposition of an additive functional. Such a process can be built by taking logarithms of the multiplicative functional, a common transformation in economics. I now describe such a decomposition. While there are alternative ways to decompose time series, what follows is closest to what I will be

[1] For economic aggregates, it is necessary to subtract of the date zero logarithms in order that $Y_0 = 0$.

[2] This latter implication gives the key ingredient of a more general definition of a multiplicative functional.

[3] Interestingly, Box and Tiao (1977) anticipated the potentially important notion of long run co-movement in their method of extracting canonical components of multivariate time series.

interested in. An additive functional can be decomposed into three components:

$$\begin{array}{cccccc} Y_t & = & \nu t & + & \hat{Y}_t & -g(X_t) + g(X_0) \\ & & \uparrow & & \uparrow & \uparrow \\ & & \textbf{linear trend} & & \textbf{martingale} & \textbf{stationary} \end{array} . \tag{5}$$

This decomposition gives a way to identify shocks with "permanent" consequences. Recall that the best forecast of the future values of a martingale is the current value of that martingale. Thus permanent shocks are reflected in the increment to the martingale component of (5). It helps to isolate the exposure of economic time series to macroeconomic risk that dominates the fluctuation of Y over long time horizons.

The remainder of this section is organized as follows. I first verify formally the martingale property $\hat{Y}$, and then I give operational ways to construct this decomposition. I end the section with two examples. The first example give the continuous-time counterpart to this decomposition for a model with linear stochastic dynamics. This example illustrates the construction of permanent shocks that is typically used in conjunction with vector autoregressive methods. The second example introduces stochastic volatility. This example allows for volatility to fluctuate over time in a manner that can be highly persistent. Thus a particular form of nonlinearity is introduced into the analysis, a form that has received considerable attention in both the macroeconomics and asset-pricing literatures.

My first formal statement of decomposition (5) is:

Theorem 3.1. *Suppose that Y is an additive functional with increments that have finite second moments. In addition, suppose that*

$$\lim_{\tau \to \infty} \frac{1}{\tau} E\left(Y_t | X_0 = x\right) = \nu,$$

and

$$\lim_{\tau \to \infty} E\left(Y_\tau - \nu\tau | X_0 = x\right) = g(x),$$

where the convergence is in mean square. Then Y can be represented as:

$$Y_t = \nu t + \hat{Y}_t - g(X_t) + g(X_0). \tag{6}$$

where $\{\hat{Y}_t\}$ is an additive martingale.

Proof. Let $Y_t^* = Y_t - \nu t$. Let $\mathcal{F}_t$ be the sigma algebra generated by by the X process between time 0 and time t. As a consequence of the Law of Iterated Expectations and the mean-square convergence,

$$
\begin{aligned}
g(X_t) + Y_t^* &= \lim_{\tau\to\infty} E\left(Y_{t+\tau}^* - Y_t^* | X_t\right) + Y_t^* \\
&= \lim_{\tau\to\infty} E\left(Y_{t+\tau}^* | \mathcal{F}_t\right) \\
&= \lim_{\tau\to\infty} E\left[E\left(Y_{t+\tau}^* - Y_{t+\epsilon}^* | \mathcal{F}_{t+\epsilon}\right) + Y_{t+\epsilon}^* | \mathcal{F}_t\right] \\
&= E\left[g(X_{t+\epsilon}) + Y_{t+\epsilon}^* | \mathcal{F}_t\right]
\end{aligned}
$$

Thus $\{Y_t^* + g(X_t)\}$ is a martingale with initial value $g(X_0)$. After subtracting $g(X_0)$,

$$
\hat{Y}_t = Y_t^* + g(X_t) - g(X_0)
$$

remains a martingale, but it has initial value zero as required for an additive functional. □

I next show how to use the local evolution of the additive functional to construct the components of this decomposition. Recall the representation given in (3):

$$
Y_t = \int_0^t \beta(X_u) du + \int_0^t \xi(X_u) \cdot dW_u + \sum_{0 \le u \le t} \lambda(X_u, X_{u-}),
$$

and the construction of $\tilde{\beta}$ in formula (4. Then

$$
\tilde{Y}_t = \int_0^t \xi(X_u) \cdot dW_u + \sum_{0 \le u \le t} \lambda(X_u, X_{u-}) - \int_0^t \tilde{\beta}(X_u) du
$$

is a local martingale. In what follows let

$$
\hat{\beta} = \beta + \tilde{\beta}.
$$

Theorem 3.2. *Suppose*

i) X is a stationary, ergodic Markov process;

ii) $\tilde{Y}$ is a square integrable martingale;

iii) $\hat{\beta}(X_t)$ has a finite second moment;

iv) There is a solution g to

$$
g(x) = \int_0^\infty E\left(\hat{\beta}(X_t) - E\left[\hat{\beta}(X_t)\right] | X_0 = x\right);
$$

Then $\hat{Y}$ given by $Y_t - \nu t + g(X_t) - g(X_0)$ is a martingale with stationary, square integrable increments with $\nu = E\left[\tilde{\beta}(X_t)\right]$.

This theorem gives an algorithm for computing ν from the local evolution of Y and the stationary distribution for X. It remains to compute the function g of the Markov state. Since $\hat{Y}$ is a martingale, its increments should not be predictable. As a consequence,

$$\hat{\beta}(x) - \nu + \lim_{t \downarrow 0} \frac{1}{t} E\left[g(X_t) - g(x) | X_0 = x\right] = 0,$$

which gives an equation for g that depends on the local evolution of X. The calculation of the expected time derivative:

$$\lim_{t \downarrow 0} \frac{1}{t} E\left[g(X_t) - g(x) | X_0 = x\right] = \mathbb{A} g(x)$$

defines the so called generator $\mathbb{A}$ for the Markov process. Specifying generator $\mathbb{A}$ is one way to represent the transition dynamics for Markov process. In the case of a multivariate diffusion, this equation is known to be a second-order differential equation as an implication of Ito's Lemma. There are well known extensions to accommodate jumps. Using the generator, the function g satisfies

$$\mathbb{A} g = \nu - \hat{\beta}. \tag{7}$$

For the diffusion model, this leads to solving:

$$\frac{\partial g(x)}{\partial x} \cdot \mu(x) + \frac{1}{2} \text{trace}\left[\sigma(x)\sigma(x)' \frac{\partial^2 g(x)}{\partial x \partial x'}\right] = \nu - \hat{\beta}(x). \tag{8}$$

The local evolution of the martingale $\hat{Y}$ is given by:

$$\xi(X_t) dW_t + \left[\frac{\partial g(X_t)}{\partial x}\right]' \sigma(X_t) dW_t,$$

where the first term is contributed by the local evolution of $\tilde{Y}$ and the second term by the local evolution of $g(X)$.

More generally, to obtain a solution g to a long-run forecasting problem, it suffices to solve equation (7) depicted using the local evolution of the Markov process. Much is known about such an equation. As argued by Bhattacharya (1982) and Hansen and Scheinkman (1995), when X is ergodic this equation has at most one solution. When X is exponentially ergodic, there always exists a solution.[4]

[4]These references suppose that X is stationary. Hansen and Scheinkman (1995) use an L^2 notion of

Following Gordin (1969), by extracting a martingale we can produce a more refined analysis. Specifically, an implication of the Martingale Central Limit Theorem is that

$$\lim_{t\to\infty} \frac{1}{\sqrt{t}}(Y_t - \nu t) \approx \frac{1}{\sqrt{t}}\hat{Y}_t \Rightarrow \text{normal}$$

is normally distributed with mean zero.[5] In addition to central limit approximation, there are other important applications of this decomposition. For linear time series, Beveridge and Nelson (1981) and others use this decomposition to identify $\hat{Y}_t$ as the permanent component of a time series. When there are multiple additive functionals under consideration and they have common martingale components of lower dimension, then one obtains the cointegration model of Engle and Granger (1987). Linear combinations of the vector of additive functionals will have a time trend and martingale component that are identically zero. Blanchard and Quah (1989) use such a decomposition to identify permanent shocks. The martingale increments are innovations to supply or technology shocks.

I now consider some examples.

Example 3.3. *Suppose that*

$$\begin{aligned} dX_t &= AX_t dt + BdW_t, \\ dY_t &= \nu dt + HX_t dt + FdW_t \end{aligned}$$

where A has eigenvalues with strictly negative real parts and W is multivariate Brownian standard motion. In this example $\hat{\beta}(x) = \nu + Hx$, and g satisfies the partial differential equation:

$$\frac{\partial g(x)}{\partial x} \cdot (Ax) + \frac{1}{2}\text{trace}\left[BB'\frac{\partial^2 g(x)}{\partial x \partial x'}\right] = -Hx$$

which is a special case of (8). This equation has a linear solution:

$$g(x) = -HA^{-1}x$$

The surprise movement or "innovation" to $g(X_t)$ is $-HA^{-1}BdW_t$. Thus in this example,

$$\hat{Y}_t = \int_0^t \left(F - HA^{-1}B\right) dW_u$$

exponential ergodicity using the implied stationary distribution of X as a measure. Bhattacharya (1982) establishes a functional counterpart to the central limit theorem using these methods. In both cases strong dependence in X can be tolerated provided there exists a solution to (7).

[5]See Billingsley (1961) for the discrete-time martingale central limit. Moreover, there are well known functional extensions of this result.

is the martingale component.

Next I consider a model with stochastic volatility.

Example 3.4. *Suppose that X and Y evolve according to:*

$$\begin{aligned} dX_t^{[1]} &= A_1 X_t^{[1]} dt + \sqrt{X_t^{[2]}} B_1 dW_t, \\ dX_t^{[2]} &= A_2 (X_t^{[2]} - 1) dt + \sqrt{X_t^{[2]}} B_2 dW_t \\ dY_t &= \nu dt + H_1 X_t^{[1]} dt + H_2 (X_t^{[2]} - 1) dt + \sqrt{X_t^{[2]}} F dW_t. \end{aligned}$$

Both $X^{[2]}$ and Y are scalar processes. The process $X^{[2]}$ is an example of a Feller square root process, which I use to model the temporal dependence in volatility. I restrict $B_1 {B_2}' = 0$ implying that $X^{[1]}$ and $X^{[2]}$ are conditionally uncorrelated. The matrix A_1 has eigenvalues with strictly negative real parts and A_2 is negative. Moreover, to prevent zero from being attained by $X^{[2]}$, I assume that $A_2 + \frac{1}{2}|B_2|^2 < 0$. I have parameterized this process to have mean one when initialized in its stationary distribution, which for my purposes is essentially a normalization. In this example g solves the partial differential equation:

$$\frac{\partial g(x_1, x_2)}{\partial x} \cdot \begin{bmatrix} A_1 x_1 \\ A_2 (x_2 - 1) \end{bmatrix} + \frac{x_2}{2} \text{trace} \left(\begin{bmatrix} B_1 {B_1}' & 0 \\ 0 & |B_2|^2 \end{bmatrix} \frac{\partial^2 g(x_1, x_2)}{\partial x \partial x'} \right) = -H_1 x_1 - H_2 (x_2 - 1),$$

which is a special case of (8). The solution is:

$$g(x_1, x_2) = -H_1 (A_1)^{-1} x_1 - H_2 (A_2)^{-1} (x_2 - 1).$$

The local innovation in $g(X_t)$ is $\sqrt{X_t^{[2]}}[-H_1(A_1)^{-1} B_1 - H_2 (A_2)^{-1} B_2] dW_t$. Thus in this example the martingale component for Y is given by:

$$\hat{Y}_t = \int_0^t \sqrt{X_u^{[2]}} \left[F - H_1 (A_1)^{-1} B_1 - H_2 (A_2)^{-1} B_2 \right] dW_u.$$

This example has the same structure as example 3.3 except that the Brownian motion shocks are scaled by $\sqrt{X_t^{[2]}}$ to induce volatility that varies over time. While example 3.3 is fully linear, example 3.4 introduces a nonlinear volatility factor. More generally, additive functionals do not have to be linear functions of the Markov state or linear functions of Brownian increments. Nonlinearity can be built into the drifts (conditional means) or the diffusion coefficients (conditional variances). Under these more general constructions, the function g used to represent the transient component will not be a linear function of the

Markov state.[6]

Even when such nonlinearity is introduced, conveniently the sum of two additive functionals is additive and the sum of the martingale decompositions is the martingale decomposition for the sum of the additive functionals provided the component martingale differences are constructed using a common information structure.

For the purposes of valuation, in what follows I will use multiplicative functionals. Such functionals can be represented conveniently as the logarithms of additive functionals. One strategy at our disposal is to decompose and then exponentiate. Thus for $M_t = \exp(Y_t)$:

$$M_t = \exp(\nu t) \exp\left(\hat{Y}_t\right) \frac{\exp[-g(X_t)]}{\exp[-g(X_0)]}$$

for the decomposition given in (6). While such a factorization is sometimes of value, for the purposes of my analysis, it is important that I construct an alternative factorization. The exponential of a martingale is not a martingale. If the process is lognormal, then this assumption can be used to transform $\exp(\hat{Y})$ into a martingale by scaling it by an exponential function of time. Later I will illustrate this outcome. More generally, I will construct an alternative multiplicative decomposition that will be of direct use.[7]

Prior to our development of an alternative decomposition, I discuss some limiting characterizations that will interest us.

4 Limiting characterizations of stochastic growth or discounting

In representing an additive functional (Y), either β or its counterpart $\hat{\beta}$ that adjusts for jump predictability, could be viewed as a state-dependent *local* growth rate or discount rate for the exponentiatial of this functional ($\exp(Y)$). I now explore the relation between the average local growth rate and a long-run counterpart to this growth rate. As I will show, compounding has nontrivial consequences when the local growth rate is state-dependent. Here I am interpreting growth liberally so as to include discounting as well. For instance, what I develop in this section is also germane to the study of long-term implications of compounding of short-term interest rates, a topic that has been explored in the literature on climate policy (see for instance Newell and Pizer (2003, 2004)) and in the literature on the long-term behavior of stochastic discount factor (see for instance Alvarez and Jermann

[6]The Markov assumption is also not necessary for such a decomposition.

[7]While this additive decomposition is linked to a Central Limit Theorem, this alternative decomposition has much closer ties to the theory of large deviations rather than central limit theory.

(2000)).

Log-linear relations, either exact or approximate, are convenient for many purposes. For studying the links between macroeconomics and finance, however, they are limiting for at least two reasons. First, asset pricing investigates how risk exposure is priced. It is the components of this risk exposure that are linked to macroeconomic shocks that are valued. Characterizing risk exposure necessarily leads to the study of volatility and characterizing value necessarily leads to the study of covariation. Second, models that feature time variation in risk exposure or the risk prices require the introduction of nonlinearity in the underlying stochastic process modeling. Even if it is the long-run implications that are featured, probability tools that consider nonlinear implications of a stochastic structure are required.

4.1 Some interesting limiting behavior

For a multiplicative functional M, define its asymptotic growth (or decay) rate as:

$$\lim_{t \to \infty} \frac{1}{t} \log E\left[M_t | X_0 = x\right] = \rho(M)$$

provided that this limit is well defined. I will be interested in the stronger approximation result that

$$\lim_{t \to \infty} \exp\left[-\rho(M)t\right] E\left[M_t f(X_t) | X_0 = x\right] \propto e(x) \tag{9}$$

which justifies calling $\rho(M)$ a growth rate and provides a more refined approximation. Moreover, while the coefficient in this limit will depend on the choice of f, there is a limiting form of state dependence captured by the function e, which and is independent of f. I will say more about the admissible functions f and about the coefficients in this approximation in our subsequent analysis. The essential point is that the rate $\rho(M)$ will not depend on the function f as long as it resides in a potentially rich class of such functions. Eventually, I will show how to represent the proportionality coefficient in the limit of (9) as as a linear functional of f.

The function e is positive, and it solves:

$$E\left[M_t e(X_t) | X_0 = x\right] = \exp\left[\rho(M)t\right] e(x) \tag{10}$$

for all $t \geq 0$. This is an eigenvalue equation, which I will say more about later. While there may be multiple eigenvalues associated with alternative strictly positive eigenfunctions, typically at most one of these eigenvalue-eigenfunction pairs is of interest to us. The resulting eigenvalue $\rho(M)$ is referred to as the *principal eigenvalue* and the associated eigenfunction e is the *principal eigenfunction.*

4.2 Operator families

A key step in my analysis is the construction of a family of operators from a multiplicative functional M. Formally, with any multiplicative functional M we associate a family of operators:

$$\mathbb{M}_t f(x) = E\left[M_t f(X_t) | X_0 = x\right] \tag{11}$$

indexed by t. When M has finite first moments, this family of operators is at least well defined on the space L^∞ of bounded functions.

I use alternative constructions of M, and feature depictions of M as a product of components. The stochastic process components have explicit economic interpretations including stochastic discount factor processes, macroeconomic growth trajectories, or growth processes used to represent hypothetical cash flows to be priced. My use of stochastic discount factor processes to reflect valuation is familiar from empirical asset pricing. (For instance, see Harrison and Kreps (1979), Hansen and Richard (1987), Cochrane (2001), and Singleton (2006).) A stochastic discount factor process decays asymptotically in contrast to a growth process. Such decay is needed for an infinitely lived equity with a growing cash flow to have a finite value as in the case of equity. In contrast to this earlier literature, I am interested in the stochastic process of discount factors over alternative horizons t as a way the dynamics in valuation in the presence of stochastic growth.

