Modeling the Long Run: 
Valuation in Dynamic Stochastic Economies

Lars Peter Hansen*
April 15, 2008

Abstract
In this paper I propose to augment the toolkit for economic dynamics and asset valuation with methods that will reveal economic import of long-run stochastic structure. These tools enable informative decompositions of a model’s dynamic implications for valuation. The methods I feature the long-term behavior of valuation operators that explicitly incorporate stochastic growth. The valuation operators are indexed by the gap of time between when a payoff is realized and when it is priced. Appropriately adapted Perron-Frobenius theory gives a characterization of the valuation behavior when this gap becomes large. Using such methods I provide operational decompositions of value implications of economic models including measures of parameter sensitivity and characterizations of long-run risk prices.

1 Introduction
In this paper I propose to augment the toolkit for economic dynamics and econometrics with methods that will reveal economic import of long run stochastic structure. These tools enable informative decompositions of a model’s dynamic implications for valuation. They are the outgrowth of my observation and participation in an empirical literature that aims to understand the low frequency links between financial market indicators and macroeconomic aggregates.

*Prepared for the Fisher-Schultz Lecture at the 2006 European Meetings of the Econometric Society. I appreciate helpful comments by Jaroslav Borovicka, Valentin Haddad and Mark Hendricks. Portions of this work are joint with John Heaton, Nan Li and Jose Scheinkman, and very much influenced by related work I have done with Xiaohong Chen and Tom Sargent.
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 run value implications of economic models?
- 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. Implications at higher frequencies are either skimmed over; or additional model components, often ad hoc are added
in hopes of expanding the empirical relevance. 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-run approximation and to understand better what proposed model fixups do to long-run 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 depicted over a single period in a discrete time model or over an instant of time in a continuous time model. I will extend the log-linear analysis in Hansen et al. (2008) by deriving the long run counterpart to this familiar exercise by performing a sensitivity analysis that recovers prices of exposure to the component parts of long run (growth rate) risk. These same methods facilitate long-run welfare comparisons in explicitly dynamic and stochastic environments.

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

1.1 Game plan

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

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

ii) Decomposition of Additive functionals (section 3): I produce decomposition of additive functionals into a linear trend, a martingale and a stationary component. This decom-
position nest familiar decompositions for the macroeconomic time series literature and
the stochastic process literature on central limit approximation. This decomposition
identifies permanent shocks and increments to the martingale component.

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. Stochastic growth is 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
currency. 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) \textit{Long-run approximation} (section 5): I measure long-run growth and the associated value
decay through the construction of principal eigenvalues and principal eigenfunctions. I
use an extended version of Perron-Frobenius theory to establish this approximation. As I
will show, these objects give us a convenient characterization of long-run behavior. They
will also give us a way to formally define permanent and transitory model components.

v) \textit{Sensitivity and long-run pricing} (section 6): Of special interest is how the long-run
attributes of valuation change when we alter the growth processes or when we alter the
stochastic discount factor used to represent valuation. I show formally how to conduct
a sensitivity analysis with two applications in mind. We consider changes in the risk
exposure of hypothetical growth processes which give rise to long-run risk prices. I also
explore how long run values and rates of return are predicted to change as the attributes
of the economic environment are modified.

vi) \textit{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 feature some specifications for which
quasi-analytical characterizations are possible. I show when specific features of asset
pricing models have transient implications.
1.2 A revealing special case

Prior to my formal development, I will jump ahead a bit and illustrate steps iii and iv.

I characterize long-run stochastic growth (or decay) by posing and solving an approximation problem. Decay is relevant in my study of valuation because of the role of discounting. The approximation problem we consider borrows its origins from what is known as Perron-Frobenius theory of matrices and operators. To apply this idea in our setting we construct operators indexed by time and study the long run behavior of these operators. This allows us to do two things, a) extract a growth rate, a dominant eigenvalue, and b) construct a family of functions of the Markov state for which the large time version of the operators applied to those functions has an approximate one factor structure. The one factor is the principal eigenfunction. It is necessarily positive.

When a Markov process has a finite number of states, the mathematical problem that we study can be formulated in terms of matrices with nonnegative entries. Consider an \( n \times n \) matrix \( B \) with entries \( \{b_{ij}\} \). The entries of this matrix are constructed based in part on probabilities of transiting between the Markov states. Other inputs are state dependent growth rates, state dependent discount rates or both. The off diagonal entries are positive. Associated with this matrix, form an indexed family of matrices by calendar time:

\[
M_t = \exp(tB).
\]

The date \( t \) operator \( M_t \) governs the expected growth, discount or the composite of both over an interval of time \( t \). It will typically not be a probability matrix in our applications. (Column sums will not be restricted to be unity.) The matrix \( B \) connects the family operators in a special way. Given an \( n \times 1 \) vector \( f \), Perron-Frobenius theory characterizes limiting behavior of \( \frac{1}{t} \log M_t f \) by solving:

\[
Be = \rho e
\]

where \( e \) is an \( n \times 1 \) column eigenvector restricted to have strictly positive entries and \( \rho \) is a real eigenvalue. When the matrix \( B \) is not symmetric, we also consider the transpose problem

\[
B'e^* = \rho e^*
\]

where \( e^* \) also has positive entries. Depending on the application, \( \rho \) can be positive or negative. Importantly, \( \rho \) is larger than the real part of any other eigenvalue.

