Modeling the Long Run:
Valuation in Dynamic Stochastic Economies[1]

by

Lars Peter Hansen
University of Chicago and NBER
lhansen@uchicago.edu

August 5, 2008

Abstract

I explore the equilibrium value implications of economic models that incorporate reactions to a stochastic environment. I propose a dynamic value decomposition (DVD) designed to distinguish components of an underlying economic model that influence values over long horizons from components that impact only the short run. To quantify the role of parameter sensitivity and to impute long-term risk prices, I develop an associated perturbation technique. Finally, I use DVD methods to study formally some example economies and to speculate about others. 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 used to identify shocks with long-run economic consequences.

[1]Prepared for the Fisher-Schultz Lecture at the 2006 European Meetings of the Econometric Society. I appreciate helpful comments by Emilio Barucci, Jaroslav Borovicka, Buz Brock, Valentin Haddad, Mark Hendricks, Eric Janofsky, Jim Marrone, Nour Meddahi, Monika Piazzesi, Phil Reny, Chris Rogers, Jose Scheinkman, Grace Tsiang, Harald Uhlig, Jessica Wachter and Neng Wang, Amir Yaron. This paper is very much influenced by my joint work with John Heaton, Nan Li and Jose Scheinkman on this topic and by related work I have done with Xiaohong Chen and Tom Sargent. This material is based upon work supported by the National Science Foundation under Award Number SES0519372.

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-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 of 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. (2008a) and Bansal 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. Given my focus on valuation, 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 modifications 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 Overview

There are a variety of applications of the dynamic value decomposition (DVD) methods that I describe. They can be used to

i) construct model-based measures of consumption or cash-flow duration;

ii) construct risk-adjusted measures of long-term premia;

iii) characterize long-term risk-exposure dynamics;

iv) characterize of the long-term risk-price dynamics;

v) make local and global model comparisons.

A valuation model assigns prices to consumption processes or cash-flows for each hypothetical payoff horizon. Cash flows that make substantial contributions to value far into the future are said to have *high duration*. This duration depends on how the cash flow grows and how that cash flow is discounted. There is a premia for a cash flow at each horizon relative to a riskless counterpart. The risk premia depend on the risk exposures (the cash flow's dependence on the underlying shocks) and on the risk prices (the marginal compensation to investors that bear the risk associated with these shocks). In my applications I feature long-term risk-price dynamics and model comparisons.

My characterization of *risk-price dynamics* is based on a valuation counterpart to the impulse-response functions featured in macroeconomic dynamics. I construct a *risk-price trajectory* by computing a derivative for each payoff horizon, which is the marginal change in a risk premia induced by a marginal change in risk exposure of a cash flow constructed from underlying macroeconomic shocks. To feature risk-price dynamics while abstracting from *risk-exposure dynamics*, I value martingale cash flows built from alternative shocks. By design, the expected future cash-flows are independent of the prediction horizon.

In one of my applications in section 7, I compare the risk price implications for a model in which investors have power utility to a counterpart model with recursive utility using a risk-sensitive parameterization of preferences. For analytical simplicity I restrict the elasticity of intertemporal substitution of the investors to be unity for this second model.

The investors in this model care about the intertemporal composition of risk in contrast to those in the first model. I show how the limiting risk prices are the same for the two models by setting appropriately the parameters of investor preferences.

Following Bansal and Yaron (2004), I consider a model with predictability in consumption growth rates. While I consider a three shock specification, the analysis of the shock to the consumption growth rate is particularly revealing. Such a shock has an impact that builds gradually as is reflected in the continuous-time impulse response function for the logarithm of consumption depicted in the top panel of figure 1. When investors have time-separable power utility, the risk-price trajectory also builds gradually over the alternative investment horizons as depicted by the dashed line in the lower panel of figure 1. It converges to a limit that I will characterize. The convergence is slower than for the impulse response function. In this example economy, the impulse-response function converges exponentially. In contrast, I will show that the risk-price trajectory displays a hyperbolic convergence. Notice that the risk price trajectory eventually becomes almost coincident with the hyperbolic curve depicted by lower solid line in figure 1. I will provide formal characterizations of both the limit point of the risk-price trajectory and the hyperbolic curve that approximates the trajectory as the payoff horizon becomes long. While much of asset pricing literature focuses on short-term or local risk prices, the methods in this paper reveal pricing features for longer payoff horizons.

In addition to exposing the pricing dynamics for a given model, DVD methods facilitate model comparisons. Investor preferences are forward looking in the recursive utility model, and this is reflected in a larger and flatter risk-price trajectory depicted as the solid line in figure 1. Even though the model with power utility investors has the same long-term limiting risk price, its risk-price trajectory starts near zero and only approximates its limit for extremely long payoff horizons.[2]

1.2 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

[2] Details about computations including the parameter choices are given in section 7 and appendix C

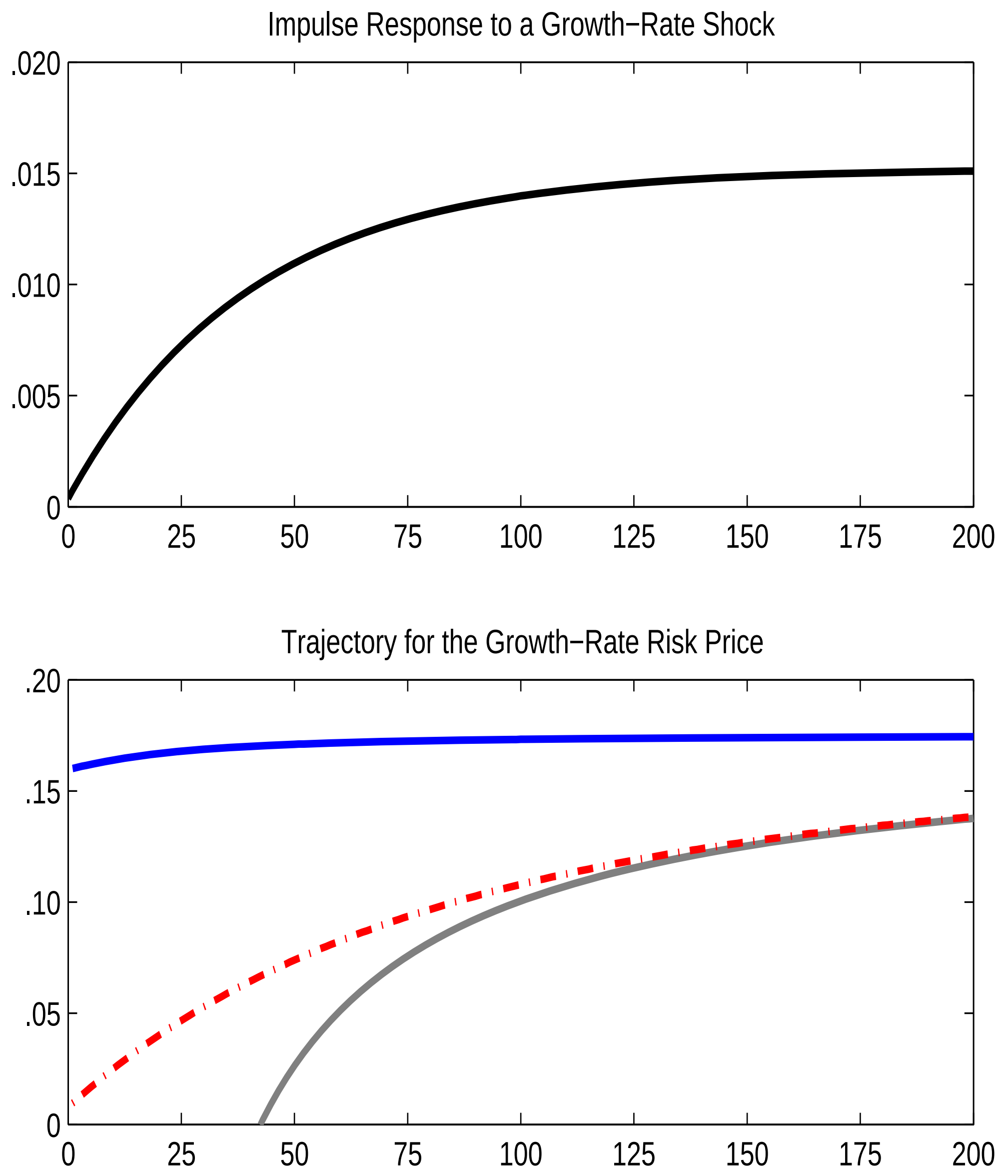

Figure 1: Impulse responses and risk price trajectories for the growth-rate shock. The horizontal axis is given in quarterly time units. The top panel gives the impulse-responses of the logarithm of consumption using the same parameter values as in figure 4. The bottom panel reproduces risk prices depicted in figure 4. The lower solid curve in the bottom panel is the trajectory for a model with power utility investors, the upper solid curve is for a model with recursive utility investors, and the dashed curve is the hyperbolic approximation to this trajectory.

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:

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

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* (section 5): I measure long-run growth and the associated decay in value 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):

$$
\begin{array}{cccccc}
M_t & = & \exp(\rho t) & \hat{M}_t & \left[\frac{e(X_0)}{e(X_t)}\right] & \\
 & & \uparrow & \uparrow & \uparrow & \quad (2) \\
 & & \textbf{exponential trend} & \textbf{martingale} & \textbf{ratio}. &
\end{array}
$$

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, factorization (2) is distinct from (1) because the exponential of a martingale is not itself a martingale. In performing calculations, I use the martingale to change the underlying probability measure in a way that supports a convenient characterization of long-run behavior in valuation. In my applications, M's will be constructed from explicit economic models and hypothetical changes in stochastic growth trajectories.

v) *Sensitivity and long-run pricing* (section 6): Of special interest is how the long-run attributes of valuation change when I make small alterations in a) growth processes or b) stochastic discount factors used to represent valuation. I show formally how to conduct a sensitivity analysis with these two applications in mind. The applications I feature show how marginal changes in the risk exposure of hypothetical growth processes alter risk premia. This calculation gives me an operational way to construct risk prices as a function of the investment horizon. Analogous calculations applied to the stochastic discount factor show how long-term 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 7): 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 sidestep some numerical issues by exploring specifications for which there are quasi-analytical characterizations. In the applications I produce *pricing* counterparts to impulse functions familiar in literature on economic dynamics. Impulse responses characterize how economic variables respond over time to the underlying shocks. In contrast, I assess the valuation consequences for consumption trajectories or cash flows that are "exposed" to macroeconomic shocks. I assign risk prices to shocks at alternative investment horizons by valuing hypothetical cash flows constructed from the respective macroeconomic shocks. Such price calculations require economic models that are "structural" to give me the necessarily latitude to extrapolate value implications. The methods accommodate nonlinearities present in the underlying Markov dynamics. The pricing consequences of these nonlinearities are revealed by the methods I describe.

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.[3]

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, $\epsilon \int \eta(dx^*|X_{t-})$ is the approximate probability

[3] Alternative and distinct approaches to characterizing asset price dynamics with similar motivations can be found in Daniel and Marshall (2005), Dybvig et al. (1996) and Otrok et al. (2002).

that there will be a jump for small time interval ϵ. 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 I will next show, 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 be interested in other specifications. Conveniently, the sum of two additive functionals is additive.

I consider a family of such functionals parameterized by (β, ξ, χ) 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) $\chi : \mathcal{E} \times \mathcal{E} \to \mathbb{R}$, $\chi(x, x) = 0$.

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

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 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 $\chi(X_t, X_{t-})$. The term $\sum_{0 \le u \le t} \chi(X_u, X_{u-})$ thus reflect the impact of "large shocks". This component is not necessarily a martingale because the jumps may have a predictable component. The integral

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

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

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

I will be primarily interested in specifications of χ for which this constructed process is a martingale. In summary, an additive functional 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.[4] 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 $t < u \le t + \tau$.[5] 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

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

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

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).[6] 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 multiplicative functionals, I consider the decomposition of an additive functional. My initial investigation of additive functionals is consistent with the common practice of building models that apply to logarithms of macroeconomic or financial times series. While there are alternative ways to decompose time series, what follows is closest to what I will be interested in. An additive functional can be decomposed into three components:

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

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 gives the continuous-time counterpart to

[6]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.