Why feature multiplicative functionals? The operator families that interest us are necessarily related. They must satisfy one of two related and well known *laws*: the Law Iterated Expectations and the Law of Iterated Values. The Law of Iterated Values imposes temporal consistency on valuation. In the case of models with frictionless trade at all dates, it is enforced by the absence of arbitrage. In the frictionless market model prices are modeled as the output from forward-looking operators:

$$\mathbb{S}_t f(x) = E\left[S_t f(X_t) | X_0 = x\right].$$

In this expression S is a stochastic discount factor process and $f(X_t)$ is a contingent claim to a consumption numeraire expressed as a function of a Markov state at date t and $\mathbb{S}_t f$ depicts its current period value. Thus $\mathbb{M}_t = \mathbb{S}_t$ and $M = S$. The Law of Iterated Values restricted to this Markov environment is:

$$\mathbb{S}_t \mathbb{S}_\tau = \mathbb{S}_{t+\tau} \tag{12}$$

for $t \geq 0, \tau \geq 0$ where $\mathbb{S}_0 = \mathbb{I}$, the identity operator. To understand this, the date t price assigned to a claim $f(X_{t+\tau})$ is $\mathbb{S}_\tau f(X_t)$. The price of buying a contingent claim at date 0 with *payoff* $\mathbb{S}_\tau f(X_t)$ is given by the left-hand side of (12) applied to the function f. Instead of this two-step implementation, consider the time zero purchase of the contingent claim $f(X_{t+\tau})$. Its date zero purchase price is given by the right-hand side of (12).

Alternatively, suppose that $\mathbb{E}_t$ is a conditional expectation operator for date t associated with a Markov process. This is true by construction when $M = 1$, because in this case:

$$\mathbb{E}_t f(x) = E\left[f(X_t)|X_0 = x\right]$$

As we will see other choices of M can give rise to expectation operators provided that we are willing to alter the implicit Markov evolution. The Law of Iterated Expectations or the Chain Rule of Forecasting implies:

$$\mathbb{E}_t\mathbb{E}_\tau = \mathbb{E}_{t+\tau}$$

for $\tau \geq 0$ and $t \geq 0$. In the case of conditional expectation operators, $\mathbb{E}_t 1 = 1$ but this restriction is not necessarily satisfied for valuation operators.

These laws are captured formally as statement that the family of operators should be a semigroup.

Definition 4.1. *A family of operators* $\{\mathbb{M}_t\}$ *is a (one-parameter) semigroup if a)* $\mathbb{M}_0 = \mathbb{I}$ *and b)* $\mathbb{M}_t\mathbb{M}_\tau = \mathbb{M}_{t+\tau}$ *for* $t \geq 0$ *and* $\tau \geq 0$.

I now answer the question: Why use multiplicative functionals to represent operator families? I do so because a multiplicative functional M guarantees that the resulting operator family $\{\mathbb{M}_t : t \geq 0\}$ constructed using (11) is a one parameter semigroup.

In valuation problems, stochastic discount factors are only one application of multiplicative functionals. Multiplicative functionals are also useful in building cash flows or claims to consumption goods that grow over time. While X may be stationary, the cash flow

$$C_t = G_t f(X_t)\tilde{G}_0$$

displays stochastic growth when G is a multiplicative functional. I include the adjustment $\tilde{G}_0$ because I normalized the the growth process to be one at date zero. Scaling by $\tilde{G}_0$ is a way to ensure that the initialization $G_0 = 1$ is indeed only a normalization. Moreover, shifting the vantage point from time zero to time t,

$$\frac{C_{t+\tau}}{G_t} = \left(\frac{G_{t+\tau}}{G_t}\right) f(X_{t+\tau}) \left(G_t\tilde{G}_0\right).$$

I study cash flows of this type by building an operator that alters the transient contribution to the cash flow $f(X_t)$. This leads us to study

$$\mathbb{P}_t f(x) = E\left[G_t S_t f(X_t) | X_0 = x\right].$$

The value assigned to C_t is given by $\tilde{G}_0 \mathbb{P}_t f(X_0)$ because $\tilde{G}_0$ is presumed to be in the date zero information set. Importantly, it is the product of two multiplicative functionals that we use for representing the operator $\mathbb{P}_t$: $M = GS$.

4.3 Products and co-dependence

Covariances play a prominent role in representing risk premia in asset valuation. I will suggest a long-run counterpart that is motivated by studying the behavior of products of multiplicative functionals. While the product of two multiplicative functionals is multiplicative, it is not true that

$$\rho\left(M^{[1]} M^{[2]}\right) = \rho\left(M^{[1]}\right) + \rho(M^{[2]}).$$

Co-dependence is important when characterizing even the limiting behavior of the product $M^{[1]} M^{[2]}$. In fact the discrepancy:

$$\rho\left(M^{[1]} M^{[2]}\right) - \rho\left(M^{[1]}\right) - \rho(M^{[2]}). \tag{13}$$

will be used to give a long-run version of a risk premium. If $M_t^{[1]}$ and $M_t^{[2]}$ happen to be jointly log normal for each t, then (13) is equal to the limiting covariance between the corresponding logarithms:

$$\lim_{t \to \infty} \frac{1}{t} \text{Cov}\left(Y_t^{[1]}, Y_t^{[2]}\right)$$

where $M^{[j]} = \exp(Y^{[j]})$ for $j = 1, 2$. While this illustrates that co-dependence plays a central role in $\rho\left(M^{[1]} M^{[2]}\right)$, we will not require log-normality in what follows.

To motivate such calculations, for a given horizon the risk premium on a cash-flow $G_t f(Xt)$ paid out at date t and valued at date zero is measured by:

$$\begin{array}{ccccc} \frac{1}{t} \log E\left[G_t f(X_t) | X_0 = x\right] & - & \frac{1}{t} \log E\left[S_t G_t f(X_t) | X_0 = x\right] & + & \frac{1}{t} \log E\left[S_t | X_0 = x\right] \\ \textbf{log} & & \textbf{log} & & \textbf{log} \\ \textbf{expected payoff} & - & \textbf{price} & - & \textbf{riskfree return} \end{array} \tag{14}$$

where f is specified such that the respective logarithms are well defined. The term:

$$\frac{E\left[G_t f(X_t) | X_0 = x\right]}{E\left[S_t G_t f(X_t) | X_0 = x\right]}$$

is the expected return on the investment over the horizon t, and

$$\frac{1}{E\left[S_t|X_0 = x\right]}$$

is the expected return on a riskless investment.

By letting t shrink to zero and computing marginal changes in the risk exposure, we obtain local risk premia familiar in the continuous-time asset pricing literature. The methods described in this paper permit me to study the limit as the investment horizon is made arbitrarily long and to explore the corresponding changes in risk exposure. Provided that f is inconsequential to the limit, the long-horizon limit is

$$\rho(G) + \rho(S) - \rho(SG).$$

In the log-normal case this limiiting risk premia will turn out to be the negative of the covariance of the increments of the martingale components of $\log G$ and $\log S$ (see calculations in subsection 6.2).

Prior to proceeding, I comment a bit on the previous literature. The study of the dynamics of risk premia is familiar from the work of Wachter (2005), Lettau and Wachter (2007) and Hansen et al. (2008). Hansen et al. (2008) characterize the resulting limiting risk premia and the associated risk prices in a log-linear environment.[8] Hansen and Scheinkman (2007) extend this approach to fundamentally nonlinear models with a Markov structure. The perturbation method of section 7 gives a way to compute risk prices in nonlinear Markov environment.[9] Later I will be extending this characterization to produce long-term risk prices and long-term risk return tradeoffs.

4.4 Local versus global

In the decomposition of an additive functional, the linear trend coefficient ν is averages the local state dependent growth rate. I now explore the relation between the local, state

[8] Hansen et al. (2008) also consider the limiting behavior of holding period returns. This limit includes contributions from the principle eigenfunction and the principal eigenvalue of the associated valuation operator for pricing cash flows with stochastic growth components.

[9] Wachter (2005) develops a computational approach based on pricing what she calls "zero-coupon" equity, which in our notation is $E\left[S_t G_t f(X_t)|X_0 = x\right]$. Her algorithm has component prices that converges to zero as the horizon is extended. By using an adaptation of the so-called "power method", these prices can be rescaled to have nondegenerate limit. The limiting function is a principal eigenfunction of the type that I have described. The power method rescales each iteration and hence adjusts for the asymptotic decay. The limit of this rescaling reveals the eigenvalue. Using this more refined characterization of the limit could improve computational performance, and the results of Hansen and Scheinkman (2007) provide justification for the limit approximation.

dependent growth rate and the long-run counterpart.

Consider for the moment a special class of multiplicative functionals:

$$M_t = \exp\left[\int_0^t \beta(X_u)du\right].$$

Such functionals are special because they are smooth, or locally riskless. The multiplicative functional has a state dependent growth rate given by $\beta(x)$. If $\beta(x)$ were constant (state independent), then the long-run growth rate $\rho(M)$ and the local growth rate would coincide. When β fluctuates, $\log(M_t)$ will have a well defined average growth rate where the average is computed using the stationary distribution for X. Jensen's inequality prevents us from just exponentiating this average to compute $\rho(M)$.

The limit $\rho(M)$ is a key ingredient in the study of large deviations. While $\frac{1}{t}\int_0^t \beta(X_u)du$ may obey a Law of Large Numbers and converge to its unconditional expectation under the stationary distribution, more can be said about small probability departures from this law. Large deviation theory seeks to characterize these departures by evaluating expectations under the stationary distribution for an alternative probability measure assigned to X. The same tools used in large deviation theory allow me study the link between β and ρ where β is the local growth or decay rate and ρ is the long-run counterpart.

Let Q be a probability distribution over the state space $\mathcal{E}$ of the Markov process X. Following Donsker and Varadhan (1976), Dupuis and Ellis (1997) and others, construct a *rate function* $\mathbb{I}(Q)$ to measure the discrepancy between the original stationary distribution and Q. See appendix A for a construction of this measure and for a discussion of how it relates to some of my discussion that follows. The function $\mathbb{I}$ is convex in the probability measure Q, and it is used to construct what is called a Laplace principle that characterizes the limit:

$$\rho(M) = \sup_Q \int \beta(x)dQ - \mathbb{I}(Q) \geq E\left[\beta(X_t)\right] \tag{15}$$

for alternative choices of β. The inequality follows because $\mathbb{I}(Q) = 0$ when Q is the stationary distribution of the Markov process X.

This optimization problem is inherently static, with the dynamics loaded into the construction of convex function $\mathbb{I}$. The function $\mathbb{I}$ is constructed independent of the choice of β. Recall that β is the local growth rate of M and its associated semigroup. The long-run limiting growth rate of a multiplicative functional and its associated semigroup exceeds on average the local growth rate integrated against the stationary distribution of the underlying

Markov process. Optimization problem (15) characterizes formally this difference.[10]

This analysis applies to stochastic growth and to stochastic discounting. For instance, these methods provide a general way to characterize the link between short-term and short-term discounting as posed by Newell and Pizer (2003, 2004). Motivated by problems in climate policy analysis, they study discounting abstracting from stochastic growth and the associated risk adjustments. Their focus is on the long-run consequences for valuation of fluctuations in discount rates. Optimization problem 15 gives a general answer to the question they pose for Markov valuation problems. For this application, let $-\beta$ be the short-term (in this case instantaneous) interest rate that is state dependent and hence fluctuates. Then $-\rho(M)$ is long-term counterpart. Optimization problem shows that in the long-run it is a distorted average of $-\beta$ that is germane for discounting. Given the maximization over alternative probability distributions, $-\rho(M)$ will be less than the average of $-\beta$. The magnitude of this discrepancy depends on the potency of the convex penalty function $I(Q)$.[11] As we will see the resulting penchant for small long-term discount rates can be undermined by taking account of risk.

For more general multiplicative functionals, the local growth rate is defined as:

$$\beta^*(x) = \lim_{t \downarrow 0} \frac{E\left(M_t | X_0 = x\right) - 1}{t}$$

provided that this limit exists. When $M_t = \exp(A_t)$ and

$$A_t = \int_0^t \beta(X_u) du + \int_0^t \xi(X_u) \cdot dW_u + \sum_{0 \le u \le t} \lambda(X_u, X_{u-})$$

as in (3), the local growth rate is

$$\beta^*(x) = \beta(x) + \frac{1}{2}|\xi(x)|^2 + \int \left(\exp[\lambda(y, \cdot)] - 1\right) \eta(dy|x).$$

Direct exposure to Brownian motion risk and jump risk alters the local growth rate.

The multiplicative functional M can be decomposed into two component multiplicative

[10]Large deviation theory exploits problem (15) because $\rho(M)$ implies a bound of the form:

$$\text{Prob}\left\{\frac{1}{t}\int_0^t \beta(X_u) \ge k\right\} \le \exp\left(t\left[\rho(M) - k\right]\right)$$

for large t. This bound is only revealing when $k > \rho(M)$. Our interest in $\rho(M)$ is different, but the probabilistic bound is also intriguing.

[11]While this analysis allows for nonlinearity in the Markov dynamics, it does not include the case in which the process $\{\beta(X_t)\}$ is nonstationary except through a sequence of approximating models.

functionals:

$$M_t = \exp\left(\int_0^t \beta^*(X_u)du\right) M_t^* \tag{16}$$

where M^* is a local martingale.[12] Both components are multiplicative functionals. When this local martingale is a martingale, it is associated with a distorted probability distribution for X.[13] The probability twisting associated with M^* preserves the Markov structure. The entropy measure discussed previously is now constructed relative to the probability distribution associated with M^*. This extension permits M processes that are not locally predictable, provided that we change probability distributions in accordance with M^*. The long-run growth rate $\rho(M)$ remains the solution to a convex optimization problem:

$$\rho(M) = \sup_Q \left[\int \beta^*(x)dQ - \mathbb{I}^*(Q)\right] \tag{17}$$

where $\mathbb{I}^*$ is constructed using the change in probability measure.[14] While the inequality associated with optimization problem 15 is satisfied, it is satisfied only after the change measure. The average local growth rate could be greater than the long-run growth rate computed under the original probability measure when there the multiplicative functional is exposed locally to risk. In this sense the variability channel featured by Newell and Pizer (2003, 2004) could be even be more than off set by the presence of local exposure of the discount or growth factors to risk.[15] This latter "exposure" is what captures local risk premia, that is risk-premia for short-term investments as I will characterize shortly. Risk adjustments also have long-run consequences as reflected in the formula:

$$\rho(GS) - \rho(S)$$

for the long-term risk-adjusted rate of return.

[12] In the case of supermartingales, this decomposition can be viewed as a special case of one obtained by Ito and Watanabe (1965). They show that any multiplicative supermartingale can be represented as the following product of two multiplicative functionals:

$$M_t = M_t^\ell M_t^d$$

where $\{M_t^\ell : t \geq 0\}$ is a nonnegative local martingale and $\{M_t^d : t \geq 0\}$ is a decreasing process whose only discontinuities occur where $\{X_t : t \geq 0\}$ is discontinuous.

[13] Applied to valuation problem without growth this distorted probability distribution is the risk neutral distribution familiar from mathematical finance.

[14] The link between this optimization problem and the eigenvalue problem is well known in the literature on large deviations in the absence of a change of measure, for instance see Donsker and Varadhan (1976), Balaji and Meyn (2000) and Kontoyiannis and Meyn (2003).

[15] This off-set is important to produce an upward sloping term structure of interest rates.

5 A revealing special case

Prior to a more formal development, I illustrate calculations using an environment with an underlying continuous-time Markov chain that visits only a finite number of states.

I characterize long-run stochastic growth (or decay) by posing and solving an approximation problem using what is called a *principal* eigenvector and eigenvalue. The principal eigenvector has positive entries. As I will illustrate, there is a well-defined sense in which this eigenvector dominates over long valuation horizons. The approximation problem I will study more generally in the next section borrows its origins from what is known as Perron-Frobenius theory of matrices.

When a Markov process has an n states, the mathematical problem that we study can be formulated in terms of matrices. To model a jump process, consider a matrix $\mathbb{N}$ with all nonnegative entries as a way to encode the conditional measure $\eta(dx^*|x)$. Recall that this measure encodes both the jump intensity (the likelihood of a jump) of the underlying Markov chain X and the jump distribution (conditioned on a jump where will the process jump to). The matrix of transition probabilities for X over an interval of time t is known to be given by $\exp(t\mathbb{A})$ where

$$\mathbb{A} = \mathbb{N} - \text{diag}\{\mathbb{N}\mathbf{1}_n\}$$

where $\mathbf{1}_n$ is an n-dimensional vector of ones and $\text{diag}\{\cdot\}$ is a diagonal matrix with the entries of the vector argument located in the diagonal positions. Notice in particular that $\mathbb{A}$ has all positive entries in the off-diagonal positions, and it satisfies $\mathbb{A}\mathbf{1}_n = \mathbf{0}_n$. This property is the local counterpart to the requirement that the entries in any row of $\exp(t\mathbb{A})$ are the transition probabilities conditioned on the state associated with the selected row. That is, $\exp(t\mathbb{A})\mathbf{1}_n = \mathbf{1}_n$.

For a multiplicative functional associated with an n-state jump process, state dependent growth or decay rates are modeled using β and κ. Recall that β dictates the growth or decay absent any jump and κ dictates how the multiplicative function jumps as a function of the jumps in the underlying Markov process jumps. For this discrete-state problem, I represent the function β as vector $\mathbf{b}$. Similarly, I represent function $\exp[\kappa(x^*, x)]$ as an n by n matrix $\mathbb{K}$ with positive entries. Form an n by n matrix $\mathbb{B}$

$$\mathbb{B} = \mathbb{K} \times \mathbb{N} - \text{diag}\{\mathbb{N}\mathbf{1}_n\} + \text{diag}\{\mathbf{b}\}$$

where $\times$ used in the matrix multiplications denotes element-by-element multiplication. This construction $\mathbb{B}$ of modifies $\mathbb{A}$ to include state dependent growth (or decay) associated with the corresponding multiplicative functional. The off-diagonal entries of $\mathbb{B}$ are all positive, but typically $\mathbb{B}\mathbf{1}_n$ is not equal to $\mathbf{0}_n$ in contrast to $\mathbb{A}\mathbf{1}_n$.