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

$$\lim_{t \to \infty} \frac{1}{t} \log M_t f = \rho$$

$$\lim_{t \to \infty} \exp(-\rho t) M_t f = (f \cdot e^*) e.$$ 

for any vector $f$ where we have normalized $e^*$ so that $e^* \cdot e = 1$. This formally defines $\rho$ as the long-run growth rate of the family of matrices $\{M_t : t \geq 0\}$. The eigenvector $e$ gives the direction that dominates in the long run.

I will use an extension of this method to determine model specifications which have important long-run affects on the matrix $B$ used in modeling instantaneous transition. Long run implications can be disguised in the construction of the local transitions. Our aim is to see through this disguise.

In this investigation, I will use operators and functions instead of matrices and vectors to accommodate continuous-state Markov processes. This will lead me to characterize dominance in a more general way. I will explore several different constructions of the operator counterpart to $B$, reflecting alternative hypothetical economic environments or alternative economic inputs. I will be interested in how $\rho$ and $e$ change as we alter $B$ in ways that are motivated explicitly through economic considerations. The operators I consider can have a complicated eigenvalue structure. I will avoid characterizing fully this structure, but instead I will use martingale methods that exploit representations of the operator families as I next describe.

I exploit representation results for operators based on stochastic processes built conveniently from the underlying Markov process. The representation that interests us is:

$$M_t f(x) = E \left[ M_t f(X_t) | X_0 = x \right]$$

(1)

where $X = \{X_t : t \geq 0\}$ is the Markov stochastic process and $M = \{M_t : t \geq 0\}$ is a positive stochastic process constructed as a functional of the Markov process $X$ in a restricted way. The operator $M_t$ maps a function $f$ of the Markov state into a function $f^*$ of the Markov state. Prior to forming the conditional expectation, the function of the Markov state $f(X_t)$ is scaled by the date $t$ component of the positive process $M$.

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

In addition to giving me a convenient way to incorporate economic structure in the analysis of valuation and growth, representation (1) also allows me to develop and exploit time series decompositions of the process $M$ as a device to deconstruct the family of operators $\{M_t : t \geq 0\}$. Decomposing the process $M$ into multiplicative components that separately reflect the growth, martingale and transient contributions will help us to characterize long run implications. In what follows, I will move freely between the operators $\{M_t : t \geq 0\}$ and the stochastic process $M$ used to represent them.

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 $X$ 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:

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

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 $\eta(dy|x)$ where $\int \eta(dy|X_{t-})$ is the jump intensity. That is for small $\epsilon$, $\epsilon \int \eta(dy|X_{t-})$ is the approximate probability that there will be a jump. The conditional measure $\eta(dy|x)$ 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 \((\mu, \sigma, \eta)\).

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

### 2.2 Convenient functions of the Markov process

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

An additive functional \(A\) is constructed so that \(A_{t+\tau} - A_t\) depends on \(X_u\) for \(t < u \leq t + \tau\). It is dependent on the underlying Markov process. For convenience, it is initialized at \(A_0 = 0\). Even if the underlying Markov process is asymptotically stationary, an additive functional will typically not be. Instead it will have increments that are asymptotically stationary, an additive functional can be normally distributed, but I will also interested other specifications. Conveniently, the sum of two additive functionals is additive.

Rather than give a general definition of an additive functional, I describe formally a family of such functionals parameterized by \((\beta, \xi, \lambda)\) where:

1. \(\beta : \mathcal{E} \to \mathbb{R}\) and \(\int_0^t \beta(X_u) du < \infty\) for every positive \(t\);
2. \(\xi : \mathcal{E} \to \mathbb{R}^m\) and \(\int_0^t |\xi(X_u)|^2 du < \infty\) for every positive \(t\);
3. \(\lambda : \mathcal{E} \times \mathcal{E} \to \mathbb{R}, \lambda(x, x) = 0\).

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

(2)

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

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

\[^1\]This latter implication gives the key ingredient of a more general definition of a multiplicative functional.
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).\textsuperscript{2} 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.

3.1 Additive decomposition

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

An additive functional can be decomposed into three components:

\textbf{Theorem 3.1.} Suppose that $Y$ is an additive functional with increments that have finite second moments. In addition, suppose that

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

and

$$\lim_{t \to \infty} E (Y_t - \nu t | X_0 = x) = g(x),$$

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

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

(3)

where $\{\hat{Y}_t\}$ is a martingale.