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*

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

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 the X process between time 0 and time t. As a consequence of the Law of Iterated Expectations and the mean-square convergence,

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

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):

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

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

$$\tilde{Y}_t = \int_0^t \xi(X_u) \cdot dW_u + \sum_{0 \le u \le t} \chi(X_u, X_{u-}) - \int_0^t \tilde{\beta}(X_u)du \tag{7}$$

is a local martingale. In what follows let

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

I now have the ingredients for the following result.

Theorem 3.2. *Suppose*

i) X is a stationary, ergodic Markov process;

ii) $\tilde{Y}$ given in (7) is a square integrable martingale;

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

iv) There is a solution g to

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

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

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,

$$\hat{\beta}(x) - \nu + \lim_{t \downarrow 0} \frac{1}{t} E\left[g(X_t) - g(x)|X_0 = x\right] = 0, \tag{8}$$

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 a 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, in light of (8) the function g satisfies

$$\mathbb{A} g = \nu - \hat{\beta}. \tag{9}$$

For the diffusion model, this leads to solving:

$$\frac{\partial g(x)}{\partial x} \cdot \mu(x) + \frac{1}{2} \text{trace}\left[\sigma(x)\sigma(x)' \frac{\partial^2 g(x)}{\partial x \partial x'}\right] = \nu - \hat{\beta}(x). \tag{10}$$

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 (9) 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.[7]

Following Gordin (1969), by extracting a martingale we can produce a more refined

[7] These references suppose that X is stationary. Hansen and Scheinkman (1995) use an L^2 notion of 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 (9).

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.[8] 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 a multivariate standard Brownian 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 (10). 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,

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

is the martingale component.

[8]See Billingsley (1961) for a discrete-time martingale central limit theorem. Moreover, there are well known functional extensions of this result. For instance, see Hall and Heyde (1980).

Next I consider a model with stochastic volatility.

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

$$\begin{aligned} dX_t^{[1]} &= A_{11}X_t^{[1]}dt + A_{12}(X_t^{[2]} - 1) + \sqrt{X_t^{[2]}}B_1 dW_t, \\ dX_t^{[2]} &= A_{22}(X_t^{[2]} - 1)dt + \sqrt{X_t^{[2]}}B_2 dW_t \\ dY_t &= \nu dt + H_1 X_t^{[1]}dt + H_2(X_t^{[2]} - 1)dt + \sqrt{X_t^{[2]}}F dW_t. \end{aligned}$$

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_{11} has eigenvalues with strictly negative real parts and A_{22} is negative. Moreover, to prevent zero from being attained by $X^{[2]}$, I assume that $A_{22} + \frac{1}{2}|B_2|^2 < 0$. 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:

$$\frac{\partial g(x_1,x_2)}{\partial x} \cdot \begin{bmatrix} A_{11}x_1 + A_{12}(x_2 - 1) \\ A_{22}(x_2 - 1) \end{bmatrix} + \frac{x_2}{2}\text{trace}\left(\begin{bmatrix} B_1 B_1{}' & 0 \\ 0 & |B_2|^2 \end{bmatrix} \frac{\partial^2 g(x_1,x_2)}{\partial x \partial x'} \right) = -H_1 x_1 - H_2(x_2 - 1),$$

which is a special case of (10). The solution is:

$$g(x_1,x_2) = -H_1(A_{11})^{-1}x_1 - \left[H_2 - H_1(A_{11})^{-1}A_{12}\right](A_{22})^{-1}(x_2 - 1).$$

The local innovation in $g(X_t)$ is $\sqrt{X_t^{[2]}}\left[-H_1(A_{11})^{-1}B_1 - \left[H_2 - H_1(A_{11})^{-1}A_{12}\right](A_{22})^{-1}B_2\right] dW_t$. Thus in this example the martingale component for Y is given by:

$$\hat{Y}_t = \int_0^t \sqrt{X_u^{[2]}}\left[F - H_1(A_{11})^{-1}B_1 - \left[H_2 - H_1(A_{11})^{-1}A_{12}\right](A_{22})^{-1}B_2\right] dW_u.$$

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.[9]

Even when such nonlinearity is introduced, conveniently the sum of two additive functionals is an additive functional. Moreover, the sum of the martingale components is the martingale component for the sum of the additive functionals provided that the constructions use a common information structure.

As I have already argued, the additive decomposition is valuable for identifying shocks with durable consequences or for characterizing long-run consequences of shocks. Additive decompositions have direct ties to the study of log-linear relations, either exact or approximate, are convenient for many purposes. For the purposes of valuation, in what follows I will use multiplicative functionals. Such functionals can be represented conveniently as the exponentials of additive functionals. One strategy at my disposal is first to decompose additive functionals and then to exponentiate the components. 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. Except in degenerate examples, 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 via a different approach that will be of direct use.[10]

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 this section I describe the relation between a *local* growth rate of a multiplicative functional M and its *long-term* or asymptotic counterpart. The local growth rate is defined

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

[10]While the additive decomposition is linked to a Law of Large Numbers and a Central Limit Theorem, this alternative decomposition has much closer ties to the Theory of Large Deviations.

as:

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

provided that this limit exists. Since $M_t = \exp(Y_t)$ and

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

as in (3), the local growth rate is computed to be

$$\beta^*(x) = \beta(x) + \frac{1}{2}|\xi(x)|^2 + \int \left(\exp[\chi(y, \cdot)] - 1\right) \eta(dy|x) \tag{11}$$

using, when necessary, continuous-time stochastic calculus. Notice that direct exposure to Brownian motion risk and jump risk contributes to this local growth rate.

By contrast, define the 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. Compounding has nontrivial consequences for the long-term growth rate when the local growth rate is state-dependent. I will characterize this asymptotic limit and explore the relation between the average local growth rate and the asymptotic growth rate. 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 discount rates that are state dependent. This topic has been explored in the study of climate policy (see for instance Newell and Pizer (2003, 2004)) and in the study of the long-term behavior of stochastic discount factors (see for instance Alvarez and Jermann (2000)).

4.1 A revealing special case

Prior to a more general 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 its origins from what is known as Perron-Frobenius theory of matrices.

Example 4.1. *When a Markov process has 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 state X and the jump distribution (conditioned on a jump, where the process will 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*

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

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

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

where $\times$ used in the matrix multiplications denotes element-by-element multiplication. This construction of $\mathbb{B}$ 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 $\mathbb{A}\mathbf{1}_n$. 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 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 $\{\mathbb{M}_t : t \geq 0\}$. Specifically, the vector

$$\beta^* = \mathbb{B}\mathbf{1}_n$$

contains the state dependent growth rates defined more generally in formula (11).

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

$$\mathbb{B}'\mathbf{e}^* = \rho \mathbf{e}^* \tag{12}$$

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,

$$\lim_{t \to \infty} \frac{1}{t} \log \mathbb{M}_t \mathbf{f} = \rho \tag{13}$$

$$\lim_{t \to \infty} (\log \mathbb{M}_t \mathbf{f} - t\rho) = \log(\mathbf{f} \cdot \mathbf{e}^*)\mathbf{1}_n + \log \mathbf{e} \tag{14}$$

for any vector $\mathbf{f}$ for which $\mathbf{f} \cdot \mathbf{e}^* > 0$, where I have normalized $\mathbf{e}^*$ so that $\mathbf{e}^* \cdot \mathbf{e} = 1$. Formally, (13) defines ρ as the long-run growth rate of the family of matrices $\{\mathbb{M}_t : t \geq 0\}$. In the remainder of this section, I develop and apply a generalization of this limit. After adjusting for this growth rate, (14) gives an example of a more refined approximation that I explore in section 5. The logarithm of the eigenvector $\mathbf{e}$ exposes the impact of state-dependent compounding of growth or discounting over long horizons.

To accommodate continuous Markov states, I will use operators in place of the matrices. These operators are introduced in the next subsection section. They have a complicated eigenvalue structure because I allow a more general specification of the underlying Markov process. For instance, there may be multiple eigenvalues associated with distinct positive eigenfunctions. Typically at most one of these eigenvalue-eigenfunction pairs is of interest to us. As in the finite-state example, 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:

$$\mathbb{M}_t f(x) = E\left[M_t f(X_t) | X_0 = x\right] \tag{15}$$

indexed by t. When M has finite first moments, this family of operators is at least well defined on the space of bounded functions.[11]

I use alternative constructions of M where is sometimes the product of multiple components. The stochastic process components have explicit economic interpretations including stochastic discount factor processes, macroeconomic growth trajectories, or growth processes used to represent hypothetical financial claims to be priced. My use of stochastic discount factor processes to reflect valuation is familiar from empirical asset pricing. For instance, see Hansen and Richard (1987), Cochrane (2001), and Singleton (2006). A stochastic discount factor for a given payoff horizon discounts the future and it adjusts for risk when used to assign values to a future payoff. A stochastic discount factor *process* assigns values to a cash flow *process* such as a consumption process that will be realized in future dates or a dividend process on an infinitely-lived security. This process typically decays asymptotically. Such decay is needed for an infinitely-lived security 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 to study the dynamics of valuation in the presence of stochastic growth.

[11] See Hansen and Scheinkman (2008) for a a more general and explicit formulation of the domain of such operators.

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 of 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:

$$\mathbb{S}_t \mathbb{S}_\tau = \mathbb{S}_{t+\tau} \tag{16}$$

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 (16) 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 (16).[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:

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

I will show that 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.

[12] Lewis (1998), Linetsky (2004) and Boyarchenko and Levendorskii (2007) use operator methods of this type to study of the term structure of interest rates and option pricing.

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

Definition 4.2. *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 (15) 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 τ,

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

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

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

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 $\mathbb{P}_t$: $M = SG$. Exploiting the recursive and time invariant nature of Markov pricing, the value assigned to $C_{t+\tau}$ at time τ is $G_\tau \tilde{G}_0 \mathbb{P}_t f(x)$.

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,

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

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

$$\rho\left(M^{[1]}M^{[2]}\right) - \rho\left(M^{[1]}\right) - \rho(M^{[2]})\,. \tag{17}$$

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 (17) 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.

Here is an application of this apparatus to a dynamic extension of a familiar asset pricing problem. The risk premium on a cash-flow $G_t f(X_t)$ paid out at date t and valued at date zero is measured by:

$$\begin{array}{ccccc} \frac{1}{t}\log E\left[G_t f(X_t)|X_0 = x\right] & - & \frac{1}{t}\log E\left[S_t G_t f(X_t)|X_0 = x\right] & + & \frac{1}{t}\log E\left[S_t|X_0 = x\right] \\ \textbf{log} & & \textbf{log} & & \textbf{log} \\ \textbf{expected payoff} & - & \textbf{price} & - & \textbf{riskfree return} \end{array} \tag{18}$$

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, I reproduce the local risk premia familiar in the continuous-time asset pricing literature. In contrast, the purpose of the methods described in this paper is 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

$$\textbf{risk premium} = \rho(G) + \rho(S) - \rho(SG). \tag{19}$$

In the log-normal case this limiting risk premia will turn out to be the negative covariance of the increments in the martingale components of $\log G$ and $\log S$ (see calculations in example 5.2 in subsection 5.2).

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), Hansen et al. (2008a) and Bansal et al. (2008). Hansen et al. (2008a) characterize the resulting limiting risk premia and the associated risk prices in a log-linear environment.[13] Hansen and Scheinkman (2008) extend this approach to fundamentally nonlinear models with a Markov structure. The perturbation method of section 6 gives a way to compute risk prices in nonlinear Markov environment.[14] Later I will extend 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 averages the local state dependent growth rate: $\nu = E\left[\beta^*(X_t)\right]$. I now explore the relation between the local, state dependent growth rate and the long-run counterpart $\rho(M)$.