As in section 4, I form an indexed family of operators, in this case matrices, indexed by the time horizon by exponentiating the matrix $t\mathbb{B}$:

$$\mathbb{M}_t = \exp(t\mathbb{B}).$$

The date t matrix $\mathbb{M}_t$ reflects the expected growth, discounting or the composite of both over an interval of time t. The entries of $\mathbb{M}_t$ are all nonnegative, and I presume that for some time horizon t, the entries are strictly positive. The matrix $\mathbb{M}_t$ is typically not be a probability matrix in our applications, however. (Column sums are not unity.) Instead $\mathbb{M}_t$ reflects continuous compounding of stochastic growth or discounting over a horizon t. The matrix $\mathbb{B}$ encodes the instantaneous contributions to growth or discounting, and it *generates* the family of matrices $\{\mathbb{M}_t : t \geq 0\}$.

Given an $n \times 1$ vector $\mathbf{f}$, Perron-Frobenius theory characterizes limiting behavior of $\frac{1}{t} \log \mathbb{M}_t \mathbf{f}$ by first solving:

$$\mathbb{B}\mathbf{e} = \rho \mathbf{e}.$$

where $\mathbf{e}$ is a column eigenvector restricted to have strictly positive entries and ρ is a real eigenvalue. Consider also the transpose problem

$$\mathbb{B}'\mathbf{e}^* = \rho \mathbf{e}^* \tag{18}$$

where $\mathbf{e}^*$ also has positive entries. Depending on the application, ρ can be positive or negative. Importantly, ρ is larger than the real part of any other eigenvalue of the matrix $\mathbb{B}$.

Taking the exponential of a matrix preserves the eigenvectors and exponentiates the eigenvalues. As a consequence, $\mathbb{M}_t$ has an eigenvector given by $\mathbf{e}$ and with an associated eigenvalue equal to $\exp(\rho t)$. The multiplication by t implies that the magnitude of $\exp(\rho t)$ relative to the other eigenvalues of $\mathbb{M}_t$ becomes arbitrarily large as t gets large. As a consequence,

$$\lim_{t \to \infty} \frac{1}{t} \log \mathbb{M}_t \mathbf{f} = \rho$$

$$\lim_{t \to \infty} \exp(-\rho t) \mathbb{M}_t \mathbf{f} = (\mathbf{f} \cdot \mathbf{e}^*) \mathbf{e}$$

for any vector $\mathbf{f}$ where we have normalized $\mathbf{e}^*$ so that $\mathbf{e}^* \cdot \mathbf{e} = 1$. This formally defines ρ as the long-run growth rate of the family of matrices $\{\mathbb{M}_t : t \geq 0\}$. The eigenvector $\mathbf{e}$ gives the direction that dominates in the long run. Thus this long-run characterization via eigenvector analysis exposes the impact of state-dependent compounding of growth or discounting over long horizons.

I will use an extension of this method to determine model specifications which have important long-run effects on the matrix $\mathbb{B}$ used in modeling instantaneous transitions. Long-term implications can be disguised in the construction of the local transitions. My aim is to see through this disguise. I will explore several different constructions of the operator counterpart to $\mathbb{B}$, reflecting alternative hypothetical economic environments or alternative economic inputs. The constructions are specifications of the analogs to the vector $\mathbf{b}$ and the matrix $\mathbb{K}$ in this example. I will be interested in how ρ and $\mathbf{e}$ change as I alter $\mathbb{B}$ in ways that are motivated explicitly through economic considerations.

In the next section I will develop a decomposition of the valuation dynamics (DVD) that is particularly convenient when the Markov process has continuous states. A essential component is the construction of a martingale that I use to change the probability measure. This change of measure allows me to appeal to well known convergence properties of Markov processes to study valuation in the presence of growth. I use the solution to (18) to build a new generator of a Markov chain:

$$\hat{\mathbb{A}} = \left[\text{diag}\left\{\mathbf{e}\right\}\right]^{-1} \mathbb{B}\text{diag}\left\{\mathbf{e}\right\} - \rho\mathbb{I}$$

Notice that since the matrix $\mathbb{A}$ has positive off-diagonal entries, the same is true of the matrix $\hat{\mathbb{A}}$. Moreover, since (18) is satisfied, $\hat{\mathbb{A}}\mathbf{1}_n = \mathbf{0}_n$. As a consequence, $\left\{\exp(t\hat{\mathbb{A}}) : t \geq 0\right\}$ is the family (semigroup) of transition matrices for an Markov chain, one that is useful in approximating long-term consequences in a more general stochastic setting.[16]

The operators I consider in the next section can have a complicated eigenvalue structure because I allow a more general specification of the underlying Markov process. I will avoid characterizing fully this structure, but instead I will use martingale methods that exploit representations of the operator families as I next describe. I will construct the martingale component of original multiplicative functional M using the principal eigenfunction and eigenvalue counterparts to $\mathbf{e}$ and ρ.

[16]In the applied mathematics literature this new chain is sometimes referred to as the *twisted* chain.

6 Multiplicative factorization

I now propose a multiplicative factorization of stochastic growth functionals with three components: a) deterministic growth rate, b) a positive martingale, c) a transient component:

$$
\begin{array}{ccccc}
M_t & = & \exp(\rho t) & \hat{M}_t & \left[\frac{\hat{e}(X_t)}{\hat{e}(X_0)}\right] \\
 & & \uparrow & \uparrow & \uparrow \\
 & & \textbf{exponential trend} & \textbf{martingale} & \textbf{transient}
\end{array}
\tag{19}
$$

Component a) governs the long-term growth or decay. It is constructed from a principal eigenvalue. I will use component b), the positive martingale, to build an alternative probability measure. This alternative measure gives us a tractable framework for a formal study of approximation. Component c) is built directly from the principle eigenfunction and characterizes transient departures in behavior that are distinct from martingale behavior.

This decomposition is suggestive. All three components are themselves multiplicative functionals, but with very different behavior. Consider the separate components. The term $\exp \rho t)$ captures exponential growth. The multiplicative martingale has expectation unity for all t and in this sense is not *expected* to grow. The third component appears *transient* when the underlying Markov process is stationary. While the stochastic inputs of the martingale $\hat{M}$ will be long lasting, perhaps the same is not true for the transient component. Although positive, this martingale will typically not converge to a limiting random variable with unit expectation. For instance, its logarithm can have stationary increments.

This component-by-component analysis turns out to be misleading. The components are correlated and this correlation can have an important impact on the long-run expected behavior of the process. Thus I am led to ask: Is this decomposition unique? When is this decomposition useful? The answers to these questions are intertwined.

6.1 Decomposition

I build the decomposition as follows. First I solve:

$$E\left[M_t e(X_t)|X_0 = x\right] = \exp(\rho t)e(x) \tag{20}$$

for any t where e is strictly positive as in (10). The function e can be viewed as a *principal eigenfunction* of the semigroup with ρ being the corresponding eigenvalue. Since this equation holds for any t, it can be localized by computing:

$$\lim_{t \downarrow 0} \frac{E\left[M_t e(X_t) | X_0 = x\right] - \exp(-\rho t) e(x)}{t} = 0, \tag{21}$$

which gives an equation in e and ρ to be solved. The local counterpart to this equation is

$$\mathbb{B} e = \rho e, \tag{22}$$

where

$$\lim_{t \downarrow 0} \frac{E\left[M_t e(X_t) - e(x) | X_0 = x\right]}{t} = \mathbb{B} e$$

The operator $\mathbb{B}$ is the so-called *generator* of the semigroup constructed with the multiplicative functional M. It is an operator on a space of appropriately defined functions. Heuristically, it captures the local evolution of the semigroup. In the case of a diffusion model, this generator is known to be a second-order differential operator:

$$\mathbb{B} f = \left(\beta + \frac{1}{2}|\xi|^2\right) f + (\sigma \xi' + \mu) \cdot \frac{\partial f}{\partial x} + \frac{1}{2} \text{trace}\left(\sigma \sigma' \frac{\partial^2 f}{\partial x \partial x'}\right).$$

It is convenient to express the corresponding eigenvalue equation in terms of $\log e$ after dividing the equation by e:

$$\rho = \left(\beta + \frac{1}{2}|\xi|^2\right) + (\sigma \xi' + \mu) \cdot \frac{\partial \log e}{\partial x} + \frac{1}{2} \text{trace}\left(\sigma \sigma' \frac{\partial^2 \log e}{\partial x \partial x'}\right) + \frac{1}{2} \left(\frac{\partial \log e}{\partial x}\right)' \sigma \sigma' \left(\frac{\partial \log e}{\partial x}\right)$$

We have seen the finite-state counterpart to this equation in section 5.

Typically it suffices to solve the local equation (22) to obtain a solution to (20). See Hansen and Scheinkman (2007) for a more detailed discussion of this issue. In the finite-state Markov model of section 5, convenient and well known sufficient conditions exist for there to be a unique (up to scale) positive eigenfunction satisfying (20). More generally, however, this uniqueness will not hold. Instead I will obtain uniqueness from additional considerations.

Given a solution to (20), I construct a martingale via:

$$\hat{M}_t = \exp(-\rho t) M_t \left[\frac{e(X_t)}{e(X_0)}\right],$$

which is itself a multiplicative functional. The multiplicative decomposition (19) follows immediately by letting $\hat{e} = \frac{1}{e}$ and solving for M in terms of $\hat{M}$, ρ and $\hat{e}$.

6.2 Additive versus multiplicative decomposition

There are important differences in the study of additive and multiplicative functionals. It can be misleading to simply exponentiate the decomposition of an additive functional because of the dependence between components. This dependence can change the configuration of permanent and transitory components.

In special cases, however, the two are related.

Example 6.1. *Consider again example 3.3 and recall the additive functional:*

$$dY_t = \nu dt + HX_t dt + FdW_t.$$

Form

$$M_t = \exp(Y_t).$$

While the exponential of a martingale is not a martingale, in this case the exponential of the additive martingale will become a martingale provided that we multiply the additive martingale by an exponential function of time. This simple adjustment exploits the lognormal specification as follows:

$$\hat{M}_t = \exp\left(\hat{Y}_t - \frac{t}{2}|F - HA^{-1}B|^2\right).$$

is a martingale. The growth rate for M *is:*

$$\rho(M) = \nu + \frac{|F - HA^{-1}B|^2}{2}$$

In this case it is easy to go from a martingale decomposition of an additive functional to that of a multiplicative functional.

An equivalent way to proceed is to build e *as an exponential of a linear function of* x*, and to seek a solution to (22). It may be verified that* $e(x) = \exp(-HA^{-1}x) = \exp[g(x)]$ *is the solution to this equation for* $\rho = \nu + \frac{|F-HA^{-1}B|^2}{2}$*. Thus* e *is obtained by exponentiating the function* g *used in the additive martingale construction. The eigenvalue* ρ *includes an extra volatility adjustment as is typical in log-normal models.*

Consider two log-normal functionals $M^{[1]}$ *and* $M^{[2]}$ *parameterized by* (η_i, F_i, H_i) *for* $i = 1, 2$*. A simple calculation reveals that*

$$\rho(M^{[1]}M^{[2]}) - \rho(M^{[1]}) - \rho(M^{[2]}) = (F_1 - H_1A^{-1}B) \cdot (F_2 - H_2A^{-1}B),$$

which is the covariance of the increments to the martingale components of $\log M^{[1]}$ *and* $\log M^{[2]}$.

In this log-normal example there is a simple link between the additive decomposition and the multiplicative factorization. This link breaks down when volatility is state dependent as is apparent in my next example.

Example 6.2. *Consider again Example 3.4, and recall the additive functional:*

$$dY_t = \nu dt + H_1 X_t^{[1]} dt + H_2 (X_t^{[2]} - 1) dt + \sqrt{X_t^{[2]}} F dW_t.$$

Form

$$M_t = \exp(Y_t).$$

Guess a solution $e(x) = \exp(\alpha \cdot x)$ *where* $x = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$ *and* $\alpha = \begin{bmatrix} \alpha_1 \\ \alpha_2 \end{bmatrix}$. *To compute* $\rho(M)$, *I solve a special case of (22):*

$$\nu + {x_1}' ({A_1}' \alpha_1 + {H_1}') + (x_2 - 1)(A_2 \alpha_2 + H_2) + \frac{1}{2} x_2 |\alpha' B + F|^2 = \rho.$$

which I derive as a special case of (22). Thus the coefficients on x_1 *and* x_2 *are zero when:*

$$\begin{aligned} {A_1}' \alpha_1 + {H_1}' &= 0 \\ A_2 \alpha_2 + H_2 + \frac{1}{2} |{\alpha_1}' B_1 + \alpha_2 B_2 + F|^2 &= 0. \end{aligned} \tag{23}$$

The first equation can be solved for α_1 *and the second one for* α_2 *given* α_1. *The solution to the first equation is:*

$$\alpha_1 = -({A_1}')^{-1} {H_1}'$$

The second equation is quadratic in α_2, *so there may be two solutions. Specifically,*

$$\alpha_2 = -\left(\frac{B_2 \cdot F + A_2}{|B_2|^2} \right) \pm \frac{\sqrt{|B_2 \cdot F + A_2|^2 - |B_2|^2 \left(|F - H_1 (A_1)^{-1} B_1|^2 + 2H_2 \right)}}{|B_2|^2}, \tag{24}$$

provided that the term under the square root sign is positive. Notice in particular that this term will be positive for sufficiently small $|B_2|$. *We will have cause to select one of these solutions as the interesting one. Finally,*

$$\rho = \nu - (A_2 \alpha_2 + H_2).$$

6.3 Martingales and changes in probabilities

Why might positive multiplicative martingales be of interest? A positive martingale scaled to have unit expectation is known to induce an alternative probability measure. This trick is a familiar one from asset pricing, but it is valuable in many other contexts. Since $\hat{M}$ is a martingale, I form the distorted or twisted expectation:

$$\hat{E}\left[f(X_t)|X_0\right] = E\left[\hat{M}_t f(X_t)|X_0\right].$$

For each time horizon t, I define an alternative conditional expectation operator. The martingale property is needed so that the resulting family of conditional expectation operators obeys the Law of Iterated Expectations. It insures consistency between the operators defined using $\hat{M}_{t+\tau}$ and $\hat{M}_t$ for expectations of random variables that are in the date t conditioning information sets. Moreover, with this (multiplicative) construction of a martingale, the process remains Markov under the change in probability measure.

Definition 6.3. *The process X is* **stochastically stable** *under the measure $\hat{\cdot}$ if*

$$\lim_{t\to\infty} \hat{E}\left[f(X_t)|X_0 = x\right] = \hat{E}\left[f(X_t)\right] \tag{25}$$

for any f for which $\hat{E}(f)$ is well defined and finite.[17]

Theorem 6.4. *Given a multiplicative functional M, suppose that e and ρ satisfy equation (21) and that X is stochastically stable under the $\hat{\cdot}$ probability measure. Then*

$$E\left[M_t f(X_t)|X_0 = x\right] = \exp(\rho t)\hat{E}\left[\frac{f(X_t)}{e(X_t)}|X_0 = x\right] e(x).$$

Moreover,

$$\lim_{t\to\infty} \exp(-\rho t) E\left[M_t f(X_t)|X_0 = x\right] = \hat{E}\left[f(X_t)\hat{e}(X_t)\right] e(x)$$

provided that $\hat{E}\left[f(X_t)\hat{e}(X_t)\right]$ is finite where $\hat{e} = 1/e$.

This theorem gives a method for long-run approximation, which is quite distinct from log-linear methods that approximate around a steady state. Instead a martingale component of M is used to the change the probability measure, approximation can proceed using tools from the study of Markov processes with stable stochastic dynamics. Notice that the stability condition is presumed to hold under the distorted or twisted probability distribution. Establishing this property allows us to ensure that the dependence between the martingale and transient components is limited sufficiently so that we may think of ρ as the exponential

[17] This is stronger than ergodicity because it rules out periodic components. Ergodicity requires that time series averages converge but not necessarily that conditional expectation operators converge. Under ergodicity the time series average of the conditional expectation operators would converge but not necessarily the conditional expectation operators.

growth rate. In other words, this is necessary for

$$\rho = \rho(M)$$

defined previously.

It follows from Theorem 6.4 that once we scale by the growth rate ρ, we obtain a one-factor representation of long-term behavior. Changing the function f simply changes the coefficient on the function e. Thus the state dependence is approximately proportional to e as the horizon becomes large. For this method to justify our previous limits, we require that $f\hat{e}$ have a finite expectation under the $\hat{\cdot}$ probability measure. The class of functions f for which this approximation works depends on the stationary distribution for the Markov state of the $\hat{\cdot}$ probability measure and the function $\hat{e}$. These functions of the Markov state have transient contributions to valuation since for these components:

$$\lim_{t\to\infty} \frac{1}{t} \log E\left[M_t f(X_t)|X_0\right] = \rho(M).$$

Definition 6.5. *For a given multiplicative functional M, a process $f(X)$ is* transient *if X is stochastically stable under the probability measure implied by the martingale component and $\hat{E}[f(X_t)\hat{e}(X_t)]$ is well defined and finite.*

The family of f's that define transient processes determines the sense in which the principal eigenvalue and function dominate in the long run. How rich this collection will be is problem specific. As we will see, there are important examples when this density has a fat tail which limits the range of the approximation. On the other hand, the process X can be strongly dependent under the $\hat{\cdot}$ probability measure.