\textit{Proof.} Let $Y_t^* = Y_t - \nu t$. As a consequence of the Law of Iterated Expectations and the

\textsuperscript{2}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.
mean-square convergence,

\[ g(X_t) + Y_t^* = \lim_{s \to \infty} E \left( Y_{t+s}^* | X_t \right) \]
\[ = \lim_{s \to \infty} E \left[ E \left( Y_{t+s}^* - Y_{t+s}^* | X_{t+s} \right) + Y_{t+s}^* | X_t \right] \]
\[ = E \left[ g(X_{t+s}) + Y_{t+s}^* | X_t \right] \]

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.

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

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

with mean zero by the martingale central limit theorem.\(^3\) 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 martingale component that is identically zero. Blanchard and Quah (1989) use such a decomposition to identify shocks. The martingale increments are innovations to supply or technology shocks.

Recall the representation given in (2):

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

Form

\[ \tilde{\beta}(x) = \int_{\mathcal{E}} \lambda(y, x) \eta(dy|x). \]

Then

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

---

\(^3\)See Billingsley (1961) for the discrete-time martingale central limit. Moreover, there are well known functional extensions of this result.
is a local martingale. In what follows let
\[ \hat{\beta} = \beta + \tilde{\beta}. \]

**Theorem 3.2.** Suppose

i) \( X \) is a stationary, ergodic Markov process;

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

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

iv) There is a solution \( g \) to
\[
g(x) = \int_0^\infty E \left( \hat{\beta}(X_t) - E \left[ \hat{\beta}(X_t) \right] \mid X_0 = x \right);\]

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

This theorem gives an algorithm for computing \( \nu \) 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,
\[
\hat{\beta}(x) - \nu + \lim_{t \downarrow 0} \frac{1}{t} E \left[ g(X_t) - g(x) \mid X_0 = x \right] = 0,
\]
which gives an equation for \( g \) that depends on the local evolution of \( X \). The calculation
\[
\lim_{t \downarrow 0} \frac{1}{t} E \left[ g(X_t) - g(x) \mid X_0 = x \right] = A g(x)
\]
defines the so called generator \( A \) for the Markov process. Specifying generator \( A \) is one way to represent the transition dynamics for Markov process. In the case of a multivariate diffusion, this equation is known to be a second-order differential equation as an implication of Ito’s Lemma. There are well known extensions to accommodate jumps. Using the generator, the function \( g \) satisfies
\[
A g = \nu - \hat{\beta}.
\]
(4)

For the diffusion model, this leads to solving:
\[
\frac{\partial g(x)}{\partial x} \cdot \mu(x) + \frac{1}{2} \text{trace} \left[ \sigma(x) \sigma(x)'^T \frac{\partial^2 g(x)}{\partial x \partial x'} \right] = \nu - \hat{\beta}(x).
\]
(5)
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 $\hat{Y}$ 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 (4) depicted using the local evolution of the Markov process. Much is known about such an equation. As argued by Bhattacharya (1982) and Hansen and Scheinkman (1995), when $X$ is ergodic this equation has at most one solution. When $X$ is exponentially ergodic, there always exists a solution.\(^4\)

I now consider some examples.

**Example 3.3.** Suppose that

$$dX_t = AX_tdt + BdW_t,$$
$$dY_t = \nu dt + HX_tdt + FdW_t$$

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

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

which is a special case of (5). 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,

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

is the martingale component.

\(^4\)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 (4).
Example 3.4. Suppose that \( X \) and \( Y \) evolve according to:

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

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

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

which is a special case of (5). The solution is:

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

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

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

This example has the same structure as example 3.3 except that the Brownian motion shocks are scaled by \( \sqrt{X_t^2} \) to induce volatility that varies over time. While example 3.3 is fully linear, example 3.4 introduces a nonlinear volatility factor. More generally, additive functionals do not have to be linear functions of the Markov state or linear functions of Brownian increments. Nonlinearity can be built into the drifts (conditional means) or the diffusion coefficients (conditional variances). Under these more general constructions, the function \( g \) used to represent the transient component will not be a linear function of the Markov state.\(^5\)

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

In what follows we will use multiplicative functionals, processes whose logarithms can be represented conveniently as additive functionals. One strategy at our disposal is to decompose then exponentiate. Thus for $M_t = \exp(Y_t)$:

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

for the decomposition given in (3). While such a factorization is sometimes of value, for the purposes of my analysis, it is important that I construct an alternative factorization. The exponential of a martingale is not a martingale. If the process is lognormal, then this assumption can be used to transform $\exp(\hat{Y})$ into a martingale by scaling it by an exponential function of time. Later I will illustrate this outcome. Alternatively, the conditional normality of a diffusion process could be exploited, but this requires the use of a state dependent growth rate. Instead I will construct an alternative multiplicative decomposition that will be of direct use. As we will see, from a mathematical perspective, this decomposition has much closer ties to the theory of large deviations rather than central limit theory.