Consider for the moment a special class of multiplicative functionals:

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

[13]Hansen et al. (2008a) 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.

[14]Wachter (2005) develops a computational approach based on pricing what she calls "zero-coupon" equity, which in our notation is $E\left[S_tG_tf(X_t)|X_0 = x\right]$. Her algorithm has component prices that converge 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 by providing an approximation for the infinite sum of terms starting from a given horizon forward. The results of Hansen and Scheinkman (2008) described in section 5 provide justification for the limit approximation.

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 η, but Jensen's inequality prevents us from just exponentiating this average to compute $\rho(M)$.

The limit $\rho(M)$ in this special case is a key ingredient in the study of large deviations. While $\frac{1}{t}\int_0^t \beta(X_u)du$ obeys a Law of Large Numbers and converges to its unconditional expectation η, 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 asymptotic 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* $\mathbb{J}(Q)$ 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 $\mathbb{J}$ is convex in the probability measure Q, and it is used to construct what is called a Laplace principle that characterizes the limit:

Problem 4.3.

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

The inequality follows because $\mathbb{J}(Q) = 0$ 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 $\mathbb{J}$. The function $\mathbb{J}$ 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 (4.3) characterizes formally this difference.[15]

[15] Large deviation theory exploits problem (4.3) because $\rho(M)$ implies a bound of the form:

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

This analysis applies to stochastic growth and to stochastic discounting provided that the associated multiplicative functional is locally predictable. For instance, these methods provide a general way to characterize the link between short-term and long-term discounting as posed by Newell and Pizer (2003, 2004). In this application $-\beta$ is the instantaneous discount rate and $-\rho$ is the long-term discount rate used in policy evaluation. Motivated by problems in climate policy analysis, they study discounting abstracting from stochastic growth and adjustment for risk. Their focus is on the long-run consequences for valuation of fluctuations in short-term discount rates. Specifically they argue that $-\rho$ can be substantially smaller than $-E\beta$. Optimization problem 4.3 gives a general statement characterization this inequality for Markov valuation problems. This optimization problem shows that in the long-run it is a distorted average of $-\beta(X_t)$ 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 $J(Q)$.[16] As we will see the resulting penchant for small long-term discount rates can be undermined by taking account of risk.

Recall from (11) that for more general multiplicative functional M that local growth rate β^* includes adjustments for local exposure to Brownian motion and jump risk. The multiplicative functional M can be decomposed into two component multiplicative functionals:

$$M_t = \exp\left(\int_0^t \beta^*(X_u)du\right) M_t^* \tag{20}$$

where M^* is a local martingale.[17] Both components are multiplicative functionals. When this local martingale is a martingale, it is associated with a distorted probability distribution for X.[18] The probability twisting associated with M^* preserves the Markov structure. The entropy measure discussed previously is now constructed relative to the probability

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.

[16] 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.

[17] 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.

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

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:

Problem 4.4.

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

where $\mathbb{J}^*$ *is constructed using the change in probability measure.*[19]

While optimization problem 4.4 guarantees that $\rho(M)$ exceeds an average of β^*, this average is computed using a change of measure. Thus 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 offset by the presence of local exposure of the discount or growth factors to risk.[20] For instance, the "exposure" of the stochastic discount factor to risk is what captures local risk premia, that is risk-premia for short-term investments, as I will characterize shortly. Risk adjustments also have long-term consequences as reflected in the formula:

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

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

Instead of directly solving problem 4.4, in the next section I will develop a method to compute ρ because this method will also provide a more refined characterization of the valuation dynamics.

5 Multiplicative factorization

So far, I have focused on the behavior of growth or decay rates. Since convergence will be slow in some example economies, and as a consequence it is important to develop a more refined approximation. I now propose a refinement based on a multiplicative factorization of stochastic growth or discount functionals with three components: a) an exponential

[19]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).

[20]This offset is important to produce an upward sloping term structure of interest rates.

function of time, b) a positive martingale, c) a ratio of function of the Markov process at zero and t:

$$\begin{array}{cccc} M_t \; = & \exp(\rho t) & \hat{M}_t & \left[\frac{e(X_0)}{e(X_t)}\right] \\ & \uparrow & \uparrow & \uparrow \\ & \textbf{growth or} & \textbf{martingale} & \textbf{state} \\ & \textbf{decay} & \textbf{change in probability} & \textbf{dependence} \end{array} \tag{21}$$

This decomposition constitutes provides the ingredients for a decomposition of value dynamics (a DVD). 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.[21]

The dynamics of risk pricing and premia are best understood in terms of this alternative measure as reflected in the formula:

$$\log \mathbb{M}_t f(x) - \rho t = \log e(x) + \log \hat{E}\left[f(X_t)\hat{e}(X_t)|X_0 = x\right] \tag{22}$$

where $\hat{E}$ is used to denote the expectation operator under the twisted measure and $\hat{e} = \frac{1}{e}$. Thus after adjusting for the growth (or decay) rate ρ, the implied values of a cash flow as a function of the investment horizon is represented conveniently in terms of the dynamics under the *twisted* probability measure.

This alternative probability measure gives me a framework for a formal study of long-term approximation on which I can use the existing toolkit for the study of Markov processes that are "stochastically stable."

Definition 5.1. *The process X is* **stochastically stable** *under the measure $\hat{}$ if*

$$\lim_{t\to\infty} \hat{E}\left[f(X_t)|X_0 = x\right] = \hat{E}\left[f(X_t)\right] \tag{23}$$

for any f for which $\hat{E}(f)$ is well defined and finite.[22]

Under this stochastic stability, the limit as the investment horizon t becomes large of

[21] As an alternative approach, Rogers (1997) proposes a convenient parameterization of a martingale from which one can construct examples of multiplicative functionals of the form (21). Given my interest in structural economic models of asset pricing and in products of stochastic discount factor and growth functionals, I will instead explore factorizations of pre-specified multiplicative functionals.

[22] 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.

(22) is

$$\lim_{t\to\infty} [\log \mathbb{M}_t f(x) - t\rho] = \log e(x) + \log \hat{E}\left[f(X_t)\hat{e}(X_t)\right]. \quad (24)$$

The second term on the right-hand side computes the unconditional expectation of $f\hat{e}$ under the twisted probability measure, provided of course that this expectation is finite and positive. This generalizes result (14) for the finite-state Markov chain example 4.1. Limit (24) gives a hyperbolic approximation for $\log \mathbb{M}_t f(x)$ as a function of the investment horizon:

$$\frac{1}{t}\log \mathbb{M}_t f(x) \approx \rho + \frac{1}{t}\Big(\log e(x) + \log \hat{E}\left[f(X_t)\hat{e}(X_t)\right]\Big).$$

By applying this approximation to $M = SG$, I have an operational method for characterizing how a DVD converges to its limit value.

Component c) is built directly from the principle eigenfunction e. It captures state dependence and it provides a way to characterize the convergence to the limiting growth (or decay) rate ρ. The choice of f contributes a constant term $\log \hat{E}\left[f(X_t)\hat{e}(X_t)\right]$, while the principle eigenfunction e determines the state dependence independent of f. At this juncture I call this component transient by analogy to the decomposition of an additive functional. Later I will relate this term to components of models that are transient from the standpoint of valuation.

Multiplicative factorization (21) is also a decomposition of a stochastic growth or discount process. All three components are themselves multiplicative functionals, but with very different behavior. The term $\exp(\rho t)$ is not stochastic. The multiplicative martingale has expectation unity for all t and thus is not *expected* to grow. When applied to stochastic discount factor processes, Alvarez and Jermann (2000) interpret the martingale as reflecting the role of permanent macroeconomic shocks in understanding the term structure of interest rates. The third component appears "transient" when the underlying Markov process X is stationary. While the stochastic inputs of the martingale $\hat{M}$ will be long lasting, perhaps the same is not true for this third component.[23]

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 behavior of the original M process. Thus I am led to ask: Q1: Is factorization (21) unique? Q2: When is this factorization useful? The answers to these two questions are intertwined. The remainder of the section answers these questions and develops more fully the construction

[23] Although positive, the martingale component of the factorization will typically not converge to a limiting random variable with unit expectation.

and implications of (21).

5.1 Factorization

I build the factorization as follows. First I solve:

$$E\left[M_t e(X_t) | X_0 = x\right] = \exp(\rho t) e(x) \tag{25}$$

for any t where e is strictly positive as in (12). 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:

$$\lim_{t \downarrow 0} \frac{E\left[M_t e(X_t) | X_0 = x\right] - \exp(-\rho t) e(x)}{t} = 0, \tag{26}$$

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

$$\mathbb{B} e = \rho e, \tag{27}$$

where

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

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 4.1.

Typically it suffices to solve the local equation (27) to obtain a solution to (25). See Hansen and Scheinkman (2008) for a more detailed discussion of this issue. In the finite-state Markov model of section 4.1, convenient and well known sufficient conditions exist for there to be a unique (up to scale) positive eigenfunction satisfying (25). More generally, however, this uniqueness will not hold. Instead I will obtain uniqueness from additional considerations.

Given a solution to (25), 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 (21) follows immediately by letting $\hat{e} = \frac{1}{e}$ and solving for M in terms of $\hat{M}$, ρ and $\hat{e}$.

5.2 Additive versus multiplicative decomposition

There are important differences in the study of additive and multiplicative functionals and decompositions. For instance, principal eigenfunction e in the multiplicative factorization (21) is not the exponential of the the function g used in the additive decomposition (5):

$$\exp[g(x)] \neq e(x).$$

In special cases, however, the two are related.

Example 5.2. *Consider again example 3.3 and recall the additive functional:*

$$dY_t = \nu dt + H X_t dt + F dW_t.$$

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 example 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 (27). 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. This is typical in log-normal models. In this case the relevant variance, $|F - HA^{-1}B|^2$, is that of the additive martingale scaled by the time interval of evolution.

This adjustment illustrates both the discount-rate reduction of Newell and Pizer (2003, 2004) in a special case and the ambiguity of this comparison in general. When M is a discount factor process constructed with $F = 0$, the long-term discount rate is $-\nu - \frac{|HA^{-1}B|^2}{2}$ and the short-term (instantaneous) rate is $-\nu - HX_t$. On average, the long-term rate is smaller, consistent with the outcome of problem 4.3. If, however, the stochastic discount factor has $F \neq 0$, the long-term rate is $-\nu - \frac{|F-HA^{-1}B|^2}{2}$ and the short-term rate is $-\nu - HX_t - \frac{|F|^2}{2}$. The ordering of short-term and long-term discount rates now depends on the relative magnitude of the instantaneous volatility, $|F|$, and the long-term volatility, $|F - HA^{-1}B|$, of the logarithm of the stochastic discount factor process.

Consider now 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 $\log M^{[2]}$. When $M^{[1]}$ is a stochastic discount factor and $M^{[2]}$ is a stochastic growth factor, from formula (19) it follows that the negative of this covariance is the long-term risk premium.

In log-normal example 5.2 there is a simple link between the additive decomposition and the multiplicative factorization. My next example shows that this link breaks down when volatility is state dependent.

Example 5.3. *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 (27):*

$$\nu + x_1{}' \left(A_{11}{}' \alpha_1 + H_1{}' \right) + (x_2 - 1) \left(A_{12}{}' \alpha_1 + A_{22} \alpha_2 + H_2 \right) + \frac{1}{2} x_2 |\alpha' B + F|^2 = \rho.$$

which I derive as a special case of (27). Thus the coefficients on x_1 *and* x_2 *are zero when:*

$$\begin{aligned} A_{11}{}' \alpha_1 + H_1{}' &= 0 \\ A_{12}{}' \alpha_1 + A_{22} \alpha_2 + H_2 + \frac{1}{2} |\alpha_1{}' B_1 + \alpha_2 B_2 + F|^2 &= 0. \end{aligned} \tag{28}$$

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:*

$$\alpha_1 = -(A_{11}{}')^{-1} H_1{}'$$

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

$$\begin{aligned} \alpha_2 = & - \left(\frac{B_2 \cdot F + A_{22}}{|B_2|^2} \right) \\ & \pm \frac{\sqrt{|B_2 \cdot F + A_{22}|^2 - |B_2|^2 \left(|F - H_1 (A_{11})^{-1} B_1|^2 + 2H_2 - 2H_1 (A_{11})^{-1} A_{12} \right)}}{|B_2|^2}, \end{aligned} \tag{29}$$

provided that the term under the square root sign is positive. Notice in particular that this term will be positive for sufficiently small $|B_2|$. *Finally,*

$$\rho = \nu - (A_{12}{}' \alpha_1 + A_{22} \alpha_2 + H_2).$$

In contrast to example 5.2, e is not the exponential of the function g used in the additive martingale construction in example 5.3. Moreover, in example 5.3 there are two possible solutions for e that are exponentials of linear functions of the state vector. I will have cause to select one of these solutions as the interesting one to use in approximation. Finally, twice the difference between ν and ρ is no longer interpretable as a long-run variance.