There is an extensive set of tools for studying the stability of Markov processes that can be brought to bear on this problem. For instance, see Meyn and Tweedie (1993) for a survey of such methods based on the use of Foster-Lyapunov criteria. See Rosenblatt (1971), Bhattacharya (1982) and Hansen and Scheinkman (1995) for alternative approaches based on mean-square approximation. While there may be multiple representations of the form (19), Hansen and Scheinkman (2007) show that there is at most *one* such representation for which the process X is stochastically stable.

Recall that in example 6.2 we found two solutions for α_2 by solving the quadratic equation (23). As an implication of the Girsanov Theorem, associated with each solution is as alternative probability measure under which

$$dW_t = \sqrt{X_t^{[2]}}\left(F' + B_1{}'\alpha_1{}' + B_2{}'\alpha_2\right)dt + d\hat{W}_t.$$

where $\hat{W}_t$ is a multivariate standard Brownian motion under the twisted measure. The implied *twisted* evolution equation for $X_t^{[2]}$ is:

$$\begin{aligned} dX_t^{[2]} &= A_2 X_t^{[2]} dt + \left(B_2 F' + |B_2|^2 \alpha_2\right) X_t^{[2]} dt + \sqrt{X_t^{[2]}} d\hat{W}_t \\ &= \pm \left[\sqrt{|B_2 F' + A_2|^2 - |B_2|^2 \left(|F - H_1 (A_1)^{-1} B_1|^2 + 2H_2\right)}\right] (X_t^{[2]} - 1) dt + \sqrt{X_t^{[2]}} d\hat{W}_t. \end{aligned}$$

where in the second representation I have substituted from solution (24). I select the "minus" solution to achieve stochastic stability.

6.4 Long-run behavior of multiplicative martingales

As I have shown, the martingale component $\hat{M}$ is valuable as a means of changing the probability measure and studying approximation as the time horizon becomes large. The martingale is useful provided that implies a stochastic evolution that is stochastically stable. This change of measure is what causes me to find a multiplicative martingale to be valuable. From another perspective, the multiplicative martingale can have degenerate or unusual behavior in the limit. This behavior does not resemble the central limit approximation I deduced for an additive martingale.

Since a multiplicative martingale is positive, it is bounded from below. By the Martingale Convergence Theorem $\hat{M}$ converges to a limiting random variable that I denote $\hat{M}_\infty$. While

$$E\left(\hat{M}_t | X_0 = x\right) = 1$$

for all t, it may be that $E\left(\hat{M}_\infty | \mathcal{F}_0\right) \leq 1$ and is often zero. For instance, it is zero in the log-normal example 3.3. While the martingale induces an alternative "twisted " probability measure, it does so in a way that is not absolutely continuous in the limit as the t becomes arbitrarily large. The twisted probability of limit events may assign positive probability to events that previously had measure zero. The multiplicative martingale remains valuable as a change of measure when the stochastic dynamics are stable even though the martingale itself may converge to zero.

I obtain a more refined characterization of the behavior following an approach initiated by Chernoff (1952).[18] Specifically I bound a threshold probability by taking expectations of a dominating function:

$$\frac{1}{t} \log Pr\left\{\hat{M}_t \geq \exp(\mathsf{b}) | X_0 = x\right\} \leq \frac{1}{t} \log E\left[(\hat{M}_t)^{\mathsf{a}} | X_0 = x\right] - \frac{\mathsf{ab}}{t} \leq 0$$

[18]See Newman and Stuck (1979) for a continuous-time Markov version of this formulation.

for any $0 \leq \mathsf{a} \leq 1$. Provided that the left-hand side limit is strictly negative, I have an exponential bound on the threshold probability for the multiplicative martingale as the horizon is extended. This bound may be optimized by the choice of a. Notice that $\hat{M}^{\mathsf{a}}$ is itself a multiplicative functional (in fact a multiplicative supermartingale) and can be studied using the methods described in this paper. Such bounds give a precise sense in which large positive movements in $\hat{M}$ over long horizons are unlikely. Notice that as the horizon gets large the contribution of b to the bound on the right-hand side becomes inconsequential. The limiting exponential decay rate does not depend on the chosen threshold. Thus while $\hat{M}$ is used productively as a change in probability measure used in computing limiting growth and decay rates, the process itself has a tendency to become small under some perspectives.

6.5 Transient model components

I now explore what it means for there to be temporary growth components or temporary components to stochastic discount factors. I focus on a stochastic discount factor process implied by an asset pricing model, but there is an entirely analogous treatment of a stochastic growth functional.

Consider a benchmark valuation model represented by a stochastic discount factor or a benchmark growth process or the product of the two components. I now ask what modifications are transient? The tools I describe in section 6 give an answer to this question.

Consider a benchmark multiplicative functional M. Recall our multiplicative decomposition:

$$M_t = \exp(\rho t)\hat{M}_t \frac{\hat{e}(X_t)}{\hat{e}(X_0)}.$$

Moreover suppose that under the associated $\hat{\cdot}$ probability measure X satisfies a stochastic stability condition 6.3. Consider an alternative model of the form:

$$M_t^* = M_t \frac{\hat{f}(X_t)}{\hat{f}(X_0)}$$

for some $\hat{f}$ where M is used to represent a benchmark model and M^* an alternative model. As argued by Bansal and Lehmann (1997) and others, a variety of asset pricing models can be represented like this with the time-separable power utility model used to construct M. Function $\hat{f}$ may be induced by changes in the preferences of investors such as habit persistence or social externalities.

I use the multiplicative decomposition of M to construct an analogous decomposition for

M^*. Given the decomposition:

$$M_t = \exp(\rho t)\hat{M}_t \frac{\hat{e}(X_t)}{\hat{e}(X_0)},$$

the corresponding decomposition for M^* is:

$$M_t^* = \exp(\rho t)\hat{M}_t \frac{\hat{e}(X_t)\hat{f}(X_t)}{\hat{e}(X_0)\hat{f}(X_0)}.$$

While the martingale component remains the same, the set of transitory components is altered because $\{f(X_t)\}$ will be transitory in this alternative representation when:

$$\hat{E}\left[f(X_t)\hat{e}(X_t)\hat{f}(X_t)\right] < \infty.$$

In particular, this restriction depends on $\hat{f}$. Later I will explore applications and show the importance of this restriction.

7 Perturbation calculation

A risk price is a marginal change in a risk premium induced by a marginal change in risk exposure. It is a derivative of particular kind obtained by parameterizing the exposure to risk, including growth rate risk. We consider limiting versions of risk prices by parameterizing multiplicative functionals as they depend on risk exposure. Relatedly, an economic model depends on an underlying set of parameters, some of which have important consequences to valuation. For instance, Hansen et al. (2008) considers a parameterized set of valuation models that depend on the intertemporal elasticity of the investors. As a precursor, or sometime an alternative, to solving models for alternative parameter configurations, derivatives computed at a baseline configuration of parameters reveal how sensitive the valuation model is to small changes in the parameters.

In these applications the multiplicative functionals used in constructing the semigroup depend on model parameters. Thus I consider $M(\mathsf{a})$ as a parameterized family. The parameterizations can capture a variety of alternative features of the underlying economic model. It can be a preference parameter as in the work of Hansen et al. (2008), or it can be a parameter that governs the exposure to a source of long-term risk that is to be valued. It is informative to explore sensitivity to changes in a variety of features of the underlying economic model. With a perturbation analysis, it is possible to exploit a given solution to a model in the study of sensitivity to model specification. Perturbing $M(\mathsf{a})$ by changing a is equivalent to perturbing the operators associated with this process. My choice of scalar parameterization is made for notational convenience. The multivariate extension is straightforward.

In the Hansen et al. (2008) application, $\mathsf{a} + 1$ is a common intertemporal substitution parameter across investors. The aim is to study how long-run risk premia change with a. Thus the stochastic discount factor process is depicted as $S(\mathsf{a})$ and $M(\mathsf{a}) = S(\mathsf{a})G$ where G is the stochastic growth component of a hypothetical or real cash flow. As an alternative, a could parameterize the long-run risk exposure of a hypothetical cash flow. In this case $M(\mathsf{a}) = SG(\mathsf{a})$. For instance, suppose that $G(\mathsf{a})$ is a parameterized family of multiplicative martingales. The simplest such example is:

$$\log G_t(\mathsf{a}) = \int_0^t \xi(X_u;\mathsf{a})dW_u - \frac{1}{2}\int_0^t |\xi(X_u;\mathsf{a})|^2 du.$$

For future reference note that

$$\frac{d}{d\mathsf{a}} \log G_t(\mathsf{a}) = \int_0^t D\xi(X_u;\mathsf{a})[dW_u - \xi(X_u;\mathsf{a})du] \tag{26}$$

where $D\xi(X_u;\mathsf{a})$ is the partial derivative of ξ with respect to a. While long-term expected returns will often not be linear in the risk exposure vector $\xi(\mathsf{a})$, the derivative of the long-term expected rate of return with respect to a gives the marginal change in value induced by a marginal change in risk exposure. It gives a local price of risk. By exploring alternative parameterizations of risk exposures, I can infer which directions are of most concern to investors as reflected by an underlying economic model.

In this section I will show that these perturbations have a simple structure when the focus is on long-run implications. Specifically, I compute:

$$\frac{d}{d\mathsf{a}}\rho(\mathsf{a}). \tag{27}$$

Calculation (27) turns out to be straightforward. First solve the principal eigenvalue problem for $\mathsf{a} = 0$ and use the solution to construct a probability measure $\hat{\cdot}$ as we described previously. The formula for the derivative is:

$$\frac{d}{d\mathsf{a}}\rho(\mathsf{a})|_{\mathsf{a}=0} = \frac{1}{t}\hat{E}\left[\frac{d\log M_t(\mathsf{a})}{d\mathsf{a}}\bigg|_{\mathsf{a}=0}\right] \tag{28}$$

which can be evaluated for any choice of t including choices that are arbitrarily small. Since $\log M(\mathsf{a})$ is an additive functional so is its derivative, $\frac{d\log M(\mathsf{a})}{d\mathsf{a}}$. Interestingly, I obtain the derivative of ρ by computing the average of the average trend growth of $\frac{d\log M(\mathsf{a})}{d\mathsf{a}}$ under the twisted $\hat{\cdot}$ probability measure.

In the remainder of this section, I give a heuristic derivation for this formula. Many readers may just prefer to accept this formula including the limiting version as the investment horizon becomes small.

7.1 Basic derivation

Consider first the finite horizon calculation. Let $M(\mathsf{a})$ be a parameterized family of multiplicative functionals. There is an associated parameterized family of valuation functionals:

$$\mathbb{M}_t(\mathsf{a})f(x) = E\left[M_t(\mathsf{a})f(X_t)|X_0 = x\right].$$

Then under some regularity conditions,

$$\begin{aligned} \frac{d}{d\mathsf{a}} \log \mathbb{M}_t(\mathsf{a})f(x)|_{\mathsf{a}=0} &= \frac{E\left(M_t(0)\left[\frac{d}{d\mathsf{a}}\log M_t(\mathsf{a})\big|_{\mathsf{a}=0}\right]f(X_t)|X_0 = x\right)}{E\left[M_t(0)f(X_t)|X_0 = x\right]} \\ &= \frac{\hat{E}\left(\left[\frac{d}{d\mathsf{a}}\log M_t(\mathsf{a})\big|_{\mathsf{a}=0}\right]\frac{f(X_t)}{e(X_t;0)}|X_0 = x\right)}{\hat{E}\left[\frac{f(X_t)}{e(X_t;0)}|X_0 = x\right]} \end{aligned} \tag{29}$$

where the $\hat{\cdot}$ is computed under the twisted probability distribution. The notation $e(x, \mathsf{a})$ denotes the principal eigenfunction associated with $\mathbb{M}_t(\mathsf{a})$.

To study the limiting counterpart, recall the decomposition:

$$M_t(\mathsf{a}) = \exp\left[\rho(\mathsf{a})t\right]\hat{M}_t(\mathsf{a})\frac{e(X_0;\mathsf{a})}{e(X_t;\mathsf{a})}$$

where I have used our parameterization of M and the fact that parameterizing M in terms of a is equivalent to parameterizing the components. Consider first the martingale component. Here I borrow an insight from maximum likelihood estimation. Note that

$$E\left[\hat{M}_t(\mathsf{a})|X_0 = x\right] = 1$$

for all a. The derivative of this expectation with respect to a is necessarily zero. Thus

$$\hat{E}\left[\frac{d}{d\mathsf{a}}\log \hat{M}_t(\mathsf{a})|_{\mathsf{a}=0}|X_0 = x\right] = E\left[\frac{d}{d\mathsf{a}}\hat{M}_t(\mathsf{a})|X_0 = x\right] = 0.$$

Many readers familiar with statistics will have a feeling of familiarity. This argument is essentially the usual argument from maximum likelihood estimation for why a score vector for a likelihood function has mean zero where $\frac{d}{d\mathsf{a}}\log \hat{M}_t(\mathsf{a})$ evaluated at $\mathsf{a} = 0$ is the score of the likelihood over an interval of time t.

Now use the decomposition and differentiate $\log M_t(\mathsf{a})$

$$\frac{d}{d\mathsf{a}} \log M_t(\mathsf{a}) = t\frac{d\rho(\mathsf{a})}{d\mathsf{a}} + \frac{d}{d\mathsf{a}} \log \hat{M}_t(\mathsf{a}) - \frac{d}{d\mathsf{a}} \log e(X_t; \mathsf{a}) + \frac{d}{d\mathsf{a}} \log e(X_0; \mathsf{a}).$$

Take expectations and use the fact that X is stationary under the $\hat{\cdot}$ probability measure to obtain derivative formula (28). See appendix B for a discussion of the principal eigenfunction.

7.2 Using the local evolution

I now make formula (28) operational in continuous time. Under the $\hat{\cdot}$ change of measure, I let $\hat{\xi}(X_t)dt$ denote the drift of the Brownian motion W implying that new drift for X is

$$\hat{\mu}(x) = \mu(x) + \sigma(x)\hat{\xi}(x).$$

I let

$$\hat{\eta}(dy|x) = \exp[\hat{\lambda}(y, x)]\eta(dy|x)$$

denote the new measure used to capture local evolution of the jump component to the Markov process. Recall that this conditional measure encodes the jump intensity and the jump distribution conditioned on a jump taking place.

The functional $\log M_t(\mathsf{a})$ is an additive functional, and its derivative is as well. Recall the continuous time model of Y we specified in equation (3):

$$Y_t(\mathsf{a}) = \int_0^t \beta(X_u; \mathsf{a})du + \int_0^t \xi(X_u; \mathsf{a}) \cdot dW_u + \sum_{0 \leq u \leq t} \lambda(X_u, X_{u-}; \mathsf{a})$$

and form $M(a) = \exp[Y(a)]$. It is most convenient to take limits of (28) as $t \to 0$. This entails computing an average local mean under the distorted distribution:

$$\begin{aligned} \frac{d}{d\mathsf{a}}\rho(\mathsf{a})|_{\mathsf{a}=0} &= \hat{E}\left(\frac{d}{d\mathsf{a}} \left[\beta(X_t; \mathsf{a}) + \xi(X_t; \mathsf{a}) \cdot \hat{\xi}(x)\right]\Big|_{\mathsf{a}=0}\right) \\ &\quad + \hat{E}\left[\int \frac{d}{d\mathsf{a}}\lambda(y, X_t; \mathsf{a})\Big|_{\mathsf{a}=0} \exp[\hat{\lambda}(y, X_t)]\eta(dy|X_t)\right] \end{aligned} \tag{30}$$

where we have used the fact that the Brownian motion has $\hat{\xi}(X_t)dt$ as the drift under the $\hat{\cdot}$ distribution and used the conditional measure $\exp[\hat{\lambda}(y, X_t)]\eta(dy|X_t)$ to construct the $\hat{\cdot}$ the jump intensity and the jump distribution conditioned on the current Markov state.

In this section I have been a bit heuristic or cavalier about taking derivatives. Formal treatments do currently exist in the applied mathematics literature. For example Kontoyiannis and Meyn (2003) (see their Proposition 6.2) consider formally smoothness of parameterized families of operators in their formal development of large deviation results for Markov processes.

8 Applications to Asset Pricing

In my study of asset pricing, I consider two limits. One reproduces the *local* risk prices familiar from asset pricing theory by taking limits as the investment horizon shrinks to zero, and the other constructs *long-term* risk prices as limits when the investment horizon is made arbitrarily large. A "term structure" of risk prices connects theses two limits. I have already characterized long-term risk premia in the presence of stochastic growth using the formula:

$$\rho(G) + \rho(S) - \rho(SG).$$

In what follows, I let G be a multiplicative martingale as a way to abstract from cash flow dynamics. For such a G, $\rho(G) = 0$ because the is no *expected* cash flow growth. A reader may object by claiming that I have now eliminated growth altogether. Even worse almost all of the sample paths of G may converge to zero. Consider, however, a more general multiplicative specification of a cash flow. Typically it is the martingale component that determines the long-term risk prices and not the transient component. Moreover, fluctuations in growth are embedded in the martingale component, and the deterministic exponential growth component does not alter risk premia or prices at any horizon.[19] Instead of extracting martingale components from initial multiplicative growth processes, I will build them directly and explore the resulting pricing implications. Long-term risk prices are imputed from a change in risk premia induced by a marginal change in the the cash-flow risk exposure, that is a derivative. This computation uses the formulas presented in section 7.

In this section I ask: what are the long-term implications for alternative models of the stochastic discount factor? Among the models I consider are ones designed to enhance short-term risk prices and induce variation in these prices over time. I use the apparatus described in previous sections to explore the implications for long-term risk prices, and I provide revealing comparisons across some models that are currently featured in the asset pricing literature.