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

4 Limiting characterizations of stochastic growth or decay

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

A key step in our 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:

$$M_t f(x) = E [M_t f(X_t) | X_0 = x]$$ (6)

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

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

$$S_t f(x) = E [S_t f(X_t) | X_0 = x].$$

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 $S_t f$ depicts its current period value. Thus $M_t = S_t$ and $M = S$. The Law of Iterated Values restricted to this Markov environment is:

$$S_t S_\tau = S_{t+\tau}$$ (7)

for $t \geq 0, \tau \geq 0$ where $S_0 = I$, the identity operator. To understand this, the date $t$ price assigned to a claim $f(X_{t+\tau})$ is $S_\tau f(X_t)$. The price of buying a contingent claim at date 0 with payoff $S_\tau f(X_t)$ is given by the left-hand side of (7) 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 (7).

Alternatively, suppose that $E_t$ 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:

$$E_t f(x) = E [f(X_t) | X_0 = x]$$

As we will see other choices of $M$ can give rise to expectation operators provided that we are willing to alter the implicit Markov evolution. The Law of Iterated Expectations or the
Chain Rule of Forecasting implies:
\[ E_t E_\tau = E_{t+\tau} \]
for \( \tau \geq 0 \) and \( t \geq 0 \). In the case of conditional expectation operators, \( E_t 1 = 1 \) but this restriction is not necessarily satisfied for valuation operators.

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

**Definition 4.1.** A family of operators \( \{M_t\} \) is a (one-parameter) semigroup if a) \( M_0 = I \) and b) \( M_t M_\tau = M_{t+\tau} \) 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 \( \{M_t : t \geq 0\} \) constructed using (6) is a one parameter semigroup.

In valuation problems, stochastic discount factors are only one application of multiplicative functionals. Multiplicative functionals are also useful in building cash flows or claims to consumption goods that grow over time. While \( X \) may be stationary, the cash flow
\[ C_t = G_t f(X_t) \tilde{G}_0 \]
displays stochastic growth when \( G \) is a multiplicative functional. I include the adjustment \( \tilde{G}_0 \) because I normalized the the growth process to be one at date zero. Scaling by \( \tilde{G}_0 \) is a way to ensure that the initialization \( G_0 = 1 \) is indeed only a normalization. Moreover, shifting the vantage point from time zero to time \( t \),
\[ \frac{C_{t+\tau}}{G_t} = \left( \frac{G_{t+\tau}}{G_t} \right) f(X_{t+\tau}) \left( G_t \tilde{G}_0 \right). \]

I study cash flows of this type by building an operator that alters the transient contribution to the cash flow \( f(X_t) \). This leads us to study
\[ \mathbb{P}_t f(x) = E [G_t S_t f(X_t) | X_0 = x]. \]
The value assigned to \( C_t \) is given by \( \tilde{G}_0 \mathbb{P}_t f(X_0) \) because \( \tilde{G}_0 \) is presumed to be in the date zero information set. Importantly, it is the product of two multiplicative functionals that we use for representing the operator \( \mathbb{P}_t: M = GS \).

16
4.2 Some interesting limiting behavior

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

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

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

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

which justifies calling $\rho(M)$ a growth rate and provides a more refined approximation. Moreover, while the coefficient in this limit will depend on the choice of $f$, there is a limiting form of state dependence captured by the function $e$, which we will determine independently of $f$. I will say more about the admissible functions $f$ and about the coefficients in this approximation in our subsequent analysis. The essential point is that this rate will not depend on the function $f$ as long as it resides in a potentially rich class of such functions. Eventually, I will show how to represent the proportionality coefficient in the limit of (8) as as a linear functional of $f$.

The function $e$ is positive, and it solves:

$$E[M_t e(X_t) | X_0 = x] = \exp[\rho(M)t] e(x)$$

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

4.3 Products and covariation

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

$$\rho (M^{[1]} M^{[2]}) = \rho (M^{[1]}) + \rho (M^{[1]}) .$$
Covariation is important when characterizing even the limiting behavior of the product \( M^{[1]} M^{[2]} \). In fact the discrepancy:

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

will be used to give a long-run version of a risk premium. If \( M^{[1]}_t \) and \( M^{[2]}_t \) happen to be jointly log normal for each \( t \), then (10) is equal to the limiting covariance between the corresponding logarithms:

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

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

### 4.4 Local versus global

In the decomposition of an additive functional, the linear trend coefficient \( \nu \) is averages the local state dependent growth rate. I now explore the relation between the local, state dependent growth rate and the long run counterpart using some results from large deviation theory.