5.3 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.

I present 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 change the probability measure, and approximation can proceed using tools from the study of Markov processes with stable stochastic dynamics. The stability condition is presumed to hold under the distorted or twisted probability distribution.

Theorem 5.4. *Given a multiplicative functional M, suppose that e and ρ satisfy equation (26) 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 $\hat{e} = 1/e$.

It follows from Theorem 5.4 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 $\hat{e}$. The resulting functions f 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 5.5. *For a given multiplicative functional M, a function $f(X)$ 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.

As I noted previously, 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 (21), Hansen and Scheinkman (2008) show that there is at most *one* such representation for which the process X is stochastically stable.

Recall that in example 5.3 we found two solutions for α_2 by solving the quadratic equation (28). As an implication of the Girsanov Theorem, associated with each solution is an alternative probability measure under which

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

where $\hat{W}_t$ is a multivariate standard Brownian motion under the twisted measure. The implied *twisted* evolution equation for $X^{[1]}$ is:

$$dX_t^{[1]} = A_{11}X_t^{[1]}dt + A_{12}(X_t^{[2]} - 1)dt + \left(B_1F' + |B_1|^2\alpha_1\right) X_t^{[2]}dt + d\hat{W}_t,$$

and for $X^{[2]}$ is

$$
\begin{aligned}
dX_t^{[2]} &= A_{22}(X_t^{[2]}-1)dt + \left(B_2F' + |B_2|^2\alpha_2\right)X_t^{[2]}dt + \sqrt{X_t^{[2]}}d\hat{W}_t \\
&= \pm\sqrt{|B_2F'+A_{22}|^2 - |B_2|^2\left(|F-H_1(A_{11})^{-1}B_1|^2 + 2H_2 - 2H_1(A_{11})^{-1}A_{12}\right)}X_t^{[2]}dt \\
&\quad -A_{22}dt + \sqrt{X_t^{[2]}}d\hat{W}_t,
\end{aligned}
$$

where in the second representation I have substituted from solution (29). I select the "minus" solution to achieve stochastic stability.

5.4 Long-run unusual 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 makes a multiplicative martingale valuable analytically valuable as a means of long-term approximation. From another perspective, however, 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 $E\left(\hat{M}_\infty | \mathcal{F}_0\right) \le 1$ 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 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).[24] Specifically I bound a threshold probability by taking expectations

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

of a dominating function:

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

for any $0 \leq \mathsf{a} \leq 1$ and any real number b. 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 long-term approximation, the process itself can become small. For this reason, I find (21) most interesting as a DVD, than as a directly interpretable factorization of a multiplicative stochastic process.[25] In particular, the martingale gives an alternative probability measure that is convenient to represent risk premia and prices over alternative payoff horizons.

5.5 Transient model components

I now suggest 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, $M = S$, or the product of a stochastic discount functional and a reference growth functional $M = SG$. I ask: what modifications have transient implications for valuation? The tools I described in this section give an answer.

Given an M implied by a benchmark valuation model, recall our multiplicative factorization (21):

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

[25] See Alvarez and Jermann (2000) for its use as a direct decomposition of a stochastic discount factor process.

Moreover, suppose that under the associated $\hat{}$ probability measure X satisfies a stochastic stability condition 5.1. Consider an alternative model of the form:

$$M_t^* = M_t \frac{\hat{f}(X_t)}{\hat{f}(X_0)} \tag{30}$$

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 will illustrate such representations in section 7. In light of (30), a candidate factorization for M^* is:

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

The exponential and martingale components remain the same as for the factorization of M. This gives me an operational notion of a *transient* change in valuation.

Definition 5.6. *The difference in the semigroups associated with M^* and M is* **transient** *if their multiplicative factorizations share the same exponential growth (or decay) component:* $\exp(\rho t)$ *and the same martingale component: $\hat{M}$ for which the process X is stochastically stable under the implied change in probability measure.*

When the difference between the semigroups associated with M and M^* is transient, long-term components to the factorizations are the same. The principal eigenfunctions, however, can be different. For instance, the principal eigenfunction for M^* given by (30) is: $e^* = e/\hat{f}$.[26] The difference between e and e^* alters the family of transient functions because for f to be transient for M^*:

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

In particular, this restriction depends on $\hat{f}$. Thus the collection of functions for which the long-term approximation methods are applicable is altered. Moreover, even if f is transient

[26] My reference to e^* as an eigenfunction is a bit loose because I have not prespecified a space of functions that it resides in. Instead I use the formalization of Hansen and Scheinkman (2008) that defines an eigenfunction using a martingale approach.

for both M and M^*, the (state dependent) coefficient for the hyperbolic approximation differs whenever:

$$\hat{E}\left[f(X_t)\hat{e}(X_t)\right] \neq \log \hat{E}\left[f(X_t)\hat{e}(X_t)\hat{f}\right] - \log \hat{f}(x).$$

What have I established in this section? I have a constructed a factorization of a multiplicative functional first to characterize how a DVD converges to its long-term limit, and second to reveal when alternative economic models have the same long-term limiting valuations.

6 Perturbation

Derivatives computed at a baseline configuration of parameters reveal sensitivity of a valuation model to small changes in a parameters. For instance, in Hansen et al. (2008a) the pricing implications of a parameterized family of valuation models depend on the intertemporal elasticity of the investors. They compute derivatives as an alternative to solving the model for the alternative parameter configurations. A risk price is also a derivative. It is a marginal change in a risk premium induced by a marginal change in risk exposure. Thus a key to constructing a risk price is to parameterize the risk exposure of a hypothetical cash flow.

In both of these applications, the multiplicative functionals used in constructing the semigroup depend on a model parameter. Thus I consider $M(\epsilon)$ as a parameterized family of multiplicative functionals and analyze value implications in the vicinity of $\epsilon = 0$. The parameter can be a preference parameter as in the work of Hansen et al. (2008a), or it can be a parameter that governs the exposure to a source of long-term risk that is to be valued. My specific aim is to study how risk-premia associated with alternative investment horizons depend on the parameter ϵ. In the Hansen et al. (2008a) application, the stochastic discount factor process is depicted as $S(\epsilon)$ and $M(\epsilon) = S(\epsilon)G$ where G is the stochastic growth component of a hypothetical or real cash flow. In the applications I consider in section 7, ϵ parameterizes the long-run risk exposure of a hypothetical cash flow. In this case $M(\epsilon) = SG(\epsilon)$ and the derivative of risk premium gives a risk price for the respective horizon.

With a perturbation analysis, it is possible to exploit a given solution to a model in the study of sensitivity to model specification. Changing the parameter ϵ of $M(\epsilon)$ allows me to perturb the valuations associated with this process. My choice of a *scalar* parameterization is made for notational convenience. The multivariate extension is straightforward.

In my applications in section 7, $G(\epsilon)$ is a parameterized family of multiplicative martingales. I use martingales in order to feature the role of pricing dynamics, since the trajectory of expected cash flows will be identically equal to one when $G(\epsilon)$ is a multiplicative martingale. I feature diffusion models, so $G(\epsilon)$ can be expressed as

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

The respective risk price is:

$$\textbf{risk price} = -\frac{d}{d\epsilon}\frac{1}{t}\log E[S_t G_t(\epsilon)|X_0 = x]\Big|_{\epsilon=0}. \tag{31}$$

By exploring alternative parameterizations of risk exposures, I can infer which directions are of most concern to investors as reflected by pricing implications of an underlying economic model.

The mathematical finance literature includes calculations of the local sensitivity of prices of derivative securities to underlying parameters. The finite horizon risk prices that interest me have the same structure as some of the calculations in this literature.[27] The specific calculation that interest me can be obtained by adapting a formula in Fourni et al. (1999). Imitating the calculation in their Proposition 3.1, construct

$$\Delta_t = \int_0^t D\xi(X_u;0)\cdot[dW_u - \xi(X_u;0)'du] \tag{32}$$

where $D\xi(X_u;\epsilon)$ is the partial derivative of ξ with respect to ϵ. Notice that Δ is an additive functional. I use this functional to represent the derivative of interest:[28]

$$\frac{d}{d\epsilon}\log E\left[S_t G_t(\epsilon) f(X_t)|X_0 = x\right]\Big|_{\epsilon=0} = \frac{E\left[S_t G_t(0) f(X_t)\Delta_t|X_0 = x\right]}{E\left[S_t G_t(0) f(X_t)|X_0 = x\right]}. \tag{33}$$

The results in Fourni et al. (1999) have been extended to include some specifications of jumps in Davis and Johansson (2006) with the corresponding modifications of the additive

[27]Such derivatives are often referred to as the "Greeks" in the option pricing literature.

[28]Fourni et al. (1999) derive this formula after imposing some regularity conditions that I do not repeat here. Their regularity conditions are sufficient conditions and turn out not to satisfied for some of my examples. For these I examples, however, I can perform direct calculations of the sensitivities.

functional Δ.

To study limiting properties, I first solve the principal eigenvalue problem for $\epsilon = 0$ and use the solution to construct a probability measure $\hat{\cdot}$ as we described previously. Recall that in the stochastic evolution under the twisted probablity measure, dW_t becomes a multivariate standard Brownian motion with an explicit drift distortion that depends on the Markov state. With this change of measure,

$$\frac{d}{d\epsilon} \log E\left[S_t G_t(\epsilon) f(X_t) | X_0 = x\right]\Big|_{\epsilon=0} = \frac{\hat{E}\left(\frac{f(X_t)}{e(X_t)}\Delta_t | X_0 = x\right)}{\hat{E}\left(\frac{f(X_t)}{e(X_t)} | X_0 = x\right)}$$

where e is the principal eigenfunction for the semigroup constructed from $M = SG(0)$.

In the remainder of this section I will show that the limiting derivative is:

$$\frac{d}{d\epsilon}\rho(\epsilon)|_{\epsilon=0} = \frac{1}{t}\hat{E}\Delta_t \tag{34}$$

which can be evaluated for any choice of t including choices that are arbitrarily small. Notice that Δ is an additive functional. Thus, I obtain the derivative of ρ by computing the average of the average trend growth of additive functional Δ under the twisted $\hat{\cdot}$ probability measure. While many readers may just prefer to accept this formula including the limiting version as the investment horizon becomes small, for completeness I give a (heuristic) derivation.

6.1 Long-term limiting behavior

Let $M(\epsilon) = SG(\epsilon)$, and recall the decomposition:

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

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

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

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

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

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 $\frac{d}{d\epsilon}\log \hat{M}_t(\epsilon)$ evaluated at $\epsilon = 0$ is the score of the likelihood over an interval of time t.

Now use the decomposition and differentiate $\log M_t(\epsilon)$

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

Take expectations and use the fact that X is stationary under the $\hat{\cdot}$ probability measure to obtain

$$\left.\frac{d\rho(\epsilon)}{d\epsilon}\right|_{\epsilon=0} = \frac{1}{t}\hat{E}\left[\left.\frac{d}{d\epsilon}\log M_t(\epsilon)\right|_{\epsilon=0}\right]. \quad (35)$$

This argument was not specific to diffusions there it applies to models with jumps as well. In the special case of a diffusion it agrees with (34).