[19]While multiplicative martingales may have degenerate long-run behavior, we could apply Theorem 3.2 and eliminate the trend term in logarithms. This allows for central-limit-type behavior for long horizons, and it does not alter the implied risk premia and corresponding risk prices.

8.1 Stochastic discount factors

Multiplicative representations pervade the asset pricing literature. Various changes have been proposed for the familiar power utility model. There is menu of such models in the literature featuring alternative departures. Consider an initial benchmark specification:

$$S_t = \exp(-\delta t) \left(\frac{C_t}{C_0} \right)^{-\gamma}.$$

Many alterations in this model take the form:

$$S_t^* = S_t \frac{\hat{f}(X_t)}{\hat{f}(X_0)}.$$

Transient components in asset pricing models have been included to produce short run fluctuations in asset prices. As argued by Bansal and Lehmann (1997), these fluctuations may take the form of habit persistence or as an extension of the power utility model of investor preferences.

8.2 Models without consumption predictability

In this subsection I explore a simple model of consumption dynamics under which the power utility model has transparent implications. My aim is to study how risk prices change across horizons for alternative models.

Suppose that consumption is a geometric Brownian motion:

$$d \log C_t = \mu_c dt + \sigma_c dW_t,$$

where C_t is aggregate consumption. I allow the Brownian motion $\{W_t : t \geq 0\}$ to be multivariate. Construct S in accordance with the power utility model:

$$S_t = \exp\left(-\delta t - \gamma t \mu_c - \gamma \int_0^t \sigma_c \cdot dW_u \right).$$

where $\frac{1}{\gamma}$ is intertemporal elasticity of substitution and δ is the subjective rate of discount. For risk pricing, I introduce a growth functional that is a martingale:

$$G_t = \exp\left(\int_0^t \sigma_g \cdot dW_u - \frac{t}{2} |\sigma_g|^2 \right)$$

for the reasons I stated previously.

In what follows I will make comparisons between a model with investors that have preferences represented by discounted, time separable, power utility (a Breeden model) and model in which a temporally dependent social externality is introduced in the manner proposed by Campbell and Cochrane (1999). In terms of my previous notation S is the stochastic discount factor for the power utility model and S^* is the stochastic discount factor a decentralized Campbell and Cochrane (1999) model. The reference to decentralization is important because internalizing the social externality alters the stochastic discount factor. A social planner would internalize this externality and this would be reflected in the stochastic discount factor.

I will provide a precise formula for S^* subsequently, but I first describe the results. If the power parameter γ is held fixed across the model specifications, the local risk prices are known to be be very different as I will illustrate. Not only are they systematically larger with the consumption externality, they vary over time. What happens as we change horizons? To address this, I study the limit prices. Specifically, I characterize the limiting risk premia:

$$\textbf{risk premium} = \rho(S^*) + \rho(G) - \rho(S^*G)$$

and see how they change as I alter the risk exposure. The risk prices are the derivatives of the risk premium with respect to σ_g used to represent stochastic growth.

By construction, $\rho(G) = 0$. The long-term risk price vector for the Breeden (1979)'s model is $\gamma\sigma_c$. I will show that the long-term risk prices for the Campbell-Cochrane model remain the same with a very important proviso. There is a discontinuity when $\sigma_g = 0$. Specifically, I will show that

$$\textbf{risk premium} = \gamma\sigma_c \cdot \sigma_g + r - r^*. \tag{31}$$

where r is the riskless rate of interest for the power utility model and r^* is the corresponding rate for the Campbell and Cochrane (1999) model for the same value of the subjective rate of discount.[20] The magnitude of the discontinuity is given by interest rate difference. This discontinuity is depicted in figure 1 in which we depict the risk premia when investors care about external habits and when they do not. Typically the risk premia converges to zero as σ_g converges to zero, but in fact the limiting risk premium is the differential in the risk-free

[20] I do not mean to imply that an econometrician or calibrator would select the same value of δ for each model. For instance, Campbell and Cochrane (1999) and Wachter (2005) use values of the subjective rate of discount that are much larger than would be used if the Breeden (1979) model was calibrated to asset return data. Even if the subjective rate of discount for the Campbell-Cochrane model is to fit interest rates, the calculation of r using this same subjective rate of discount, although counterfactual, is a revealing input into the risk-premia formula for the Campbell-Cochrane model.

rates between the the Campbell and Cochrane (1999) model and the Breeden (1979) model. In figure 1 this discontinuity is sizable. It is the distance on the vertical axis between the circle and the dot. While this discontinuity is only present in the limit, it is indicative that risk prices are large near $\sigma_g = 0$ for valuation over long-time horizons.

In contrast when I use a specification of consumption externalities proposed by Santos and Veronesi (2006) the discontinuity is not present. Santos and Veronesi (2006) imitate the increase in local prices that are present in the Campbell-Cochrane model, but the term structure of "risk prices" can be quite different. For the Santos-Veronesi model I will provide formulas for the entire term structure and not just the limit points.

In the remainder of this subsection, I provide the details of this calculation. Uninterested readers can skip to the next subsection in which I make comparisons to a model in which investor preferences are represented by an alternative recursive utility function.

8.2.1 Risk prices in the power utility model

For the benchmark S model and the martingale specification of the growth process G, the martingale factorization is:

$$S_t G_t = \hat{M}_t \exp\left[-\delta t - \gamma\mu_c t + \frac{t}{2}|-\gamma\sigma_c + \sigma_g|^2 - \frac{t}{2}|\sigma_g|^2\right],$$

where

$$\hat{M}_t = \exp\left[\int_0^t (\sigma_g - \gamma\sigma_c) dW_u - \frac{t}{2}|-\gamma\sigma_c + \sigma_g|^2\right].$$

It follows that

$$\rho(SG) = -\delta - \gamma\mu_c + \frac{\gamma^2}{2}|\sigma_c|^2 - \gamma\sigma_c \cdot \sigma_g.$$

By setting $\sigma_g = 0$,

$$\rho(S) = -\delta - \gamma\mu_c + \frac{\gamma^2}{2}|\sigma_c|^2.$$

Thus

$$\rho(G) + \rho(S) - \rho(GS) = \gamma\sigma_c \cdot \sigma_g$$

The long-term risk prices can be computed by differentiating the right-hand side with respect to the risk exposure vector σ_g, and are thus equal to: $\gamma\sigma_c$.

The dynamics of pricing for this example is degenerate, and in particular the local risk price vector is also equal to $\gamma\sigma_c$. Specifically, the local expectedrate of return is given by

$$-\lim_{t\downarrow 0} \frac{1}{t} \log E\left[G_t S_t | X_0 = x\right] = \delta + \gamma\mu_c + -\frac{\gamma^2}{2}|\sigma_c|^2 + \gamma\sigma_c \cdot \sigma_g. \qquad (32)$$

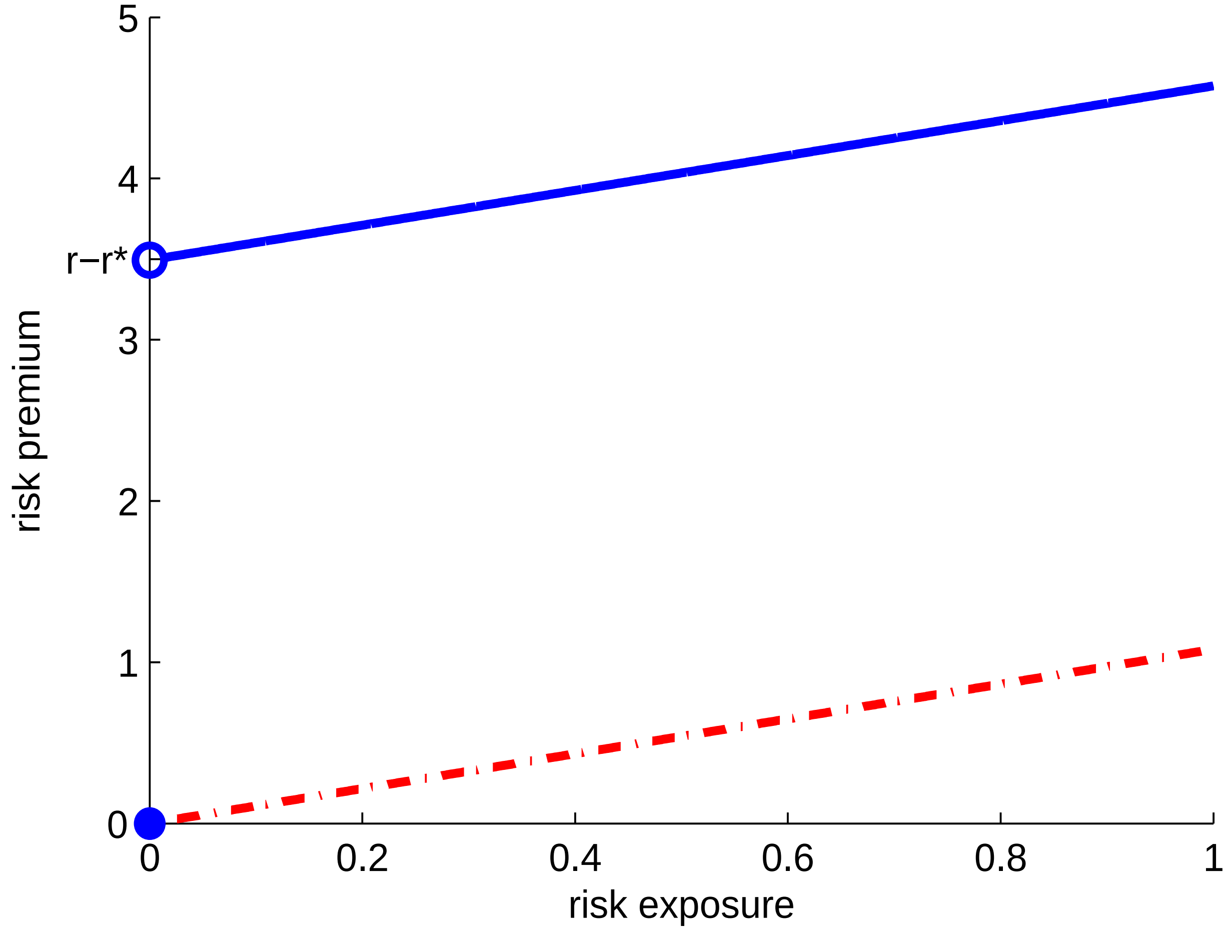

Figure 1: Risk premia as function of risk exposure. The horizontal axis is given in quarterly time units. The vertical axis is scaled by one hundred so the the risk premia are in percent. The dot-dashed line denotes the implied premia when investors have external habits, and the solid line denotes the implied premia when investors have expected utility preferences. The parameter values for the state evolution are: $F_c = 0.0054$ and $\nu_c = .0056$. I set $\gamma = 2$, and for the model with investors that have external habits I set $\theta = 0$ and $\xi = .035$.

By setting $\sigma_g = 0$ notice that the instantaneous risk-free interest rate is constant and identical to the long-term counterpart: $-\nu(S) = \delta + \gamma\mu_c - \frac{\gamma^2}{2}|\sigma_c|^2$. The vector of risk prices obtained by differentiating the local risk-premium with respect the risk-exposure vector σ_g is $\gamma\sigma_c$, which is identical to the long-run counterpart. This link between the short-run and long-run prices follows because of the separability and absence of state dependence in preferences of the investor and the lack of predictability in aggregate consumption. Later in this section I will relax the underlying assumptions and explore the short-run and long-run consequences.

In the calculations that follows I will use the multiplicative martingale $\hat{M}$ as a change of measure. As a result the the process W is no longer a standard Brownian motion but is altered to have a drift $-\gamma\sigma_c + \sigma_g$. This is an application of the Girsanov Theorem that is used extensively in mathematical finance and elsewhere.

8.2.2 An example of Campbell and Cochrane

Campbell and Cochrane (1999) modify the Breeden asset pricing model with power utility by introducing a stochastic subsistence point process C^* that shares the same stochastic growth properties as consumption. In language of time series, this process is cointegrated with consumption. The process C^* could be a social externality, which justifies its dependence on consumption shocks. Alternatively, it is a way to model exogenous preference shifters that depend on the same shocks as consumption. The resulting stochastic discount factor process is:

$$S_t^* = \exp(-\delta t) \left[\frac{(C_t - C_t^*)^{-\gamma}}{(C_0 - C_0^*)^{-\gamma}} \right]$$

We may rewrite this as:

$$S_t^* = S_t \left[\frac{(1 - C_t^*/C_t)^{-\gamma}}{(1 - C_0^*/C_0)^{-\gamma}} \right].$$

In what follows let

$$X_t = -\log(1 - C_t^*/C_t) - \mathsf{b},$$

which we model as a process that exceeds zero. Notice that adding a positive constant b to X_t preserves the positivity and it does not alter the pricing implications. It does alter investor risk aversion (see Campbell and Cochrane (1999) or the appendix C). Using this notation, write:

$$S_t^* = S_t \left[\frac{\exp(\gamma X_t)}{\exp(\gamma X_0)} \right].$$

Following Campbell and Cochrane (1999) and Wachter (2005), assume that

$$dX_t = -\xi(X_t - \mu_x)dt + \lambda(X_t)\sigma_c dW_t \tag{33}$$

where we restrict $\lambda(0) = 0$ in hopes that the zero boundary will not be attainable. Squashing the variability at zero prevents the process from being attracted to zero. After the probability distortion, the law of motion for this equation is:

$$dX_t = -\xi(X_t - \mu_x)dt + (\sigma_g - \gamma\sigma_c) \cdot \sigma_c \lambda(X_t)dt + \lambda(X_t)\sigma_c d\hat{W}_t. \tag{34}$$

We use this evolution to compute the counterpart to (32):

$$\begin{aligned} \lim_{t\downarrow 0} \frac{1}{t} \log E\left[G_t S_t^* | X_0 = x\right] &= \lim_{t\downarrow 0} \frac{1}{t} \log E\left[G_t S_t \exp[\gamma(X_t - X_0)] | X_0 = x\right] \\ &= -r - \gamma\sigma_c \cdot \sigma_g - \lim_{t\downarrow 0} \frac{1}{t} \hat{E}\left[\exp[\gamma(X_t - X_0)] | X_0 = x\right] \\ &= -r - \gamma\sigma_c \cdot \sigma_g + \gamma\xi(x - \mu_x) + \gamma(\sigma_g - \gamma\sigma_c) \cdot \sigma_c \lambda(x) \\ &\quad + \gamma^2 \frac{\lambda(x)^2 |\sigma_c|^2}{2} \end{aligned} \tag{35}$$

where r is the risk-free rate from the Breeden economy ($r = \delta + \gamma\mu_c + \frac{\gamma^2}{2}|\sigma_c|^2$) and the last equality is computed using Ito's formula.

Consider first the interest rate behavior. Campbell and Cochrane (1999), suppose the risk-rate is an affine function of the state: $r^* + \theta(x - \mu_x)$. With this outcome, the parameter θ controls the variation in the risk-free rate. To support this functional form the value of r^* is

$$r^* = r + (\theta - \gamma\xi)\mu_x.$$

The volatility function λ is given by

$$\lambda(x) = 1 - (1 + \zeta x)^{1/2}$$

where

$$\zeta \doteq \frac{2(\gamma\xi - \theta)}{\gamma^2 |\sigma_c|^2}.$$

(See appendix C.) In order that the term inside the square root be positive, $\theta < \gamma\xi$.

The local risk prices for the Campbell-Cochrane model are the entries of the vector:

$$\gamma\sigma_c - \gamma\lambda(x)\sigma_c = \gamma\left(1 + \zeta x\right)^{1/2} \sigma_c$$

which follows because (35) is affine in σ_g and the risk prices are the negative of the partial derivative with respect to σ_g. By design are state dependent and are larger than in the power utility model for given value of γ. Moreover, the state variable increment dX_t responds negatively to consumption growth shocks because $\lambda(x) < 0$. By design, risk premia are larger in bad times as reflected by unexpectedly low realizations of consumption growth. As demonstrated by Campbell and Cochrane (1999) in their closely related discrete time model, the coefficient of relative risk aversion is also enhanced and in fact it is equal to $\gamma[1-\lambda(\mu_x)]$ for $\mu_x = x$. (See also appendix C.)

Consider next the long-run behavior of value. I use evolution equation (34), and the formula for the logarithmic derivative of the density for a scalar diffusion:

$$\frac{d \log q}{dx} = \frac{2 \ \text{drift}}{\text{diffusion}} - \frac{d \log \text{diffusion}}{dx} \tag{36}$$

where the drift coefficient (local mean) is $-\xi(x-\mu_x)$ under the original measure or $-\xi(x-\mu_x) + (\sigma_g - \gamma\sigma_c)\cdot\sigma_c\lambda(X_t)$ under the twisted distribution. The diffusion coefficient (local variance) is $\lambda(x)^2|\sigma_c|^2$.

The limiting behavior is dominated by the constant term:

$$\lim_{x\to\infty} \frac{d\log q}{dx} = -\frac{\gamma^2\xi}{\gamma\xi - \theta} < 0. \tag{37}$$

As a consequence the process X is stationary under the twisted probability measure and under the original probability measure as reflected by (33) and (34) respectively. It remains to study what functions have finite moments under the twisted evolution.

When $\gamma\xi > \theta > 0$, $\exp(\gamma X_t)$ has a finite expectation under the twisted stationary density because the limit in (37) is strictly less than $-\gamma$. In contrast, when $\theta < 0$ this expectation will be infinite. Thus when $\theta > 0$ the contribution to preferences will be transient, but not when $\theta < 0$.