Consider for the moment a special class of multiplicative functionals:

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

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

The limit \( \rho(M) \) is a key ingredient in the study of large deviations. While \( \frac{1}{t} \int_0^t \beta(X_u) du \) may obey a Law of Large Numbers and converge to its unconditional expectation under the stationary distribution, more can be said about small probability departures from this law. Large deviation theory seeks to characterize these departures by evaluating expectations under the stationary distribution for an alternative probability measure assigned to \( X \). Let \( Q \) be a probability distribution over the state space \( E \) of the Markov process \( X \). Following Donsker and Varadhan (1976), Dupuis and Ellis (1997) and others, a rate function \( \mathbb{H}(Q) \) is constructed 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 $I$ is convex in the probability measure $Q$, and it is used to construct what is called a Laplace principle that characterizes the limit:

$$
\rho(M) = \sup_Q \int \beta(x)dQ - I(Q) \geq E[\beta(X_t)]
$$

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

This optimization problem is inherently static, with the dynamics loaded into the construction of convex function $I$. The function $I$ is constructed independent of the choice of $\beta$. Recall that $\beta$ 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 (11) characterizes formally this difference.$^6$

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

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

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

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

as in (2), the local growth rate is

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

Exposure to Brownian motion risk alters the local growth rate.

The multiplicative functional $M$ can be decomposed into two component multiplicative functionals:

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

$^6$Large deviation theory exploits problem (11) because $\rho(M)$ implies a bound of the form:

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

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

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

(13)

where $I^*$ is constructed using the change in probability measure. While the inequality associated with optimization problem 11 is satisfied, it is satisfied only after the change measure. The average local growth rate could be greater than the long-run growth rate computed under the original probability measure when there the multiplicative functional is exposed locally to risk.

## 5 Multiplicative factorization

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

$$M_t = \exp(\rho t) \tilde{M}_t \left[ \tilde{\epsilon}(X_t) \over \tilde{\epsilon}(X_0) \right]$$

↑ ↑ ↑

growth martingale transient

(14)

Component a) governs the long-term growth or decay. It is constructed from a principal eigenvalue. I will use component b), the positive martingale, to build an alternative probability measure. This alternative measure gives us a tractable framework for a formal study

---

7In 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.

8The 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).
of approximation. Component c) is built directly from the principle eigenfunction and characterizes transient departures in behavior that are distinct from martingale behavior.

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

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

### 5.1 Decomposition

I build the decomposition as follows. First I solve:

\[
E \left[ M_t e(X_t) | X_0 = x \right] = \exp(\rho t) e(x) \tag{15}
\]

for any \( t \) where \( e \) is strictly positive as in (9). The function \( e \) can be viewed as a principal eigenfunction of the semigroup with \( \rho \) being the corresponding eigenvalue. Since this equation holds for any \( t \), it can be localized by computing:

\[
\lim_{t \downarrow 0} \frac{E \left[ M_t e(X_t) | X_0 = x \right] - \exp(-\rho t) e(x)}{t} = 0, \tag{16}
\]

which gives an equation in \( e \) and \( \rho \) to be solved. The local counterpart to this equation is

\[
\mathbb{B} e = \rho e, \tag{17}
\]

where

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

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

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

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

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

We have seen the finite-state counterpart to this equation in subsection 1.2.

Typically it suffices to solve the local equation (17) to obtain a solution to (15). See Hansen and Scheinkman (2007) for a more detailed discussion of this issue. In the finite-state Markov model of subsection 1.2, convenient and well known sufficient conditions exist for there to be a unique (up to scale) positive eigenfunction satisfying (15). More generally, however, this uniqueness will not hold. Instead I will obtain uniqueness from additional considerations.

Given a solution to (15), 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 (14) follows immediately by letting \( \hat{e} = \frac{1}{e} \) and solving for \( M \) in terms of \( \hat{M} \), \( \rho \) and \( \hat{e} \).

5.2 Additive versus multiplicative decomposition

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

The simplest case is the log normal example.

**Example 5.1.** 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 case it is easy to go from a martingale decomposition of an additive functional to that of a multiplicative functional.

An equivalent way to proceed is to build $e$ as an exponential of a linear function of $x$, and to seek a solution to (17). 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 $\rho$ includes an extra volatility adjustment as is typical in log-normal models.

This simple link between additive and multiplicative decompositions is no longer valid when volatility is state dependent.

**Example 5.2.** Consider again Example 3.4, and recall the additive functional:

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

Form

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

Guess a solution $e(x) = \exp(\alpha \cdot x)$ where $x = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$ and $\alpha = \begin{bmatrix} \alpha_1 \\ \alpha_2 \end{bmatrix}$. To compute $\rho(M)$, I solve a special case of (17):

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

which I derive as a special case of (17). Thus the coefficients on $x_1$ and $x_2$ are zero when:

$$A_1' \alpha_1 + H_1 = 0$$

$$A_2 \alpha_2 + H_2 + \frac{1}{2} |\alpha' B_1 + \alpha_2 B_2 + F|^2 = 0. \quad (18)$$