6.2 Using the local evolution

I now make formula (35) operational by studying the limiting version as t declines to zero. 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

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

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

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

and form $M(a) = \exp[Y(a)]$. It is most convenient to take limits of (34) as $t \to 0$. This entails computing an average local mean under the distorted distribution:

$$\begin{aligned} \frac{d}{d\epsilon}\rho(\epsilon)|_{\epsilon=0} &= \hat{E}\left(\frac{d}{d\epsilon}\left[\beta(X_t;\epsilon) + \xi(X_t;\epsilon)\cdot\hat{\xi}(X_t)\right]\Bigg|_{\epsilon=0}\right) \\ &+ \hat{E}\left[\int \frac{d}{d\epsilon}\chi(y, X_t;\epsilon)\bigg|_{\epsilon=0} \exp[\hat{\chi}(y, X_t)]\eta(dy|X_t)\right] \end{aligned} \quad (36)$$

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 $\exp[\hat{\chi}(y, X_t)]\eta(dy|X_t)$ to construct the $\hat{\cdot}$ the jump intensity and the jump distribution conditioned on the current Markov state.

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 smoothness of parameterized families of operators in their formal development of large deviation results for Markov processes.

6.3 Convergence

Will the hyperbolic convergence extend to the risk prices? It carries over the in example economies that I consider, and there are good reasons to expect that a more general analysis is possible. From formulas (33) and (34),

$$\frac{d}{d\epsilon}\log E\left[S_tG_t(\epsilon)f(X_t)|X_0 = x\right]\bigg|_{\epsilon=0} - t\frac{d}{d\epsilon}\rho(\epsilon)|_{\epsilon=0} = \frac{\hat{E}\left[\frac{f(X_t)}{e(X_t)}\left(\Delta_t - \hat{E}\Delta_t\right)|X_0 = x\right]}{\hat{E}\left[\frac{f(X_t)}{e(X_t)}|X_0 = x\right]}.$$

For simplicity, suppose that $f = e$, and consider the limit:

$$\lim_{t\to\infty} \hat{E}\left(\Delta_t - \hat{E}\Delta_t|X_0 = x\right). \quad (37)$$

Recall that Δ is an additive functional, and Theorem 3.2 gives a martingale decomposition of such a functional. It is the decomposition under that change of measure that interests me. Subtracting the mean of Δ_t in expression (37) removes the linear trend component, and the conditional expectation of the martingale component is zero. Thus

$$\hat{E}\left(\Delta_t - \hat{E}\Delta_t | X_0 = x\right) = -\hat{E}\left[\hat{g}(X_t) | X_0 = x\right] + \hat{g}(x)$$

where $\hat{g}$ is the g of Theorem 3.2 for Δ under the change of measure. Thus

$$\lim_{t \to \infty} \hat{E}\left(\Delta_t - \hat{E}\Delta_t | X_0 = x\right) = -\hat{E}\hat{g}(X_t) + \hat{g}(x).$$

More generally, that is when $f \neq e$, I expect the martingale component of Δ to contribute to this limit and an additional computation is required.[29]

7 Applications

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. Intermediate time frames form a "term structure" of risk prices between these two limits. These dynamics are conveniently characterized using the properties of the twisted Markov transition described in section 5.

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 there 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.[30] Instead of extracting martingale components from initial multiplicative growth

[29]Potentially, the limit calculation could exploit Properties P2 and P6 of Malliavin calculus described in Fourni et al. (1999).

[30]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,

processes, I will build them directly and explore the resulting pricing implications. I use the construction of risk prices given in section 6 by computing derivatives of the form (31).

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 those 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. The calculations abstract from production, but they provide a dynamic characterization of a) the impact of risk aversion over alternative investment horizons, b) impact of risk prices on risk premia. These calculations will deliberately extrapolate value implications beyond the support of the data by looking at pricing implications for hypothetical cash flows at different horizons. In this sense, I will be using the models as "structural". These pricing calculations are of direct interest and they are informative for welfare cost calculations using the methods in Hansen et al. (1999) and Alvarez and Jermann (2004).

7.1 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 that emerges when investors have preferences represented using discounted power utility:

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

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.

---

and it does not alter the implied risk premia and corresponding risk prices.

7.2 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 reveal how risk prices change across horizons for alternative models.

Suppose that consumption is a geometric Brownian motion:

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

where C_t is aggregate consumption. I allow the Brownian motion $\{W_t : t \geq 0\}$ to be multivariate.[31]

Construct S in accordance with the power utility model:

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

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:

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

for the reasons I stated at the outset of this section.

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 a 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 for 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

[31] Here I allow for W to generate a larger filtration that the underlying Markov process X. What is critical for me is that the Markov dynamics are not altered with this more refined filtration. As emphasized to me by Eric Renault, this can be appropriately formulated as a statement that additional Brownian increments to be priced do not Granger-cause the underlying Markov process.

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 F_g used to represent stochastic growth.

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

$$\textbf{risk premium} = \gamma F_c \cdot F_g + r - r^*. \tag{38}$$

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.[32] This discontinuity is depicted in figure 2 in which we depict the risk premia when investors care about external habits and when they do not. Typically the risk premia converge to zero as F_g converges to zero, but in fact the limiting risk premium is the differential in the risk-free rates between the the Campbell and Cochrane (1999) model and the Breeden (1979) model. In figure 2 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 $F_g = 0$ for valuation over long time horizons. Campbell and Cochrane (1999) show that the conditional second moment of the stochastic factor diverges as the time horizon is extended. Our analysis gives a more refined characterization of the limiting behavior. The limiting risk premia remain finite and the limiting risk prices coincide with the Breeden (1979) model except at $\sigma_g = 0$. In summary, the Campbell and Cochrane (1999) model has risk-premia that remain larger in the limit as the investment horizon is increased than those in the Breeden (1979) model.

[32] 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.

The nonlinearity in this model induces an infinite risk price when evaluated at the point at a which there is no exposure to growth-rate risk.

Next I use a specification of consumption externalities proposed by Santos and Veronesi (2006). 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 is different. The limit prices are the same as in a corresponding Breeden economy and as consequence are small. Given the dramatically larger local risk prices, I study the dependence of these prices on the investment horizon. Figure 3 gives the risk price trajectory for the Santos-Veronesi model and shows the pull towards the Breeden model. The change of measure induced by the martingale dictates the transitional dynamics of the risk prices. In this model there is a state variable that measures private consumption relative to a social habit stock. The trajectories depend on this state variable. As is evident in this figure, the sensitivity of the risk prices to the Markov state vanishes much more quickly than the convergence to the limit prices. This initial convergence is dictated by the Markov process under the twisted evolution. As I argued previously the convergence of the risk premia trajectories is eventually hyperbolic in the investment horizon (see formula (24)), and the same is typically true of the risk-price trajectory. Figure 3 includes a hyperbolic function as a reference curve, and the trajectories starting from different Markov states converge to this hyperbolic function.[33] Later in this section, I show how the coefficient for this hyperbolic function depends on the stationary distribution under the twisted Markov law.

Figure 3 also illustrates the importance of state variability in the prices for shorter investment horizons. The figure considers initializations at the .25 and .75 quantiles of the stationary distribution for the Markov state. The sizeable differences in local prices, eventually vanish but the when the state is set at the .25 quantile, there is a hump shape to the risk-price trajectory.

7.3 Models with consumption predictability

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

$$d\log C_t = \mu_c dt + H_c X_t^{[1]} dt + \sqrt{X_t^{[2]}} F_c dW_t \tag{39}$$

[33]A special feature of this asset pricing model is that the approximating hyperbolic function does not depend on the Markov state.

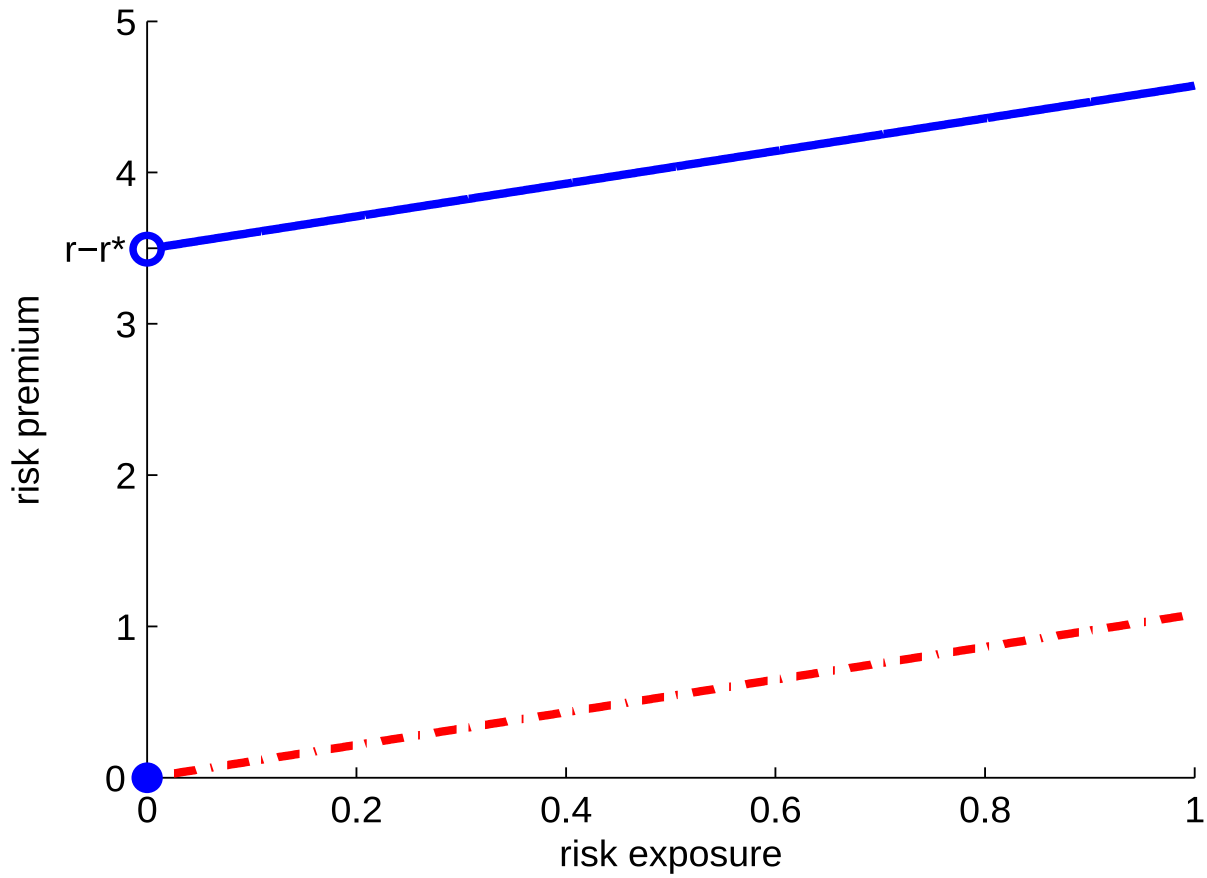

Figure 2: Risk premia as function of risk exposure. The vertical axis is scaled by one hundred so the risk premia are in percent. The dot-dashed line denotes the implied premia when investors have "external habits", and the solid line denotes the implied premia when investors have expected utility preferences. The parameter values for the state evolution are: $F_c = 0.0054$ and $\mu_c = .0056$. I set $\gamma = 2$, and for the model with investors that have external habits I set $\theta = 0$ and $\xi = .035$. The parameter ξ is continuous time counterpart to the corresponding parameter in Table 1 of Campbell and Cochrane (1999).