When $\theta = 0$, a more refined calculation is required because $\log q$ behaves like a positive scalar multiple of $-\gamma x$ for large x. This leads me to study,

$$\lim_{x\to\infty} \sqrt{x}\left(\frac{d\log q}{dx} + \gamma\right) = -2\left(\frac{\sigma_g\cdot\sigma_c}{\sigma_c\cdot\sigma_c}\right)\zeta^{-1/2} = -\frac{\sigma_g\cdot\sigma_c}{|\sigma_c|}\sqrt{\frac{2\gamma}{\xi}}.$$

For the modification in the stochastic discount factor to be transient, this term must be negative because twice this limit is the coefficient on $\sqrt{x}$ in the large x approximation of the log density plus γx. While this term is zero when σ_g is zero, it will be negative provided that the shocks to $\log G_t$ and $\log C_t$ are positively correlated.

I now characterize the limiting risk premia:

$$\textbf{risk premium} = \rho(S^*) + \rho(G) - \rho(S^*G).$$

By construction, $\rho(G) = 0$. When $\theta > 0$,

$$\textbf{risk premium} = \rho(S^*) - \rho(S^*G) = \gamma\sigma_c \cdot \sigma_g$$

as in the Breeden (1979) model. When $\theta = 0$ and $\sigma_c \cdot \sigma_g > 0$, $\rho(S^*G)$ is the same as in the Breeden (1979) model:

$$\rho(S^*G) = -\delta - \gamma\mu_g - \gamma\sigma_c \cdot \sigma_g + \frac{\gamma^2}{2}|\sigma_c|^2,$$

but $\rho(S^*)$ differs and is given by the implied real interest rate r^*. This juistifies formula (31) and figure **??**.

I have just shown that the case in which $\theta = 0$ has *special* limiting properties. Campbell and Cochrane (1999) feature this case. The instantaneous interest rate is constant and equal to r^*. The long-term counterpart is the same. Interestingly, when $\theta = 0$, $\exp(\gamma x)$ is a strictly positive solution to the eigenvalue equation:

$$E\left[S_t \exp(\gamma X_t)|X_0 = x\right] = \exp(-r^*t)\exp(\gamma x).$$

It is one of two such solutions since

$$E\left[S_t|X_o = x\right] = \exp(-rt).$$

The multiplicative martingale

$$\tilde{M}_t = \exp(rt)S_t \frac{\exp(\gamma X_t)}{\exp(\gamma X_0)}$$

implies a change in measure, but under this change of measure the process $\{X_t\}$ is stochastically unstable. See Appendix C.

What do we make of this? We constructed two alternative martingales related to the stochastic discount factor process S and hence S^*. Each martingale was built using a positive eigenfunction. Only one implies stable stochastic dynamics for X. As I argued previously, Hansen and Scheinkman (2007) show that this uniqueness is a general result.

When $\theta = 0$ the multiplicative martingale $\tilde{M}$ is the pertinent one for discount bond pricing the martingale $\hat{M}$ for pricing growth rate risk over long horizons. The discontinuity in the long-term risk premia as a function of σ_g as expressed in (31) reflects the separate roles of the two martingales in pricing. When $\theta > 0$, only the multiplicative martingale $\hat{M}$ is pertinent to pricing.

8.2.3 An example of Santos and Veronesi

Santos and Veronesi (2006) consider a related model of the stochastic discount factor. The stochastic discount factor has the form:

$$S_t^* = S_t \left(\frac{X_t + 1}{X_0 + 1} \right).$$

In this case

$$\frac{C_t^*}{C_t} = 1 - \mathsf{b}(X_t + 1)^{-\frac{1}{\gamma}}$$

for some positive number b. Changing b alters the relationship between C and C^*, but not the stochastic discount factor.

The process for X is a member of Wong (1964)'s class of Markov processes built to imply stationary densities that are in the Pearson (1916) family. Wong (1964) characterizes solutions to stochastic differential equations with a linear drift and a quadratic diffusion coefficient. One such process is the one used by Santos and Veronesi:

$$dX_t = -\xi(X_t - \mu_x)dt + \lambda(X_t)\sigma_c dW_t, \quad X_t > 0$$

where

$$\lambda(X_t) = -\kappa X_t$$

and $\mu_x > 0$.[21] The specification of local volatility is designed to keep the process X above unity. As in the Campbell-Cochrane specification, the process X responds negatively to a consumption shock.

The local risk prices are now given by

$$\gamma\sigma_c + \frac{\kappa x}{x + 1}\sigma_c.$$

In addition to being state dependent, they exceed those implied by the power utility model since the second term is always positive and they vary over time.

To study long-term pricing we again use the twisted evolution equation (34) but with this new specification of λ. The twisted law of motion for X is

$$dX_t = -\hat{\xi}\left(X_t - \hat{\mu}_x\right)dt - \kappa X_t \sigma_c \cdot d\hat{W}_t$$

[21] This process is the F process of Wong (1964).

where

$$\begin{aligned}\hat{\xi} &= \xi - \gamma\kappa|\sigma_c|^2 + \kappa\sigma_c \cdot \sigma_g \\ \hat{\mu}_x &= \left(\frac{\xi}{\hat{\xi}}\right)\mu_x\end{aligned}$$

This process remains in the class studied by Wong (1964). To explore its long-term implications, formula (36) is again informative. The logarithmic derivative of the density is

$$\frac{d\log\hat{q}(x)}{dx} = -2\left[\frac{\hat{\xi}(x-\hat{\mu}_x)}{\kappa^2|\sigma_c|^2x^2} - \frac{1}{x}\right].$$

As a consequence, in the right tail behaves like $x^{-\varsigma}$ where

$$\varsigma = 2\left(\frac{\hat{\xi}}{\kappa^2|\sigma_c|^2} + 1\right)$$

The twisted density $\hat{q}$ has a finite first moment provided that $\hat{\xi}$ is positive. The mean is given by $\hat{\mu}_x$. Thus provided that the mean reversion parameter ξ is sufficiently large

$$\xi \geq \gamma\kappa|\sigma_c|^2 - \kappa\sigma_c \cdot \sigma_g. \tag{38}$$

When inequality (38) is satisfied for an open set of values of σ_g that includes zero, the long-term risk prices agree with the power utility model.

A convenient feature of this Santos and Veronesi (2006) model is that the risk prices can be more fully characterized by "paper and pencil". In particular, the logarithm of the expected return for horizon t is:

$$-\frac{1}{t}\log E\left[G_tS_t^*|X_0 = x\right] = \gamma\sigma_c \cdot \sigma_g - \frac{1}{t}\log\hat{E}\left(X_t + 1|X_0 = x\right) + \frac{1}{t}\log(x+1).$$

Moreover,

$$\hat{E}\left(X_t + 1|X_0 = x\right) = 1 + \left[1 - \exp\left(-t\hat{\xi}\right)\right]\hat{\mu}_x + \exp\left(-t\hat{\xi}\right)x$$

The risk prices for a finite horizon are obtained by differentiating the risk premia with respect to σ_g.

In summary, provided that the mean revision parameter ξ is sufficiently large, the behavior of the long-term risk prices for the Santos-Veronesi models are quite different from those that arise in the Campbell-Cochrane specification. Their transient nature is more evident, and there is no discontinuity at $\sigma_g = 0$.

8.3 Predictability in consumption growth and volatility

Suppose now that consumption evolves according to the stochastic evolution of example 3.4 where

$$d \log C_t = \mu_c dt + H_c X_t^{[1]} dt + \sqrt{X_u^{[2]}} F_c dW_t$$

Consumption growth is predictable as captured by $H_c X_t^{[1]}$, and consumption volatility is state dependent as captured by $X_t^{[2]}$. As a point of reference consider first the Breeden (1979) model. Thus the stochastic discount factor is

$$S_t = \exp\left[-\delta t - \gamma t \mu_c - \gamma \int_0^t H_c X_u^{[1]} du - \gamma \int_0^t \sqrt{X_u^{[2]}} F_c dW_u\right].$$

Consider a growth functional constructed as a martingale:

$$G_t \;=\; \exp\left(-\frac{1}{2}|F_g|^2 t - \frac{1}{2}\int_0^t |F_g|^2 (X_u^{[2]} - 1) du + \int_0^t \sqrt{X_u^{[2]}} F_g dW_u\right).$$

In this subsection I compare the model in which investors have power utility (a Breeden model) to a counterpart model with recursive utility using a the risk-sensitive parameterization of Kreps and Porteus (1978) preferences in which the elasticity of intertemporal substitution is unity but the risk aversion parameter coincides with that of the power utility model. I will show that the limiting risk prices are the same even though the risk-free rate differs.

For both models it is straightforward to compute the "term structure" of risk prices. I depict these prices in figure 2 as a function of the investment horizon for a three shock version of the consumption dynamics. These trajectories converge to the limit prices that I will formally characterized. Only the first shock has a direct impact on consumption; the second shock alters the growth rate in consumption via a continuous-time scalar autoregression ($X^{[1]}$), and the third shock alters volatility via a Feller-square root process ($X^{[2]}$). A positive movement in the third shock diminishes consumption volatility. The formula for risk prices are given in appendix (D). In the Breeden (1979) model, the local risk prices are zero for the second two shocks. Investor preferences are forward looking in the recursive utility model, and this is reflected in nonzero local risk prices for the second two shocks. The enhanced local price of the growth rate shock illustrates the pricing mechanism featured by Bansal and Yaron (2004), and the similarity of the risk prices over long-horizons between the Breeden (1979) model and the recursive utility model illustrates a finding in Hansen et al. (2008). The forward-looking feature of recursive preferences leads to a flatter trajectory for the risk prices. The trajectory is literally flat for the first shock and the two models imply the

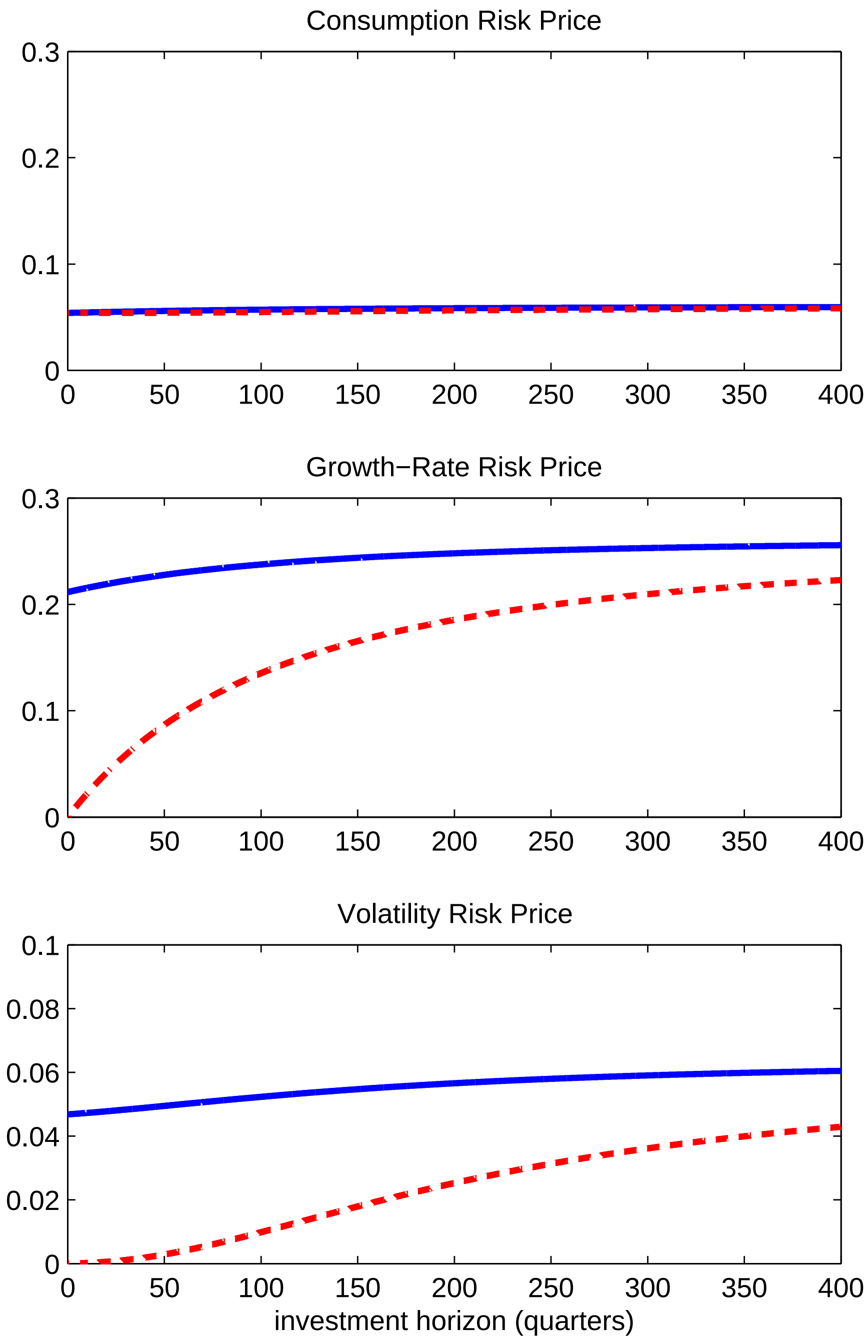

Figure 2: Risk prices indexed by investment horizon. The horizontal axis is given in quarterly time units. The solid line denotes recursive utility model, and dashed line the expected utility model. The parameter values for the state evolution are: $A_1 = -.02$, $A_2 = -.02$, $B_1 = [0 \;\; .00047 \;\; 0]$, $B_2 = [0 \;\; 0 \;\; .038]$, $H_c = [1 \;\; 0]$, $F_c = [0.0054 \;\; 0 \;\; 0]$ and $\nu_c = .0056$. The risk prices are by localizing around $G = C$. For illustrative purposes I set $\gamma = 10$. How to "calibrate" γ is a interesting question in its own right, a question that much has already been written on. I personally like the discussion in Hansen (2007). To construct these plots I set $X_0^{[1]} = 0$ and $X_0^{[2]} = 1$.

same risk prices. The coincidence of the pricing trajectories for the two models of investor preferences illustrates a point made by Kocherlakota (1990). The persistence of large local risk prices for the consumption growth rate shock over long horizons is consistent with the empirical findings of Hansen et al. (2008), although their model of consumption dynamics is different.[22] I specified stochastic volatility to be very persistent, and this is reflected in the slow rate of convergence of the risk prices. While stochastic volatility induces variation in local risk prices, the shock to volatility like the direct shock to consumption commands a relatively small risk price at all horizons.

In what follows I characterize the local prices and their long time horizon limits for the Breeden model and the Kreps-Porteus model. Thus I characterize what happens at both ends of the "term structure" of risk prices. I also argue that risk prices for the Kreps-Porteus model can be interpreted in part as model uncertainty prices using recursive versions of investor preferences that reflect a robust concern about model specification. by borrowing insights from the literature on robust control.

8.3.1 Risk prices for the Breeden model

The local risk price for dW_t is

$$\sqrt{X_t^{[2]}}\gamma F_c.$$

Let

$$dY_t = d\log S_t + d\log G_t,$$

and let

$$\begin{aligned} H_1 &= -\gamma H_c \\ H_2 &= -\frac{1}{2}|F_g|^2 \\ F &= -\gamma F_c + F_g \\ \nu &= -\delta - \gamma\mu_c - \frac{1}{2}|F_g|^2 \end{aligned}$$

Then this specification of Y is a special case of example 3.4. Applying formulas (30) and (26), the long-run risk price is the expected drift under the twisted measure induced by $\hat{M}$ of

$$F_g X_t^{[2]}dt - \sqrt{X_t^{[2]}}dW_t$$

where

$$dW_t = \sqrt{X_t^{[2]}}\left[F + (B_1)'\alpha_1^{[b]} + (B_2)'\alpha_2^{[b]}\right]dt + d\hat{W}_t$$

[22] Hansen et al. (2008) abstract from stochastic volatility and they use a discrete-time vector autoregressive model of consumption and corporate earnings to model the consumption dynamics.

where $\hat{W}$ is a multivariate standard Brownian motion under this alternative measure and the positive eigenfunction is $\exp(\alpha_1^{[b]} \cdot x_1 + \alpha_2^{[b]x_2})$. Thus the long-run risk price vector is:

$$\hat{E}\left(X_t^{[2]}\right) \quad \begin{bmatrix} \gamma F_c & - & (B_1)'\alpha_1^{[b]} & - & (B_2)'\alpha_2^{[b]} \\ \textbf{local} & & \textbf{growth} & & \textbf{volatility} \end{bmatrix}$$

By construction $X^{[2]}$ has mean one under the original probability distribution. The twisted distribution alters this mean because under the distorted probability

$$dX_t^{[2]} = A_2(X_t^{[2]} - 1)dt + B_2\left[F + (B_2)'\alpha_2^{[b]}\right] X_t^{[2]}dt + \sqrt{X_t^{[2]}}d\hat{W}_t$$

with $\hat{W}$ a multivariate standard Brownian motion. Rearranging terms in the drift coefficient gives

$$dX_t^{[2]} = \hat{A}_2\left(X_t^{[2]} - \hat{\mu}_2\right)dt + \sqrt{X_t^{[2]}}d\hat{W}_t$$

where

$$\begin{aligned} \hat{A}_2 &= A_2 + B_2\left[F + (B_2)'\alpha_2^{[b]}\right] \\ \hat{\mu}_2 &= \frac{A_2}{\hat{A}_2} = \hat{E}\left(X_t^{[2]}\right). \end{aligned}$$

I now interpret some of contributions to this price vector. The term:

$$\hat{E}\left(X_t^{[2]}\right)\gamma F_1$$

averages the *local* risk prices scaled by $\sqrt{X_t^{[2]}}$ and averages using the twisted distribution. The remaining terms are induced by the predictability in the consumption *growth rate* and consumption *volatility*. For instance,

$$-\hat{E}\left(X_t^{[2]}\right)(B_1)'\alpha_1^{[b]} = -\gamma\hat{E}\left(X_t^{[2]}\right)(B_1)'[(A_1)']^{-1}(H_c)' \tag{39}$$

reflects the temporal dependence in the growth rate of consumption, as featured in the long-term pricing calculations by Hansen et al. (2008). The third term reflects the temporal dependence in volatility.