The first equation can be solved for $\alpha_1$ and the second one for $\alpha_2$ given $\alpha_1$. The solution to
the first equation is:

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

The second equation is quadratic in \( \alpha_2 \), so there may be two solutions. Specifically,

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

(19)

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

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

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:

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

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

**Definition 5.3.** The process \( X \) is stochastically stable under the measure \( \hat{\cdot} \) if

\[ \lim_{t \to \infty} \hat{E} [f(X_t)|X_0 = x] = \hat{E} [f(X_t)] \]  

(20)

for any \( f \) for which \( \hat{E}(f) \) is well defined and finite.\(^9\)

\(^9\)This is stronger than ergodicity because it rules out periodic components. Ergodicity requires that time series averages converge but not necessarily that conditional expectation operators converge. Under ergodicity the time series average of the conditional expectation operators would converge but not necessarily the conditional expectation operators.
Theorem 5.4. Given a multiplicative functional $M$, suppose that $e$ and $\rho$ satisfy equation (16) and that $X$ is stochastically stable under the $\hat{\cdot}$ probability measure. Then

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

Moreover,

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

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

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

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

defined previously.

It follows from Theorem 5.4 that once we scale by the growth rate $\rho$, we obtain a one-factor representation of long-term behavior. Changing the function $f$ simply changes the coefficient on the function $e$. Thus the state dependence is approximately proportional to $e$ as the horizon becomes large. For this method to justify our previous limits, we require that $f\hat{e}$ have a finite expectation under the $\hat{\cdot}$ probability measure. The class of functions $f$ for which this approximation works depends on the stationary distribution for the Markov state of the $\hat{\cdot}$ probability measure and the function $\hat{e}$. These functions of the Markov state have transient contributions to valuation since for these components:

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

Definition 5.5. For a given multiplicative functional $M$, a process $f(X)$ is transient if $X$ is stochastically stable under the probability measure implied by the martingale component and $\hat{E}[f(X_t)\hat{e}(X_t)]$ is well defined and finite.

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

There is an extensive set of tools for studying the stability of Markov processes that can be brought to bear on this problem. For instance, see Meyn and Tweedie (1993) for a survey of such methods based on the use of Foster-Lyapunov criteria. See Rosenblatt (1971), Bhattacharya (1982) and Hansen and Scheinkman (1995) for alternative approaches based on mean-square approximation. While there may be multiple representations of the form (14), Hansen and Scheinkman (2007) show that there is at most one such representation for which the process \( X \) is stochastically stable.

Recall that in example 5.2 we found two solutions for \( \alpha_2 \) by solving the quadratic equation (18). As an implication of the Girsanov Theorem, associated with each solution is an alternative probability measure under which

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

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

\[
dX_t^{[2]} = A_2 X_t^{[2]} dt + (B_2 F' + |B_2|^2 \alpha_2) X_t^{[2]} dt + \sqrt{X_t^{[2]}} d\hat{W}_t
\]

where in the second representation I have substituted from solution (19). I select the “minus” solution to achieve stochastic stability.

5.4 Long-run behavior of multiplicative martingales

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

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

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

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

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

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

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

### 5.5 Transient model components

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

Consider a benchmark valuation model represented by a stochastic discount factor or a benchmark growth process or the product of the two components. I now ask what modifi-

---

10See Newman and Stuck (1979) for a continuous-time Markov version of this formulation.
Considerations are transient? The tools I describe in section 5 give an answer to this question.

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

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

Moreover suppose that under the associated \hat{\cdot} probability measure $X$ satisfies a stochastic stability condition 5.3. Consider an alternative model of the form:

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

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

I use the multiplicative decomposition of $M$ to construct an analogous decomposition for $M^*$. Given the decomposition:

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

the corresponding decomposition for $M^*$ is:

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

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

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

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

6 Perturbation calculation

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

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

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

For future reference note that

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

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

In the remainder of this section, I give a tractable formula for the derivatives of interest.

### 6.1 Finite horizon

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

$$
M_t(a)f(x) = E [M_t(a)f(X_t)|X_0 = x].
$$
Then under some regularity conditions,
\[
\frac{d}{da} \log M_t(a) f(x)|_{a=0} = \frac{E \left[ M_t(a) \frac{d}{da} \log M_t(a) f(X_t)|X_0 = x \right]_{a=0}}{E \left[ M_t(a) f(X_t)|X_0 = x \right]_{a=0}}.
\]

### 6.2 Limiting behavior

I now show that these perturbations have a simple structure when we focus on long-run implications. Specifically, I compute:

\[
\frac{d}{da} \rho(a). \tag{22}
\]

Calculation (22) turns out to be straightforward as we will now see. First solve the principal eigenvalue problem for $a = 0$ and use the solution to construct a probability measure $\hat{\cdot}$ as we described previously. The formula for the derivative is:

\[
\frac{d}{da} \rho(a)|_{a=0} = \frac{1}{t} \hat{E} \left[ \frac{d \log M_t(a)}{da} \right]_{a=0}. \tag{23}
\]