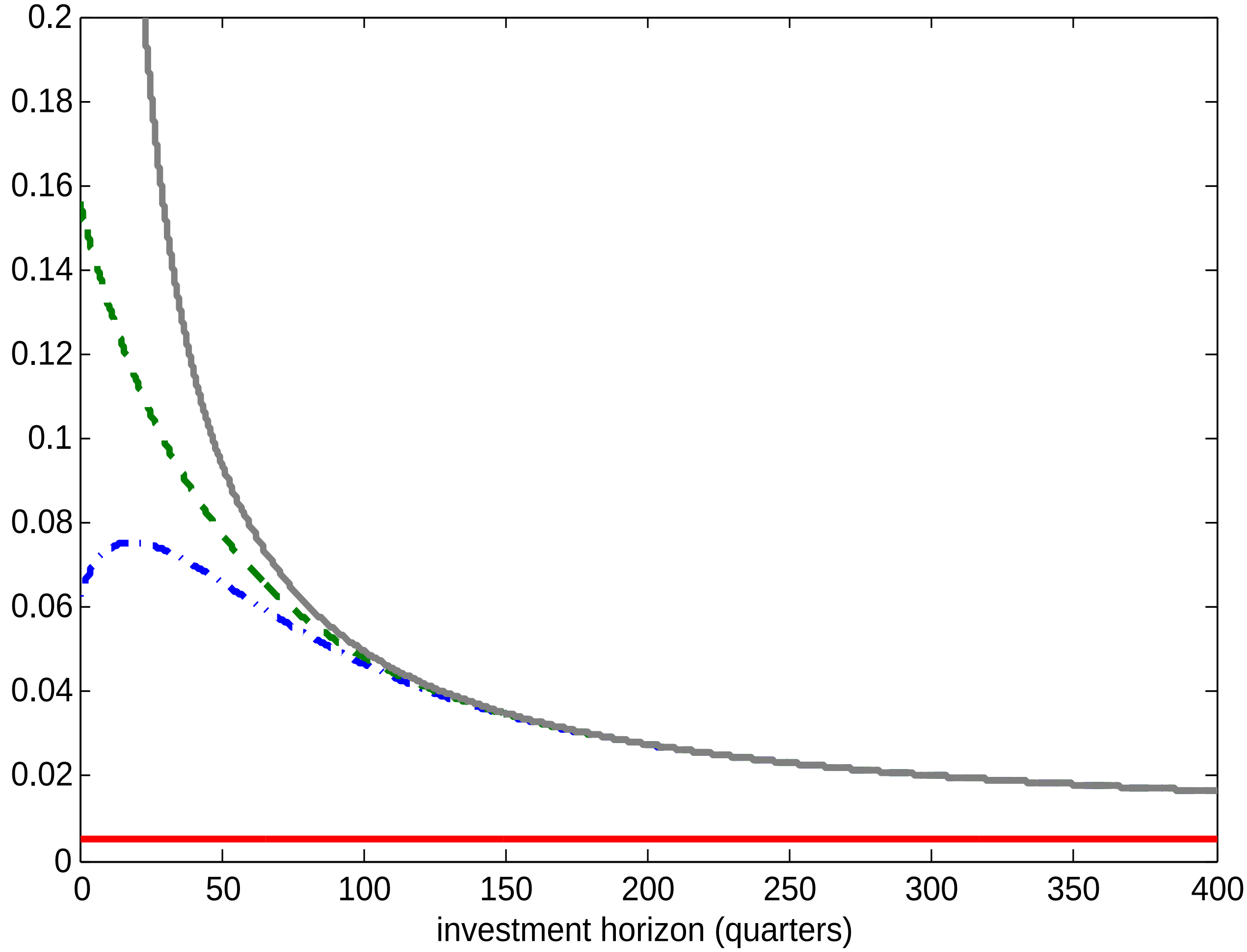

Figure 3: Risk price as a function of the investment horizon. The horizontal axis is given in quarterly time units. The bottom solid line denotes the implied prices when investors have expected utility preferences and the upper line depicts the hyperbolic function that is constructed using the invariant distributions under the twisted measures used to approximate the limiting behavior of the long-term prices when investors have external habits. The dashed line gives the risk-price trajectory obtained by setting x at the .75 quantile of its stationary distribution, the dash-dotted gives this same trajectory when x is set at the .25 quantile of its stationary distribution. The parameter values for the state evolution are: $F_c = 0.0054$ and $\mu_c = .0056$. I set $\gamma = 1$, and for the model with investors that have external habits $\xi = .04$, $\kappa = 80$ and $\mu_x = .7$. These parameters are taken from Table 1 of Menzly et al. (2004) adjusted to a quarterly time scale.

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

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

Consider a growth functional constructed as a martingale:

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

I compare the risk price implications for a 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.[34] The Kreps-Porteus investors care about the intertemporal composition of risk in contrast to investors in the Breeden model. In order that the limiting risk prices are the same, I set the risk aversion parameter for the Kreps-Porteus specification to coincide with that of the power used in the Breeden model. I justify this claim in subsection 7.5 where I show formally that the stochastic discount factor S^* for the recursive utility model

$$S_t^* = \exp(-\rho^{[r]} t) S_t \left(\frac{e^{[r]}(X_t)}{e^{[r]}(X_0)}\right) \tag{40}$$

for some scalar $\rho^{[r]}$ and some function $e^{[r]}$. This is a limiting result as the subjective rate of discount in investor preferences tends to zero. As a consequence of (40), $M^* = S^* G$ and $M = SG$ will share the same martingale component but not the same decay or growth rate in a factorization of the form (21. The shared martingale shows that the long-term risk return tradeoff is the same for the two models. On the other hand, the presence of $\rho^{[r]}$ in formula (40) implies that the long-term interest rates is different for the two models. I expect this because the elasticity of intertemporal substitution is different for the investors in the two models.

For both models it is straightforward to compute the "term structure" of risk prices. I depict the risk-price trajectories in figure 4 for a three-shock (three independent Brownian motions) version of the consumption dynamics given in equation (39). These trajectories

[34]See Schroder and Skiadas (1999) for a continuous-time formulation of the consumption-portfolio problem for an investor with such preferences.

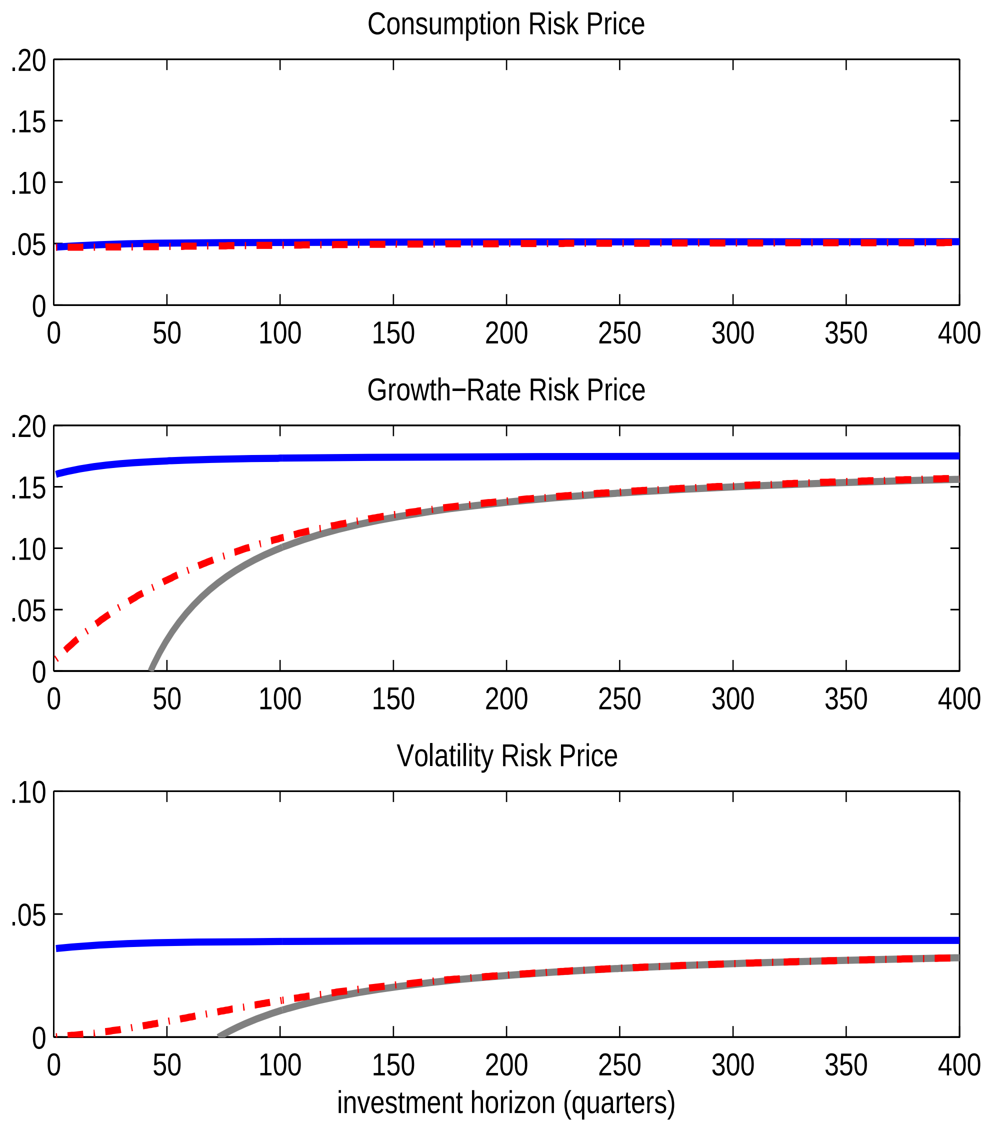

Figure 4: Risk prices indexed by investment horizon. The horizontal axis is given in quarterly time units. The upper solid line denotes recursive utility model, and dashed line the expected utility model. The lower solid line gives the the hyperbolic approximation for the expected utility model. The parameter values for the state evolution are: $A_{11} = -.025$, $A_{12} = 0$, $A_{22} = -.075$, $B_1 = [0 \;\; .00038 \;\; 0]$, $B_2 = [0 \;\; 0 \;\; -.19]$, $H_c = [1 \;\; 0]$, $F_c = [0.0047 \;\; .00076 \;\; 0]$. The risk prices are by localizing around $G = C$. For illustrative purposes I set $\gamma = 10$. See Hansen et al. (2008b) for estimation of the parameter values for this model. How to "calibrate" γ is an 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 $X_0^{[2]} = 1$.

give the risk prices as a function of the payoff horizon. Each panel corresponds to a different shock. A positive realization of the first shock increases the realized consumption growth; (a positive realization of) the second shock increases the growth rate in consumption, and (a positive realization of) the third shock diminishes consumption volatility. Shocks two and three have persistent consequences because growth rates and volatility are modeled as first-order autoregressions (see Example 3.4). The impulse-response function for the logarithm of consumption depicted in the top panel of figure 1 is for the second shock.[35] The formulas for risk-price trajectories are given in appendix C.

In the Breeden (1979) model, the local (instantaneous) risk price is negligible for shock two and zero for shock three as displayed in panels two and three of figure 4. The risk prices increase with horizon as the impact of the shocks on consumption becomes more magnified over longer horizons. There are two forces behind the convergence. Growth rates and volatility are highly persistent even under the change of measure. Moreover, as I have already argued, the eventual convergence to the limiting prices is hyperbolic in the investment horizon. This is reflected in the second and third panels, where I plot the hyperbolic function derived for the limiting approximation.

Investor preferences are forward looking in the recursive utility model, and this is evident in the nonzero local risk prices for shocks two and three as depicted in the panels two and three of figure 4. The forward-looking component to these prices is reflected in the continuation values for the consumption plans. The resulting enhancement of the local price of the growth rate shock illustrates the pricing mechanism featured by Bansal and Yaron (2004). 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. (2008a).[36] The risk-price trajectory is literally flat for the first shock and the two models imply the same risk prices. (See the first panel of 4.) The coincidence of the pricing trajectories for the two models of investor preferences illustrates a point made by Kocherlakota (1990). 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. Note

[35]The top panel plots the function $\frac{H_c}{A_1}\left[\exp(tA_{11}) - 1\right] B_1$ for the parameter values given in figure 4. In the *moving-average* representation there is a scaling by the square root of the conditional variance process $X^{[2]}$ evaluated at the appropriate calendar date. The mean-square norm of $\sqrt{X_t^{[2]}}$ is one for its stationary distribution, and thus I have not distorted the magnitude of the impulse responses. See Gallant et al. (1993) for a discussion of impulse-response functions in nonlinear environments.