8.3.2 Pricing with risk sensitive recursive utility

I now explore a limiting version of a specification of investor preferences that is known to alter local prices. This limit allows me to explore the intersection between two literatures, the literature in economics on recursive utility and the literature on risk-sensitive control theory.

Discounted version of risk-sensitive control typically solves the date zero problem of the investor (see Whittle (1990) for a discussion of the role of discounting). Hansen and Sargent (1995) give a recursive utility version of risk-sensitive control that also accommodates discounting, and Hansen et al. (2006) study this formulation in continuous time. Under this specification, there is risk-sensitive adjustment to the future continuation value of future consumption processes as in Kreps and Porteus (1978) and Epstein and Zin (1989) and it avoids some of the pitfalls of the standard specification of risk-sensitive control. In what follows I use a parameterization of Tallarini (2000) in which the elasticity of intertemporal substitution is unity. This restriction facilitates analytical characterization. I will take limits of the stochastic discount factor as the subjective rate of discount converges to zero. Since consumption grows stochastically, this will push me outside the risk-sensitive, undiscounted, ergodic control studied by Runolffson (1994). In the discounted version of recursive preferences, the stochastic discount factor is;

$$S_t^* = \exp(-\delta t) \left(\frac{C_0}{C_t} \right) \hat{V}_t \tag{40}$$

where $\hat{V}$ is a martingale component of the $\left\{ \left(\frac{V_t}{V_0} \right)^{1-\gamma} : t \geq 0 \right\}$ and V is the stochastic process of continuation values.[23]

The process V and hence $\hat{V}$ are constructed from the underlying consumption dynamics. I use a homogeneous of degree one specification of the utility recursion to construct the continuation value process V implying that any common scaling of current and future consumption results in the same scaling of the continuation value. In formula (40) for a stochastic discount factor, δ continues to be the subjective rate of discount and the inverse ratio of consumption growth reflects the unitary intertemporal elasticity of substitution in the preferences of the investor. The presence of a martingale component in the stochastic discount factor implies a change of probability and reflects the dual robust interpretation of risk-sensitive preferences.

For this model of investor preferences, the continuation value V and consumption C share the same growth components. Their ratio $\frac{V_t}{C_t}$ can be expressed a function $\tilde{f}$ of the Markov

[23]The martingale component is obtained by removing the locally predictable component of $\left\{ \left(\frac{V_t}{V_0} \right)^{1-\gamma} : t \geq 0 \right\}$ as in (16) and Ito and Watanabe (1965), where must verify that the local martingale is in fact a martingale.

state, and $\tilde{f}$ solves the equation:

$$\delta \log \tilde{f}(x) = \lim_{t \downarrow 0} \frac{E\left[\left(\frac{C_t}{C_0}\right)^{1-\gamma} \tilde{f}(X_t)|X_0 = x\right] - f(x)}{t}. \tag{41}$$

The solution gives a formula for $\left(\frac{V_t}{C_t}\right)^{1-\gamma} = \tilde{f}(x)$ from which I solve for $\frac{V_t}{C_t}$.

To study the relation between the stochastic discount factor S^* and the stochastic discount factor S for the power utility model, I take limits as δ tends to zero. The limiting stochastic discount factor remains well defined:

$$S_t^* = \left(\frac{C_0}{C_t}\right) \hat{V}_t,$$

where I abuse notation a bit by recycling notation used originally for an arbitrary positive δ. While the continuation value process becomes degenerate, the ratio $\frac{V_t}{V_0}$ remains well defined in the limit for all t as δ declines to zero. Call the resulting multiplicative process $\tilde{V}$. Construct $\tilde{C}$ in the same fashion by dividing C_t by C_0.

I now describe the construction of the multiplicative martingale $\hat{V}$ in this limiting case. Equation (41) ceases to have a solution when $\delta = 0$. Instead I look for a positive eigenfunction associated with the multiplicative functional $\tilde{C}^{1-\gamma}$. Then

$$\left(\frac{\tilde{V}_t}{\tilde{C}_t}\right)^{1-\gamma} \propto e^{[r]}(X_t) \exp(-\rho^{[r]} t).$$

where $e^{[r]}$ is a positive eigenfunction and $\rho^{[r]}$ the corresponding principal eigenvalue associated with the multiplicative functional $\tilde{C}^{1-\gamma}$.[24] The eigenfunction and value are chosen so that the implied martingale induces a change of measure with stable stochastic dynamics for X.

Recall that I constructed $\tilde{V}$ to be one at date $t = 0$. As a consequence,

$$(\tilde{V}_t)^{1-\gamma} = \exp(-\rho^{[r]} t) \left(\frac{e^{[r]}(X_t)}{e^{[r]}(X_0)}\right) (\tilde{C}_t)^{1-\gamma}$$

The right-hand side is the multiplicative martingale component of $\tilde{C}^{1-\gamma}$. Thus by extracting the martingale component of $(\tilde{C})^{1-\gamma}$, I obtain the martingale component $\hat{V} = \tilde{V}^{1-\gamma}$ for the stochastic discount factor S^* in formula (40).

[24] Even though we have introduced stochastic growth in consumption, there is direct counterpart to $\rho^{[r]}$ and $e^{[r]}$ in Runolffson (1994)'s analysis of stochastic risk sensitive control in the absence of discounting.

With this computation, I turn to studying the relation between S and S^*. Write

$$\hat{V}_t = \exp(-\rho^{[r]}t)\left(\frac{C_t}{C_0}\right)^{1-\gamma}\frac{e^{[r]}(X_t)}{e^{[r]}(X_0)}.$$

As a consequence,

$$S_t^* = \left(\frac{C_0}{C_t}\right)\hat{V}_t = \exp(-\rho^{[r]}t)S_t\left(\frac{e^{[r]}(X_t)}{e^{[r]}(X_0)}\right)$$

This representation suggests that the adjustment to preferences induces transient modifications of risk premia while altering the long-run risk free rate. I expect the interest rate differences to exists because the elasticity of substitution differs for the two models of investor preferences. The long-term risk price calculation given in (39) continues to apply to this model even though the local prices are different from the Breeden (1979) model.[25]

I next consider the local or instantaneous prices for the recursive utility model. The eigenvalue $\rho^{[r]}$ and eigenfunction $e^{[r]}$ capture the differences in the instantaneous interest rate and the eigenfunction $e^{[r]}$ alters the local risk prices *vis a vis* the Breeden (1979) model. These prices are given by

$$\sqrt{X_t^{[2]}}[\gamma F_c - (B_1)'\alpha_1^{[r]} - (B_2)'\alpha_2^{[r]}].$$

The term $\sqrt{X_t^{[2]}}[\gamma F_c$ is local price vector in the Breeden (1979) model. In the recursive utility model, it is modified because of predictability in consumption growth and volatility. The role of consumption predictability is:

$$-\sqrt{X_t^{[2]}}(B_1)'\alpha_1^r = (1-\gamma)\left(\sqrt{X_t^{[2]}}\right)(B_1)'[(A_1)']^{-1}(H_c)'.$$

and is familiar from the analysis in Bansal and Yaron (2004), Campbell and Vuolteenaho (2004) and Hansen et al. (2008). It is a recursive utility enhancement of the local risk prices based on predictability in consumption growth rates. The term

$$-(B_2)'\alpha_2^{[r]}$$

gives an adjustment for the predictability of volatility and is analogous to an adjustment in Bansal and Yaron (2004). There are counterparts to both of these adjustment in the long-term risk prices given in formula (39).

As I remarked previously, there is alternative interpretation of the risk-sensitive model

[25] Hansen et al. (2008) make this observation for a discrete-time log-linear model abstracting from stochastic volatility. Thus distorted expectation in (39) plays no role in their analysis.

of investor preferences. Under this interpretation, $\hat{V}$ is a martingale induced by solving a penalized "worst-case" problem in which the given specification of consumption dynamics is used as a benchmark model. The "robust" adjustment is made to this benchmark probability specification by solving minimization problem that penalizes changes in the probability law. For example, see Petersen et al. (2000) and Anderson et al. (2003).[26] Associated with $\hat{V}$ is a change in probability and this change alters the instantaneous interest rate and the local prices relative to model in which investors use discounted logarithmic utility to rank consumption. Thus my calculations show formally how investor concern about robustness induces approximately the same (as δ becomes small) long-term risk prices as a model in which investors are endowed with a power utility with relative risk aversion γ and no concern about robustness.

The preceding analysis exploits two important restrictions on investor preferences. The intertemporal elasticity of substitution is unity and the subjective rate of discount is zero. A natural extension is to compute two additional "derivatives" as a device to study the impact of changing investor preferences. For the long term risk prices, this can be done using the perturbation method described in section 7. Hansen et al. (2008) have used this method to explore changes in the intertemporal elasticity of substitution.[27]

The examples that I have described feature the role of investor preferences. A similar analysis applies to some models with market frictions. The solvency constraint models of Luttmer (1992), Alvarez and Jermann (2000) and Lustig (2007) have the same multiplicative martingale components as the corresponding representative consumer models without market frictions. While suggestive, a formal study along the lines of the type I have just presented for other models would reveal the precise nature of this transient adjustment to stochastic discount factors induced by solvency constraints and other forms of market imperfection.

9 Conclusion

I have contrasted two types of decompositions. The first has proved valuable as a device to aid in the identification of shocks with permanent consequences. It achieves this by extracting the martingale component in an additive decomposition. The second decomposition ia multiplicative instead of additive. I find it most useful in the study of the long-run implications of models of asset valuation. In particular, it shows how growth shocks influence economic value over long horizons. In this decomposition, the martingale component is used

[26]These papers explore stochastic perturbations in contrast to the original paper of Jacobson (1973) who developed the link to a deterministic version of robust control.

[27]In the case of the subjective rate of discount, the "derivative" will depend on which of the alternative models of investor preferences is entertained, recursive utility as in Kreps and Porteus (1978).

as a change in the probability measure and it is chosen to induce stochastic stability.

To conclude I want to be clear on two matters. First, while a concern about the role in economics in model specification is a prime motivator for this analysis, I do not mean to shift focus exclusively on the limiting characterizations. Specifically, my analysis of long-run approximation in this paper is not meant to pull discussions of transient implications off the table. Instead I mean to add some clarity into our understanding of how valuation models work by understanding better which model levers move which parts of the complex machinery. Moreover, I find the outcome of this analysis to be informative even if it reveals that some models *blur* the distinction between permanent and transitory model components.

Second, while my discussion of statistical approximation has been notably brief, I do not have to remind time series econometricians of the particular measurement challenges associated with the long run. Indeed there is a substantial literature on such issues. In part my aim is to suggest an econometric framework for the use of such measurements. But some of the measurement challenges remain. My own view is that many of the same statistical challenges that we as econometricians struggle with should be passed along to the hypothetical investors that populate our economic models. Difficulties in selecting a statistical model to use in extrapolation and associated ambiguities in inferences may well be an important component to the behavior of asset prices.

A A static max-min problem

In this appendix I develop further the static problem discussed in section 4.4 using results from the applied mathematics literature. Let $\mathcal{D}^+$ denote the strictly positive functions in $\mathcal{D}$, and let $\mathcal{Q}$ denote the family of probability measures Q on the state space $\mathcal{E}$ of the Markov process. Let $\mathbb{B}$ be the generator of the multiplicative semigroup. Following Donsker and Varadhan (1975), Donsker and Varadhan (1976) and Berestycki et al. (1994), I study the following max-min problem:

$$\varrho = \sup_{Q \in \mathcal{Q}} \inf_{f \in \mathcal{D}^+} \int \left(\frac{\mathbb{B} f}{f} \right) dQ. \tag{42}$$

Let $\mathbb{B}$ be the generator of the multiplicative semigroup. Split this generator into two components:

$$\mathbb{B} f(x) = \beta^*(x) f(x) + \mathbb{A} f(x)$$

where[28]

$$\begin{aligned} \beta^*(x) &\doteq \mathbb{B}1(x) \\ \mathbb{A} f(x) &\doteq \mathbb{B} f(x) - \beta^*(x) f(x). \end{aligned}$$

Notice that by construction $\mathbb{A} f = 0$ when f is a constant function. Suppose that $\mathbb{A}$ generates a semigroup of conditional expectations for a Markov processes. This requires additional restrictions, but these restrictions are effectively imposed on $\mathbb{B}$. I refer to β as the local growth or decay rate for the semigroup.

Consider the first the inner minimization problem of (42). Split the objective and write:

$$\inf_{f \in \mathcal{D}^+} \int \left(\beta^* + \frac{\mathbb{A} f}{f} \right) dQ.$$

Notice that the infimum over f does not depend on β^*. This in part leads Donsker and Varadhan (1975) and others to feature the optimization problem:

$$\mathbb{I}^*(Q) = \sup_{f \in \mathcal{D}_+} - \int \left(\frac{\mathbb{A} f}{f} \right) dQ \tag{43}$$

The function $\mathbb{I}^*$ is convex in Q since it can be expressed as the maximum of convex (in fact linear) functions of Q. Moreover, it can be justified as a relative measure of entropy

[28]While the function 1 does not vary over states the outcome applying $\mathbb{B}$ to 1 will typically vary with x and hence the notation $\mathbb{B}1(x)$.

between probabilities when the process implied by $\mathbb{A}$ possess a stationary distribution. The measure is relative because it depends on the generator $\mathbb{A}$ of a Markov process and measure discrepancies from the stationary distribution of this process.

I use the this representation of the solution to the inner problem to write the outer maximization problem as:

$$\sup_Q \left[\int \beta^* dQ - \mathbb{I}^*(Q) \right],$$

which is the problem posed in (17).

Suppose that the solution to the max-min problem is attained with probability measure Q^*. Consider again the inner optimization problem (43) and suppose that the supremum is attained at f^*. Let g be any other function in the domain of $\mathbb{B}$ such that $f^* + \mathsf{r}g$ is strictly positive for some r. For instance, if f^* is strictly positive and continuous, then it suffices that g be continuous, sufficiently smooth and have compact support in the interior of the state space. The first-order conditions are:

$$\int \left[\frac{\mathbb{A}g}{f^*} - \frac{g(\mathbb{A}f^*)}{(f^*)^2} \right] dQ^* = 0.$$

Let $f = \frac{g}{f^*}$, and rewrite this equation as:

$$\int \left[\frac{\mathbb{A}(f^* f)}{f^*} - \frac{f(\mathbb{A}f^*)}{f^*} \right] dQ^* = 0. \tag{44}$$

This first-order condition for r has a probabilistic interpretation. The operator

$$\begin{aligned} \mathbb{A}^* f &= \frac{\mathbb{A}(f^* f)}{f^*} - \frac{f(\mathbb{A}f^*)}{f^*} \\ &= \frac{\mathbb{B}(f^* f)}{f^*} - \frac{f(\mathbb{B}f^*)}{f^*} \end{aligned} \tag{45}$$

generates a distorted Markov process, and the first-order condition justifies Q^* as the stationary distribution of the distorted process.

To show the relation between the optimization problem and the principle eigenvalue problem, suppose that

$$\rho e = \mathbb{B}e$$

for e in $\mathcal{D}^+$. Construct a *twisted generator* using algorithm (45) with $f^* = e$, and suppose this generates a stochastically stable Markov process with stationary distribution Q^*. In particular, it satisfies (44). Notice that

$$\inf_{f\in\mathcal{D}^+}\int\left(\frac{\mathbb{B}f}{f}\right)dQ\leq\rho$$

because e is in $\mathcal{D}^+$ and ρ is an eigenvalue. Thus

$$\sup_{Q\in\mathcal{Q}}\inf_{f\in\mathcal{D}^+}\int\left(\frac{\mathbb{B}f}{f}\right)dQ\leq\rho.$$

When $Q=Q^*$, provided that e is the only solution to the inner minimization problem up to a scale factor, the upper bound is attained. As a consequence, $\rho=\varrho$ and this static problem gives an alternative construction of the principal eigenvalue.

B Derivative of a principal eigenfunction

In section 7 I showed how to compute the derivative of the principal eigenvalue. I now add to this discussion by providing the formula for the derivative of the principal eigenfunction. This derivative is useful in obtaining a more refined calculation. Recall that the principal eigenfunction is only defined up to scale. This leads me to study the derivative of logarithm of the principal eigenfunction evaluated at $\mathsf{a}=0$, denoted $D\log e$, which is well defined. I use the formula:

$$\hat{\mathbb{A}}f=\frac{1}{e}\mathbb{B}(ef)-\rho,$$

which gives the generator for the Markov process under the change of measure. Applying this formula to $f=\log e$ and differentiating the eigenfunction equation: $\mathbb{A}e=\rho e$ justifies

$$\hat{\mathbb{A}}(D\log e)(x)=\frac{d}{d\mathsf{a}}\rho\left[M(\mathsf{a})\right]|_{\mathsf{a}=0}-\frac{1}{e(x;0)}\left[\left.\frac{d}{d\mathsf{a}}\mathbb{B}(\mathsf{a})\right|_{\mathsf{a}=0}e(x;0)\right]$$

where $\mathbb{B}$ is the generator of the multiplicative semigroup when $\mathsf{a}=0$.