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

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

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

I let

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

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

The functional $\log M_t(a)$ is an additive functional, and its derivative is as well. Recall the continuous time model of $Y$ we specified in equation (2):

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

and form $M(a) = \exp[Y(a)]$. It is most convenient to take limits of (23) as $t \to 0$. This
entails computing an average local mean under the distorted distribution:

\[
\frac{d}{da} \rho(a)|_{a=0} = \hat{E}\left( \frac{d}{da} \left[ \beta(X_t; a) + \xi(X_t; a) \cdot \hat{\xi}(x) \right] |_{a=0} \right) \\
+ \hat{E} \left[ \int \frac{d}{da} \lambda(y, X_t; a) \left|_{a=0} \right. \exp[\hat{\lambda}(y, X_t)] \eta(dy|X_t) \right]
\]

(24)

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

To understand the reason for this simple formula, recall the decomposition:

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

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

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

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

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

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

Now use the decomposition and differentiate log \( M_t(a) \)

\[
\frac{d}{da} \log M_t(a) = t \frac{d\rho(a)}{da} + \frac{d}{da} \log \hat{M}_t(a) - \frac{d}{da} \log e(X_t; a) + \frac{d}{da} \log e(X_0, a).
\]

Take expectations and use the fact that \( X \) is stationary under the \( \hat{\cdot} \) probability measure to obtain derivative formula (23). See appendix B for a discussion of the principal eigenfunction.

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

In our study of asset pricing, we consider two limits. One will reproduce local risk prices familiar from asset pricing theory by taking limits as the investment horizon shrinks to zero, and the other is the limit as the investment horizon is made arbitrarily large.

For a given investment horizon I study the risk premium as measured by:

\[
\frac{1}{t} \log E \left[ G_t f(X_t) | X_0 = x \right] - \frac{1}{t} \log E \left[ S_t G_t f(X_t) | X_0 = x \right] + \frac{1}{t} \log E \left[ S_t | X_0 = x \right] \tag{25}
\]

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

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

is the expected return on the investment over the horizon \( t \), and

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

is the expected return on a riskless investment.

By letting \( t \) shrink to zero and computing marginal changes in the risk exposure, we obtain local risk prices. The methods described in this paper permit us to study the limit as the horizon is made arbitrarily large and to explore the corresponding marginal changes in risk exposure. Provided that \( f \) is inconsequential to the limit, the large horizon limit is

\[
\rho(G) + \rho(S) - \rho(SG).
\]

In what follows, I let \( G \) be a multiplicative martingale which is justified in part by our multiplicative decomposition. For such a \( G \), \( \rho(G) = 0 \). Alternative models exist for the stochastic discount factors as we now explore.

Prior to proceeding, I comment a bit on the previous literature. The study of the dynamics of risk premia is familiar from the work of Wachter (2005), Lettau and Wachter (2007) and Hansen et al. (2008). Hansen et al. (2008) characterize the resulting limiting risk premia and the associated risk prices in a log-linear environment.\(^{11}\) Hansen and Scheinkman

\(^{11}\)Hansen et al. (2008) also consider the limiting behavior of holding period returns. This limit includes
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.\textsuperscript{12}

### 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:

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

Many alterations in this model take the form:

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

Arguably transient components in asset pricing 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 as an extension of power utility.

In what follows we follow Campbell and Cochrane (1999) by exploring a simple model of consumption dynamics under which the power utility model has transparent implications.

### 7.2 Power utility without predictability

Suppose that consumption is a geometric Brownian motion:

\[ d \log C_t = \mu_c dt + \sigma_c dW_t, \]

contributions from the principle eigenfunction and the principal eigenvalue of the associated valuation operator for pricing cash flows with stochastic growth components.

\textsuperscript{12} Wachter (2005) develops a computational approach based on pricing what she calls “zero-coupon” equity, which in our notation is \( E [S_t G_t \hat{f}(X_t)|X_0 = x] \). Her algorithm has component prices that converges to zero as the horizon is extended. By using an adaptation of the so-called “power method”, these prices can be rescaled to have nondegenerate limit. The limiting function is a principal eigenfunction of the type that I have described. The power method rescales each iteration and hence adjusts for the asymptotic decay. The limit of this rescaling reveals the eigenvalue. Using this more refined characterization of the limit could improve computational performance, and the results of Hansen and Scheinkman (2007) provide justification for the limit approximation.

33
where \( c_t \) is the logarithm of aggregate consumption. I allow the Brownian motion \( \{ W_t : t \geq 0 \} \) to be multivariate. Construct \( S \) in accordance with the power utility model:

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

where \( \frac{1}{\gamma} \) is intertemporal elasticity of substitution and \( \delta \) is the subjective rate of discount. Does this new component provide a transient departure from the power utility model? For risk pricing, I introduce a growth functional that is a martingale:

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

Choosing \( G \) to be a martingale, eliminates the transient dynamics in the implied cash-flow risk exposure and allows me to focus on the the pricing dynamics as they contribute to risk premia.\(^{13}\) This martingale simplifies risk premium formula (25), as I mentioned earlier.