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

the range in the third panel is one half that of the other two panels.

In the remainder of this section, I provide the details of the calculations of risk premia and prices. Uninterested readers can skip this material.

7.4 Risk prices in the absence of consumption predictability

I consider formally the implications of three different models of investor preferences.

7.4.1 Model of Breeden

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

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

where

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

It follows that

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

By setting $F_g = 0$,

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

Thus

$$\rho(G) + \rho(S) - \rho(GS) = \gamma F_c \cdot F_g$$

The long-term risk prices can be computed by differentiating the right-hand side with respect to the risk exposure vector F_g, and are thus equal to: γF_c.

The dynamics of pricing for this example is degenerate, and in particular the local risk price vector is also equal to γF_c. Specifically, the local expected rate of return is given by

$$-\lim_{t \downarrow 0} \frac{1}{t} \log E\left[G_t S_t | X_0 = x\right] = \delta + \gamma\mu_c + -\frac{\gamma^2}{2}|F_c|^2 + \gamma F_c \cdot F_g. \tag{41}$$

By setting $F_g = 0$ notice that the instantaneous risk-free interest rate is constant and identical to the long-term counterpart: $-\nu(S) = \delta + \gamma\mu_c - \frac{\gamma^2}{2}|F_c|^2$. The vector of risk prices obtained by differentiating the local risk-premium with respect to the risk-exposure vector

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

7.4.2 Campbell and Cochrane model

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 the 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

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

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 B). Using this notation, write:

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

In light of the discussions in section 5.5, I expect the differences between the valuation implications for Campbell-Cochrane model and the Breeden model be *transient*. As I will show, however, a substantial qualification is required to make this conclusion.

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

$$dX_t = -\xi(X_t - \mu_x)dt + \lambda(X_t)F_c dW_t \tag{42}$$

where I restrict $\lambda(0) = 0$. Squashing the variability at zero prevents the process X from being attracted to zero. After the imposing the change of probability measure obtained from the Breeden model, the law of motion for this equation is:

$$dX_t = -\xi(X_t - \mu_x)dt + (F_g - \gamma F_c) \cdot F_c \lambda(X_t)dt + \lambda(X_t)F_c d\hat{W}_t. \tag{43}$$

We use this evolution to compute the counterpart to (41):

$$\begin{aligned}
\lim_{t\downarrow 0} \frac{1}{t} \log E\left[G_t S_t^* | X_0 = x\right] &= \lim_{t\downarrow 0} \frac{1}{t} \log E\left[G_t S_t \exp[\gamma(X_t - X_0)] | X_0 = x\right] \\
&= -r - \gamma F_c \cdot F_g - \lim_{t\downarrow 0} \frac{1}{t} \hat{E}\left[\exp[\gamma(X_t - X_0)] | X_0 = x\right] \\
&= -r - \gamma F_c \cdot F_g + \gamma\xi(x - \mu_x) + \gamma(F_g - \gamma F_c) \cdot F_c \lambda(x) \\
&\quad + \gamma^2 \frac{\lambda(x)^2 |F_c|^2}{2}
\end{aligned} \tag{44}$$

where r is the risk-free rate from the Breeden economy ($r = \delta + \gamma\mu_c + \frac{\gamma^2}{2}|F_c|^2$) 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

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

(See appendix B.) 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:

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

which follows because (44) is affine in F_g and the risk prices are the negative of the partial derivative with respect to F_g. By design the local risk prices are state dependent and are larger than in the power utility model for a 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. In fact it is equal to $\gamma[1 - \lambda(\mu_x)]$ for $\mu_x = x$. (See also appendix B.)

Consider next the long-run behavior of value. I use evolution equation (43), and the formula for the logarithmic derivative of the density for a scalar diffusion:

$$\frac{d \log q}{dx} = \frac{2 \;\text{ drift}}{\text{diffusion}} - \frac{d \log \text{diffusion}}{dx} \tag{45}$$

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

The limiting behavior is dominated by the constant term:

$$\lim_{x \to \infty} \frac{d \log q}{dx} = -\frac{\gamma^2 \xi}{\gamma \xi - \theta} < 0. \tag{46}$$

As a consequence the process X is stationary under the twisted probability measure and under the original probability measure as reflected by (42) and (43) 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 (46) 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,

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

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 $\log q(x)+\gamma x$. While this term is zero when F_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 $\theta>0$,

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

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

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

but $\rho(S^*)$ differs and is given by the implied real interest rate r^*. This justifies formula (38) and figure 2.[37]

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 interest rate 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

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

[37]Presumably, the risk prices evaluated at $F_g=0$ are finite at any finite horizon, but they become arbitrarily large as the valuation horizon is extended inducing a discontinuity in the limit.

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 B.

What do I make of this? I 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 shown by Hansen and Scheinkman (2008), this uniqueness is to be expected.

When $\theta = 0$ the multiplicative martingale $\tilde{M}$ is the pertinent one for pricing discount bond whereas the martingale $\hat{M}$ for pricing growth rate risk over long horizons. The discontinuity in the long-term risk premia as a function of F_g as expressed in (38) reflects the separate roles of the two martingales in pricing. When $\theta > 0$, only the multiplicative martingale $\hat{M}$ is pertinent to pricing.

7.4.3 Santos and Veronesi model

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:

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

where

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

and $\mu_x > 0$.[38] 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

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

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 (43) but with this new specification of λ. The twisted law of motion for X is

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

where

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

This process remains in the class studied by Wong (1964). To explore its long-term implications, formula (45) is again informative. The logarithmic derivative of the density is

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

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

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

The twisted density $\hat{q}$ has a finite first moment provided that $\hat{\xi}$ is positive. The mean is

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

given by $\hat{\mu}_x$. Thus provided that the mean reversion parameter ξ is sufficiently large

$$\xi \geq \gamma\kappa|F_c|^2 - \kappa F_c \cdot F_g. \tag{47}$$

When inequality (47) is satisfied for an open set of values of F_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:

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

Moreover,

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

Thus the hyperbolic approximation to the risk premia trajectory is:

$$\textbf{risk premium} \approx \gamma F_c \cdot F_g - \frac{1}{t}\left[\log(1+\hat{\xi}) + \log(1+x)\right]$$

The risk prices for a finite horizon are obtained by differentiating the risk premia with respect to F_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 $F_g = 0$.

7.5 Risk prices in the presence of consumption predictability

In what follows I characterize formally the local prices and their long time horizon limits for the Breeden model and the Kreps-Porteus model. Thus I justify what happens at both ends of the "term structure" of risk prices, and I produce the corresponding twisted measure. I specify the preferences for investors in the two models so that the long-term risk prices are the same. To support this claim, I show that the martingale components used in the changes of measure are the same for both models, and I produce the corresponding twisted probability measures. By borrowing insights from the literature on robust control, 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.

7.5.1 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 (36) and (32), the long-term 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

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

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

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

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

$$dX_t^{[2]} = A_{22}(X_t^{[2]} - 1)dt + B_2\left[F + (B_2)'\alpha_2^{[b]}\right]X_t^{[2]}dt + \sqrt{X_t^{[2]}}d\hat{W}_t$$

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

$$dX_t^{[2]} = \hat{A}_{22}\left(X_t^{[2]} - \hat{\mu}_2\right)dt + \sqrt{X_t^{[2]}}d\hat{W}_t$$

where

$$\begin{aligned} \hat{A}_{22} &= A_{22} + B_2\left[F' + (B_2)'\alpha_2^{[b]}\right] \\ \hat{\mu}_2 &= \frac{A_{22}}{\hat{A}_{22}} = \hat{E}\left(X_t^{[2]}\right). \end{aligned}$$

The twisted probability distribution will alter the $X^{[1]}$ dynamics as well.

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

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

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

$$-\hat{E}\left(X_t^{[2]}\right)(B_1)'\alpha_1^{[b]} = -\gamma\hat{E}\left(X_t^{[2]}\right)(B_1)'[(A_1)']^{-1}(H_c)' \tag{48}$$

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

7.5.2 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 Runolfsson (1994). In the discounted version of recursive preferences, the stochastic discount factor is;

$$S_t^* = \exp(-\delta t) \left(\frac{C_0}{C_t} \right) \hat{V}_t \tag{49}$$

where $\hat{V}$ is a martingale component of $\left\{ \left(\frac{V_t}{V_0} \right)^{1-\gamma} : t \geq 0 \right\}$ and V is the stochastic process of continuation values.[39]

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 (49) 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.

For this recursive utility 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 as a function $\tilde{f}$ of the Markov state, and $\tilde{f}$ solves the equation:

$$\delta \log \tilde{f}(x) = \lim_{t \downarrow 0} \frac{E\left[\left(\frac{C_t}{C_0} \right)^{1-\gamma} \tilde{f}(X_t) | X_0 = x \right] - \tilde{f}(x)}{t}. \tag{50}$$

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. While the continuation value process becomes infinite, the ratio $\frac{V_t}{V_0}$ remains well defined in the limit for

[39]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 (20) and Ito and Watanabe (1965) and verifying that the local martingale is in fact a martingale.

all t as δ declines to zero. Call the resulting multiplicative process $\tilde{V}$. Similarly, construct $\tilde{C}$ in the same fashion by dividing C_t by C_0. The limiting stochastic discount factor can be represented as:

$$S_t^* = \frac{\hat{V}_t}{\tilde{C}_t}$$

where $\hat{V}$ is a multiplicative martingale.

I will show that it is the martingale component of $\tilde{C}^{1-\gamma}$ in this limiting case. Equation (50) 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 $\tilde{C}^{1-\gamma}$.[40] The dependence on t is necessary to allow the continuation value ratio $\tilde{V}$ to grow at a different rate than the consumption ratio $\tilde{C}$. 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 $\hat{V} = \tilde{V}^{1-\gamma}$ for the stochastic discount factor S^* in formula (49).

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,

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

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

The stochastic discount factors for the Breeden model and the Kreps-Porteus model have the *same* martingale components, although the decay rates are different. 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 (48) continues to apply to this model even though the local prices are different from the Breeden (1979) model.[41]

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. (2008a). 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 (48).

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

[41] Hansen et al. (2008a) make this observation for a discrete-time log-linear model abstracting from stochastic volatility. Thus distorted expectation in (48) plays no role in their analysis.

dynamics is used as a benchmark model. The "robust" adjustment is made to this benchmark probability specification by solving a minimization problem that penalizes changes in the probability law. For example, see Petersen et al. (2000) and Anderson et al. (2003).[42] 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 premia, this can be done by applying the perturbation method described in section 6. Hansen et al. (2008a) have used this method to explore changes in the intertemporal elasticity of substitution.[43] Perturbing risk prices requires the computation of additional cross derivatives since a risk price is itself a derivative.

7.6 Other models of asset prices

The examples that I have studied feature the role of investor preferences. A similar analysis applies to some equilibrium 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 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.

Applying these methods to the disaster-recovery models of Rietz (1988), Barro (2006) and Gourio (2008) will expand on the comparisons made across specifications of the consumption dynamics. Gourio (2008) shows that adding recoveries following disasters has an important impact on local (one-period) risk premia. Recoveries make the consequences of

[42]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.

[43]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).

disasters "transient" so that the specification changes explored by Gourio (2008) will have important consequences for the entire term-structure of risk prices.

8 Conclusion

Decompositions of additive functionals have proved valuable in macroeconomic time series as an aid in identifying shocks and quantifying their impact. The increments to the martingale components of these decompositions are the permanent shocks. In this paper I have considered an alternative decomposition. To support a dynamic value decomposition (DVD), I featured multiplicative factorizations of stochastic discount and growth functionals. These factorizations allowed me to

a) characterize a long-term risk-return relation;

b) construct risk prices for alternative investment horizons and characterize their long-term behavior;

c) compare implications for valuation of alternative structural economic models.