C Reconsidering the Campbell-Cochrane Model

The instantaneous interest for the Campbell and Cochrane (1999) model is:

$$-\lim_{t\downarrow 0}\frac{1}{t}\log E\left[S_t^*|X_0=x\right]=r-\gamma\xi(x-\mu_x)-\gamma^2\frac{\lambda(x)^2|\sigma_c|^2}{2}+\gamma^2|\sigma_c|^2\lambda(x),$$

which follows from (35) by setting $\sigma_g = 0$. They suppose the risk-rate is an affine function of the state: $r^* + \theta(x - \mu_x)$. Thus

$$r^* + \theta(x - \mu_x) = r + \gamma\xi(x - \mu_x) - \gamma^2 \frac{\lambda(x)^2 |\sigma_c|^2}{2} + \gamma^2 |\sigma_c|^2 \lambda(x). \tag{46}$$

I infer the value of r^* by setting $x = 0$:

$$r^* = r + (\theta - \gamma\xi)\mu_x$$

Substituting this formula into (46), by a simple complete-the-square argument:

$$(\theta - \gamma\xi)x - \frac{\gamma^2|\sigma_c|^2}{2} = -\frac{\gamma^2|\sigma_c|^2}{2} \left[\lambda(x) - 1\right]^2 .$$

Thus

$$\lambda(x) = 1 - (1 + \zeta x)^{1/2}$$
$$\zeta \doteq \frac{2(\gamma\xi - \theta)}{\gamma^2|\sigma_c|^2}$$

Campbell and Cochrane (1999) propose that the risk exposure of C_t^* be zero when $X_t = \mu_x$. The idea is that C_t^* is locally predetermined. To understand the ramifications of this, recall that

$$C_t^* = C_t - C_t \exp(-X_t - \mathsf{b})$$

where we now will determine the coefficient b. The coefficient b is important in quantifying risk aversion. The familiar measure of relative risk aversion is now state dependent and given by

$$\textbf{risk aversion} = \gamma \exp(X_t + \mathsf{b}).$$

The local risk exposure for C_t^* is

$$C_t[1 - \exp(-X_t - \mathsf{b})]\sigma_c dB_t + C_t \exp(-X_t - \mathsf{b})\lambda(X_t)\sigma_c dB_t.$$

Thus we require that

$$1 + \exp(-x - \mathsf{b})[\lambda(x) - 1] = 0, \tag{47}$$

or

$$1 - \lambda(x) = \exp(x + \mathsf{b})$$

We seek a value b, such that this equation is satisfied for $x = \mu_x$. Squaring the equation and multiplying by $\exp(-2\mu_x)$

$$\exp(-2\mu_x)\left(1+\left[\frac{2(\gamma\xi-\theta)}{\gamma^2|\sigma_c|^2}\right]\mu_x\right)=\exp(2\mathsf{b})$$

which determines b. At this value of b, the relative risk aversion measure is $\gamma[1-\lambda(\mu_x)]$ when $x=\mu_x$.

As an extra parameter restriction, they suggest requiring that the derivative of the risk exposure with respect to x be zero at x^*:

$$\exp(-\mu_x-\mathsf{b})[1-\lambda(\mu_x)]+\exp(-\mu_x-\mathsf{b})\lambda'(\mu_x)=0,$$

or

$$\frac{1}{2}\left([\lambda(\mu_x)-1]^2\right)'=\lambda'(\mu_x)[\lambda(\mu_x)-1]=[\lambda(\mu_x)-1]^2.$$

Thus

$$\frac{\gamma\xi-\theta}{\gamma^2|\sigma_c|^2}=1+\left[\frac{2(\gamma\xi-\theta)}{\gamma^2|\sigma_c|^2}\right]\mu_x,$$

which is restriction on the underlying parameters. Specifically,

$$\mu_x=\frac{1}{2}-\frac{\gamma^2|\sigma_c|^2}{2(\gamma\xi-\theta)}$$

Notice that we may now express λ as:

$$\begin{aligned}\lambda(x)-1 &= -\left(1+\left[\frac{2(\gamma\xi-\theta)}{\gamma^2|\sigma_c|^2}\right]x\right)^{1/2}\\ &= -\left(\frac{\gamma\xi-\theta}{\gamma^2|\sigma_c|^2}+\left[\frac{2(\gamma\xi-\theta)}{\gamma^2|\sigma_c|^2}\right](x-\mu_x)\right)^{1/2}\\ &= -\left(\frac{\gamma\xi-\theta}{\gamma^2|\sigma_c|^2}\right)^{1/2}\left[1+2(x-\mu_x)\right]^{1/2}.\end{aligned}$$

as derived in Campbell and Cochrane (1999).

Finally, we consider the change measure implied by the martingale:

$$\tilde{M}_t=\exp(rt)S_t\frac{\exp(\gamma X_t)}{\exp(\gamma X_0)}$$

It implies an alternative distorted evolution:

$$\begin{aligned}dX_t &= \left[-\xi(X_t-\mu_x)-\gamma\lambda(X_t)|\sigma_c|^2\right]dt+\lambda(X_t)\sigma_c d\hat{W}_t\\ &= \left[-\xi(X_t-\mu_x)-\gamma\lambda(X_t)|\sigma_c|^2+\gamma\lambda(X_t)^2|\sigma_c|^2\right]dt+\lambda(X_t)\sigma_c d\tilde{W}_t\end{aligned}$$

$$
\begin{aligned}
&= \left[-\xi(X_t - \mu_x) + \gamma(1+\zeta X_t)|\sigma_c|^2 - \gamma(1+\zeta X_t)^{1/2}|\sigma_c|^2\right] dt + \lambda(X_t)\sigma_c d\tilde{W}_t \\
&= \left[\xi X_t + \xi\mu_x + \gamma|\sigma_c|^2 - \gamma(1+\zeta X_t)^{1/2}|\sigma_c|^2\right] dt + \lambda(X_t)\sigma_c d\tilde{W}_t
\end{aligned}
$$

where

$$
d\hat{W}_t = \gamma\lambda(X_t)^2\sigma_c dt + d\tilde{W}_t
$$

and $\tilde{W}_t$ is a standard Brownian increment under the probability measure implied by $\tilde{M}$. Given the strong pull of the drift to the right for large X_t, this evolution results in unstable stochastic dynamics.

D Finite Horizons

Consider Example 3.4 and continued in Example 6.2. The additive functional is:

$$
dY_t = \nu dt + H_1 X_t^{[1]} dt + H_2(X_t^{[2]} - 1)dt + \sqrt{X_t^{[2]}} F dW_t.
$$

Form

$$
M_t = \exp(Y_t).
$$

My aim is to compute

$$
\mathbb{M}_t 1(x) = E\left[M_t | X_0 = x\right]
$$

where the left-hand side notation reflects the fact that operator is evaluated at the unit function and this evaluation depends on the state x. I use the following formula for this computation.

$$
\mathbb{B}\mathbb{M}_t f = \frac{d}{dt}\left[\mathbb{M}_t f(x)\right] \tag{48}
$$

Guess a solution

$$
\mathbb{M}_t 1(x) = E\left[M_t | X_0 = x\right] = \exp\left[\alpha(t)\cdot x + \varrho(t)\right]
$$

where $x = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$ and $\alpha(t) = \begin{bmatrix} \alpha_1(t) \\ \alpha_2(t) \end{bmatrix}$. Notice that

$$
\mathbb{B}\exp\left[\alpha(t)\cdot x + \varrho(t)\right] = \exp\left[\alpha(t)\cdot x + \varrho(t)\right]\left(\left[\frac{d}{dt}\alpha(t)\right]\cdot x + \frac{d}{dt}\varrho(t)\right).
$$

Moreover,

$$
\begin{aligned}
\frac{\mathbb{B}\exp\left[\alpha(t)\cdot x + \varrho(t)\right]}{\exp\left[\alpha(t)\cdot x + \varrho(t)\right]} &= \nu + H_1 x_1 + H_2(x_2 - 1) \\
&\quad + {x_1}' {A_1}' \alpha_1(t) + (x_2 - 1)A_2\alpha_2(t)
\end{aligned}
$$

$$+\frac{x_2}{2}|F + \alpha_1(t)'B_1 + \alpha_2(t)B_2|^2.$$

First use (48) to produce a differential equation for $\alpha_1(t)$:

$$\frac{d}{dt}\alpha_1(t) = {H_1}' + {A_1}'\alpha_1(t) + A_2\alpha_2(t).$$

by equating coefficients on x_1. This differential equation has as its initial condition $\alpha_1(0) = 0$. Similarly, by equation coefficient on x_2,

$$\frac{d}{dt}\alpha_2(t) = H_2 + \frac{1}{2}|F + \alpha_1(t)'B_1 + \alpha_2(t)B_2|^2$$

This uses the solution for $\alpha_1(t)$ as an input. The initial condition is $\alpha_2(0) = 0$. Finally,

$$\frac{d}{dt}\varrho(t) = \nu - H_2 - A_2\alpha_2(t).$$

This is a differential equation for $\varrho(t)$ given the solution for $\alpha_2(t)$ The initial condition is $\varrho(0) = 0$. Example 6.2 has formulas for the limiting values of α_1 and α_2 as t becomes large. The function ϱ will eventually grow linearly.

References

Alvarez, Fernando and Urban Jermann. 2000. Efficiency, Equilibrium, and Asset Pricing with Risk of Default. *Econometrica* 775–798.

Anderson, Evan W., Lars P. Hansen, and Thomas J. Sargent. 2003. A Quartet of Semigroups for Model Specification, Robustness, Prices of Risk, and Model Detection. *Journal of the European Economic Association* 1:69–123.

Balaji, S. and S. P. Meyn. 2000. Multiplicative Ergodicity and Large Deviations for an Irreduciable Markov Chain. *Stochastic Processes and their Applications* 90:123–144.

Bansal, R. and B. N. Lehmann. 1997. Growth-optimal Portfolio Restrictions on Asset Pricing Models. *Macroeconomic Dynamics* 1 (2):333–354.

Bansal, R. and A. Yaron. 2004. Risks for the Long Run: A Potential Resolution of Asset Pricing Puzzles. *Journal of Finance* 59 (4):1481–1509.

Berestycki, H., L. Nirenberg, and S. R. S. Varadhan. 1994. The Principal Eigenvalue and the Maximum Principal for Second-Order Elliptic Operators in General Domains. *Communications in Pure and Applied Mathematics* 47:47–92.

Beveridge, S. and C. R. Nelson. 1981. A New Approach to the Decomposition of Economic Time Series into Permanent and Transitory Components with Particular Attention to the Measurement of the 'Business Cycle'. *Journal of Monetary Economics* 7 (2):151–174.

Bhattacharya, R. N. 1982. On the Functional Central Limit Theorem and the Law of the Iterated Logarithm. *Zeitschrift für Wahrscheinlichkeitstheorie und verwandte Gebiete* 60:185–201.

Billingsley, P. 1961. The Lindeberg-Levy Theorem for Martingales. *American Mathematical Monthly* 12.

Blanchard, O. J and D. Quah. 1989. The Dynamic Effects of Aggregate Demand and Supply Disturbances. *American Economic Review* 79 (4):655–673.

Box and Tiao. 1977. Canonical Analysis of Multiple Time Series. *Biometrika* 64:355–366.

Breeden, D. 1979. An Intertemporal Asset Pricing Model with Stochastic Consumption and Investment Opportunities. *Journal of Financial Economics* 7 (3):265–296.

Campbell, J. and J. Cochrane. 1999. By Force of Habit. *Journal of Political Economy* 107 (2):205–251.

Campbell, J. and T. Vuolteenaho. 2004. Bad Beta, Good Beta. *American Economic Review* 94 (5):1249–1275.

Chernoff, Herman. 1952. A Measure of Asymptotic Efficiency for Tests of a Hypothesis Based on the Sum of Observations. *Annals of Mathematical Statistics* 23:493–507.

Cochrane, J. H. 2001. *Asset Pricing*. Princeton, N.J.: Princeton University Press.

Donsker, Monroe E. and S. R. S. Varadhan. 1975. On a Variational Formula for the Principal Eigenvalue for Operators with Maximum Principle. *Proceedings of the National Academy of Sciences* 72 (3):780–783.

———. 1976. On the Principal Eigenvalue of Second-Order Elliptic Differential Equations. *Communications in Pure and Applied Mathematics* 29:595–621.

Dupuis, P. and R. S. Ellis. 1997. *A Weak Convergence Approach to the Theory of Large Deviations*. Wiley Series in Probability and Statistics. John Wiley and Sons.

Engle, R. and C. W. J. Granger. 1987. Co-integration and Error Correction: Representation, Estimation and Testing. *Econometrica* 55:251–276.

Epstein, L. and S. Zin. 1989. Substitution, Risk Aversion and the Temporal Behavior of Consumption and Asset Returns: A Theoretical Framework. *Econometrica* 57 (4):937–969.

Gordin, M. I. 1969. The Central Limit Theorem for Stationary Processes. *Soviet Mathematics Doklady* 10:1174–1176.

Hansen, L. P. and S. F. Richard. 1987. The Role of Conditioning Information in Deducing Testable Restrictions Implied by Dynamic Asset Pricing Models. *Econometrica* 50:587–614.

Hansen, L. P. and T. J. Sargent. 1995. Discounted Linear Exponential Quadratic Gaussian Control. *IEEE Transactions on Automatic Control* 40.

Hansen, L. P. and J. Scheinkman. 1995. Back to the Future: Generating Moment Implications for Continuous Time Markov Processes. *Econometrica* 63:767–804.

———. 2007. Long-Term Risk: An Operator Approach. The University of Chicago and Princeton University.

Hansen, L. P., J. C. Heaton, and N. Li. 2008. Consumption Strikes Back?: Measuring Long Run Risk. *Journal of Political Economy* forthcoming.

Hansen, Lars Peter. 2007. Beliefs, Doubts and Learning, Valuing Macroeconomic Risk. *American Economic Review* 97 (2).

Hansen, Lars Peter, Thomas J. Sargent, Guahar A. Turmuhambetova, and Noah Williams. 2006. Robust Control and Model Misspecification. *Journal of Economic Theory* 128:45–90.

Harrison, J. and D. Kreps. 1979. Martingales and Arbitrage in Multiperiod Security Markets. *Journal of Mathematical Economics* 20:381–408.

Ito, K. and S. Watanabe. 1965. Transformation of Markov Processes by Multiplicative Functionals. *Annals of Institute Fourier* 15:13–30.

Jacobson, David H. 1973. Optimal Stochastic Linear Systems with Exponential Performance Criteria and Their Relation to Deterministic Differential Games. *IEEE Transactions for Automatic Control* AC-18:1124–131.

Kocherlakota, Narayana R. 1990. Disentangling the Coefficient of Relative Risk Aversion from the Elasticity of Intertemporal Substitution: An Irrelevance Result. *The Journal of Finance* 45 (1):175–190.

Kontoyiannis, I. and S. P. Meyn. 2003. Spectral Theory and Limit Theorems for Geometrically Ergodic Markov Processes. *Annals of Applied Probability* 13:304–362.

Kreps, D. M. and E. L. Porteus. 1978. Temporal Resolution of Uncertainty and Dynamic Choice. *Econometrica* 46 (1):185–200.

Lettau, M. and J. Wachter. 2007. Why is Long-Horizon Equity Less Risky? A Duration-Based Explanation of the Value Premium. *Journal of Finance* 62:55–92.

Lustig, Hanno. 2007. The Market Price of Aggregate Risk and the Wealth Distribution. Working paper, UCLA.

Luttmer, E. 1992. *Asset Pricing in Economies with Frictions.* Ph.D. thesis, University of Chicago.

Meyn, S. P. and R. L. Tweedie. 1993. Stability of Markovian Processes III: Foster-Lyapunov Criteria for Continuous Time Processes. *Advances in Applied Probability* 25:518–548.

Newell, Richard G. and William A. Pizer. 2003. Discounting the Future: How Much Do Uncertain Rates Increase Valuations. *Journal of Environmental Economics and Management* 46:52–71.

———. 2004. Uncertain Discount Rates in Climate Policy Analysis. *Energy Policy* 32:519–529.

Newman, Charles M. and Barton W. Stuck. 1979. Chernoff Bounds for Discriminating Between Two Markov Processes. *Stochastics* 2:139–153.

Pearson, K. 1916. Mathematical Contributions to the Theory of Evolution, XIX: Second Supplement to a Memoir on Skew Variation. *Philosophical Transactions of the Royal Society of London. Series A* 216:429457.

Petersen, Ian R., Matthew R. James, and Paul Dupuis. 2000. Minimax Optimal Control of Stochastic Uncertain Systems with Relative Entropy Constraints. *IEEE Transactions on Automatic Control* 45:398–412.

Rosenblatt, M. 1971. *Markov Processes: Structure and Asymptotic Behavior*. Berlin: Springer-Verlag.

Runolffson, T. 1994. Optimal Control of Stochastic System with an Exponential-of-integral Performance Criterion. *Systems and Control Letters* 22:451 – 456.

Santos, T. and P. Veronesi. 2006. Habit Formation, the Cross Section of Stock Returns and the Cash-Flow Risk Puzzle. Columbia University and the University of Chicago.

Singleton, K. J. 2006. *Empirical Dynamic Asset Pricing: Model Specification and Econometric Assessment*. Priceton, N.J.: Princeton University Press.

Tallarini, Thomas D. 2000. Risk-sensitive real business cycles. *Journal of Monetary Economics* 45 (3):507–532.

Wachter, Jessica. 2005. Solving Models with External Habit. *Finance Research Letters* 2:210–226.

Whittle, P. 1990. *Risk Sensitive and Optimal Control*. West Suffix, England: John Wiley and Sons.

Wong, E. 1964. The Construction of a Class of Stationary Markoff Processes. In *Stochastic Processes in Mathematical Physics and Engineering*, edited by R. E. Bellman, no. 17 in Proceedings of Symposia in Applied Mathematics, 264–276. American Mathematical Society.