The methods I described require a "structural" model because they extrapolate value implications by featuring the the pricing of synthetically constructed martingale cash flows. Valuation of such cash flows reveals the dynamics of risk prices. While local prices are familiar to the literature on asset pricing, my aim was to explore the entire term structure of such prices. Thus I produced risk-price trajectories that can be viewed as the valuation counterpart to impulse response functions commonly used in economic dynamics. The analytical methods that I proposed here and in the antecedents, Hansen et al. (2008a) and Hansen and Scheinkman (2008), characterize what happens at longer horizons. A central part of this analysis is the extraction of martingale components used to transform the underlying probability measure. This change of measure provides a more refined characterization of risk-price trajectories and supports DVD's more generally. I suggest two ways to make model comparisons based on DVD's. One classifies when the implications of two alternative valuation models are transient. The other considers a parameterized family of valuation models and computes derivatives of long-term risk premia with respect to this parameter. These derivatives measure the sensitivity of value implications to a small change in the parameter.

While I have focused on risk-price dynamics, the methods I described can also be applied to study the dynamics of risk exposure in the presence of stochastic growth. Consumption or cash flows will typically have different risk exposures at alternative investment horizons. The methods suggested here give a model-based way to measure the resulting cash-flow duration as it contributes to measures of value. A single period return to equity with cash flows or dividend that have stochastic growth components can be viewed as a bundle or portfolios of holding period returns on cash flows with alternative payout dates. (See Lettau and Wachter (2007) and Hansen et al. (2008a).) Even though the pricing is "local", the risk exposure of the composite securities depends on how far into the future the primitive payoff will be realized. The methods described here can be adapted to study how cash flow dynamics are reflected in single-period returns of infinitely-lived securities or cash flows.

To conclude I want to be clear on two matters. First, while a concern about the role of 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. As I showed in examples from the asset pricing literature, the initial points in the risk price trajectories, the local risk prices, can be far from their limits. Thus the hyperbolic approximations I suggested provide an important refinement, and transient model components contribute to these approximations. Moreover, the outcome of the analysis is 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 absent, 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 a framework for the use of such measurements. But some of the measurement challenges remain, and I suspect the prior information about the underlying economic model will be required for sensible applications. Also, some of the same statistical challenges with which we econometricians struggle should be passed along to the hypothetical investors that populate our economic models. When decision-making agents within an economic model face difficulties in making probabilistic extrapolations of the future, the associated ambiguities in statistical 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 semigroup. Following Donsker and Varadhan (1975), Donsker and Varadhan (1976) and Berestycki et al. (1994), I study the following max-min problem:

$$\varrho = \sup_{Q\in\mathcal{Q}} \inf_{f\in\mathcal{D}^+} \int \left(\frac{\mathbb{B}f}{f}\right) dQ. \tag{51}$$

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[44]

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

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 (51). 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:

$$\mathbb{J}^*(Q) = \sup_{f\in\mathcal{D}_+} \left[-\int \left(\frac{\mathbb{A}f}{f}\right) dQ\right] \tag{52}$$

The function $\mathbb{J}^*$ is convex in Q since it can be expressed as the maximum of convex (in

[44]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 $\mathbb{B}1(x)$.

fact linear) functions of Q. Moreover, it can be justified as a relative measure of entropy between probabilities when the process implied by $\mathbb{A}$ possesses 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 this representation of the solution to the inner problem to write the outer maximization problem as:

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

which is the problem posed in (4.4).

Suppose that the solution to the max-min problem is attained with probability measure Q^*. Consider again the inner optimization problem (52) 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:

$$\int \left[\frac{\mathbb{A}(f^* f)}{f^*} - \frac{f(\mathbb{A}f^*)}{f^*} \right] dQ^* = 0. \tag{53}$$

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

$$\begin{aligned} \mathbb{A}^* f &= \frac{\mathbb{A}(f^* f)}{f^*} - \frac{f(\mathbb{A}f^*)}{f^*} \\ &= \frac{\mathbb{B}(f^* f)}{f^*} - \frac{f(\mathbb{B}f^*)}{f^*} \end{aligned} \tag{54}$$

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 (54) with $f^* = e$, and suppose this generates a stochastically stable Markov process with stationary distribution Q^*. In particular, it satisfies (53). 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 Reconsidering the Campbell-Cochrane Model

In this appendix I give some more details of my analysis of the Campbell-Cochrane model. Part of this discussion will be familiar to careful readers of Campbell and Cochrane (1999). I include some repetition because I parameterize their model in aa different (but equivalent) way.

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

$$r^* + \theta(x - \mu_x) = r + \gamma \xi (x - \mu_x) - \gamma^2 \frac{\lambda(x)^2 |\sigma_c|^2}{2} + \gamma^2 |\sigma_c|^2 \lambda(x). \tag{55}$$

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

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

Substituting this formula into (55), 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

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

or

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

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 the 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, I 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:

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

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.

C Risk premia for finite payoff horizons

In this appendix I give differential equations I solve to compute the risk-price trajectories for the model with consumption predictability. The analytical tractability is familiar from the literature on affine models (*e.g.* see Duffie and Kan (1994)).

Consider Example 3.4 and continued in Example 5.3. 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.

$$\mathbb{B}\mathbb{M}_t f = \frac{d}{dt}\left[\mathbb{M}_t f(x)\right] \tag{56}$$

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

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

Moreover,

$$\begin{aligned} \frac{\mathbb{B}\exp\left[\alpha(t) \cdot x + \varrho(t)\right]}{\exp\left[\alpha(t) \cdot x + \varrho(t)\right]} &= \nu + H_1 x_1 + H_2(x_2 - 1) \\ &\quad + x_1{}' A_{11}{}' \alpha_1(t) + (x_2 - 1)[A_{12}{}' \alpha_1(t) + A_{22} \alpha_2(t)] \end{aligned}$$

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

First use (56) to produce a differential equation for $\alpha_1(t)$:

$$\frac{d}{dt}\alpha_1(t) = {H_1}' + {A_{11}}'\alpha_1(t).$$

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,

$$\frac{d}{dt}\alpha_2(t) = H_2 + {A_{12}}'\alpha_1(t) + A_{22}\alpha_2(t) + \frac{1}{2}|F + \alpha_1(t)'B_1 + \alpha_2(t)B_2|^2$$

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

$$\frac{d}{dt}\varrho(t) = \nu - H_2 - {A_{12}}'\alpha_1(t) - A_{22}\alpha_2(t).$$

The initial condition is $\varrho(0) = 0$.

References

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

Alvarez, Fernando and Urban J. Jermann. 2004. Using Asset Prices to Measure the Welfare Cost of Business Cycles. *Journal of Political Economy* 112 (6):1223–1256.

Anderson, Evan W., Lars Peter 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, Srinivasan and Sean P. Meyn. 2000. Multiplicative Ergodicity and Large Deviations for an Irreducible Markov Chain. *Stochastic Processes and their Applications* 90:123–144.

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

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

Bansal, Ravi, Robert Dittmar, and Dana Kiku. 2008. Cointegration and Consumption Risks in Asset Returns. *Review of Financial Studies* forthcoming.

Barro, Robert J. 2006. Rare Disasters and Asset Markets in the Twentieth Century. *Quarterly Journal of Economics* 59 (4):823–866.

Berestycki, Henri, Louis Nirenberg, and S. R. Srinivasa 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, Stephen and Charles 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, Rakhi 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, Patrick. 1961. The Lindeberg-Levy Theorem for Martingales. *American Mathematical Monthly* 12.

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

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

Boyarchenko, Nina and Sergei Levendorskii. 2007. The Eigenfunction Expansion Method in Multi-Factor Quadratic Term Structure Models. *Mathematical Finance* 17:503–540.

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

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

Campbell, John Y. and Tuomo 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, John H. 2001. *Asset Pricing*. Princeton, N.J.: Princeton University Press.

Daniel, Kent and David Marshall. 2005. Equity-Premium and Risk-free-rate Puzzles at Long Horizons. *Macroeconomic Dynamics* 1 (2):452–484.

Davis, Mark H. A. and Martin P. Johansson. 2006. Malliavin Monte Carlo Greeks for Jump Diffusions. *Stochastic Processes and their Applications* 116 (1):101–129.

Donsker, Monroe E. and S. R. Srinivasa 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.

Duffie, Darrell and Rui Kan. 1994. Multi-Factor Term Structure Models. *Philosophical Transactions: Physical Sciences and Engineering* 347 (1684):577–586.

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

Dybvig, Philip H., Jr. Ingersoll, Jonathan E., and Stephen A. Ross. 1996. Long Forward and Zero-Coupon Rates Can Never Fall. *The Journal of Business* 69 (1):1–25.

Engle, Robert and Clive 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.

Fourni, Eric, Jean-Michel Lasry, Jrme Lebuchoux, Pierre-Louis Lions, and Nizar Touzi. 1999. Applications of Malliavin Calculus to Monte Carlo Methods in Finance. *Finance and Stochastics* 3 (4):391–413.

Gallant, A. Ronald, Peter E. Rossi, and George Tauchen. 1993. Nonlinear Dynamic Structures. *Econometrica* 61 (4):871–907.

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

Gourio, Francios. 2008. Disasters and Recoveries. *American Economic Review* 98 (2):68–73.

Hall, Peter and C. C. Heyde. 1980. *Martingale Limit Theory and its Application*. Probability and Mathematical Statistics. Boston: Academic Press.

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

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

Hansen, Lars Peter and Thomas J. Sargent. 1995. Discounted Linear Exponential Quadratic Gaussian Control. *IEEE Transactions on Automatic Control* 40.

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

———. 2008. Long-Term Risk: An Operator Approach. *Econometrica* forthcoming.

Hansen, Lars Peter, Thomas J. Sargent, and Jr. Tallarini, Thomas D. 1999. Robust Permanent Income and Pricing. *The Review of Economic Studies* 66 (4):873–907.

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.

Hansen, Lars Peter, John C. Heaton, and Nan Li. 2008a. Consumption Strikes Back?: Measuring Long-Run Risk. *Journal of Political Economy* 116:260–302.

Hansen, Lars Peter, Junghoon Lee, Nick Polson, and Seung Min Yae. 2008b. Assessing the Recursive Utility Model. University of Chicago.

Ito, Kiyosi and Shin Chi 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, Ioannis and Sean P. Meyn. 2003. Spectral Theory and Limit Theorems for Geometrically Ergodic Markov Processes. *Annals of Applied Probability* 13:304–362.

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

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

Lewis, Alan L. 1998. Applications of Eigenfunction Expansions in Continuous-Time Finance. *Mathematical Finance* 8 (4):349–383.

Linetsky, Vadim. 2004. Lookback Options and Diffusion Hitting Times: A Spectral Approach. *Finance and Stochastics* 8:373–398.

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

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

Menzly, Lior, Tano Santos, and Pietro Veronesi. 2004. Understanding Predictability. *Journal of Political Economy* 112 (1):1–47.

Meyn, Sean P. and Richard 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.

Otrok, Christopher, B. Ravi Kumar, and Charles H. Whiteman. 2002. Habit Formation: A Resolution of the Equity Premium Puzzle. *Journal of Monetary Economics* 49 (5):1261 – 1288.

Pearson, Karl. 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.

Rietz, Thomas A. 1988. The Equity Risk Premium: A Solution. *Journal of Monetary Economics* 22 (1):117–131.

Rogers, L. C. G. 1997. The Potential Approach to the Term Structure of Interest Rates and Foreign Exchange Rates. *Mathematical Finance* 7 (2):157–164.

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

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

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

Schroder, Mark and Costas Skiadas. 1999. Optimal Consumption and Portfolio Selection with Stochastic Differential Utility. *Journal of Economic Theory* 89:68–126.

Singleton, Kenneth 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, Peter. 1990. *Risk Sensitive and Optimal Control*. West Suffix, England: John Wiley and Sons.

Wong, Eugene. 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.