I now explore the risk prices. First I use the benchmark \( S \) model and the growth process \( G \) to construct a martingale. Since consumption is geometric Brownian motion, this is accomplished as:

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

where

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

It follows that

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

By setting \( \sigma_g = 0 \),

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

Thus

\[
\rho(G) + \rho(S) - \rho(GS) = \gamma \sigma_c \cdot \sigma_g.
\]

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

The dynamics of pricing for this example is degenerate, and in particular the local risk

\(^{13}\)While multiplicative martingales may have degenerate long-run behavior, we could apply Theorem 3.2 and eliminate the trend grow term in logarithms. This allows for central-limit-type behavior for long horizons, and it does not alter the implied risk premia and corresponding risk prices.
price vector is also equal to \( \gamma \sigma_c \). The local and long-term prices are the same as I now show. Specifically,

\[
\lim_{t \to 0} \frac{1}{t} \log E \left[ G_t S_t | X_0 = x \right] = -\delta - \gamma \mu_c + \frac{\gamma^2}{2} |\sigma_c|^2 - \gamma \sigma_c \cdot \sigma_g. \tag{26}
\]

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

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

### 7.3 An example of Campbell and Cochrane

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

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

We may rewrite this as:

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

In what follows let

\[
X_t = -\log(1 - C_t^*/C_t) - b,
\]

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

\[
S_t^* = S_t \left[ \frac{\exp(\gamma X_t)}{\exp(\gamma X_0)} \right].
\]
Following Campbell and Cochrane (1999) and Wachter (2005), assume that

$$dX_t = -\xi(X_t - \mu_x)dt + \lambda(X_t)\sigma_c dW_t$$  \hspace{1cm} (27)$$

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

$$dX_t = -\xi(X_t - \mu_x)dt + (\sigma_g - \gamma\sigma_c) \cdot \sigma_c \lambda(X_t)dt + \lambda(X_t)\sigma_c d\hat{W}_t. \hspace{1cm} (28)$$

We use this evolution to compute the counterpart to (26):

$$\lim_{t \downarrow 0} \frac{1}{t} \log E[G_t S^*_t | X_0 = x] = \lim_{t \downarrow 0} \frac{1}{t} \log E[G_t S_t \exp[\gamma(X_t - X_0)] | X_0 = x]$$

$$= -r - \gamma \sigma_c \cdot \sigma_g - \lim_{t \downarrow 0} \frac{1}{t} \hat{E}[\exp[\gamma(X_t - X_0)] | X_0 = x]$$

$$= -r - \gamma \sigma_c \cdot \sigma_g + \gamma \xi(x - \mu_x) + \gamma(\sigma_g - \gamma\sigma_c) \cdot \sigma_c \lambda(x)$$

$$+ \gamma^2 \frac{\lambda(x)^2 |\sigma_c|^2}{2} \hspace{1cm} (29)$$

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

### 7.3.1 Interest rates

The instantaneous interest is:

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

which follows from (29) by setting \(\sigma_g = 0\). 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 \(\theta\) controls the variation in the risk-free rate. Thus

$$r^* + \theta(x - \mu_x) = r + \gamma \xi(x - \mu_x) - \gamma^2 \frac{\lambda(x)^2 |\sigma_c|^2}{2} + \gamma^2 |\sigma_c|^2 \lambda(x). \hspace{1cm} (30)$$

I infer the value of \(r^*\) by setting \(x = 0\):

$$r^* = r + (\theta - \gamma \xi)\mu_x$$
Substituting this formula into (30), by a simple complete-the-square argument:

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

Thus

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

\[\zeta \equiv \frac{2(\gamma\xi - \theta)}{\gamma^2|\sigma_c|^2}\]

I take the negative square root in order that \(\lambda(0) = 0\). In order that the term inside the square root be positive, we must restrict \(\theta\) so that \(\theta < \gamma\xi\).

### 7.3.2 Local risk prices

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

\[\gamma\sigma_c - \gamma\lambda(x)\sigma_c = \gamma (1 + \zeta x)^{1/2}\sigma_c\]

which follows because (29) is affine in \(\sigma_g\) and the risk prices are the negative of the partial derivative with respect to \(\sigma_g\). By design are state dependent and are larger than in the power utility model for given value of \(\gamma\). Moreover, the state variable increment \(dX_t\) responds negatively to consumption growth shocks because \(\lambda(x) < 0\). By design, risk premia are larger in bad times as reflected by unexpectedly low realizations of consumption growth. As demonstrated by Campbell and Cochrane (1999) in their closely related discrete time model, the coefficient of relative risk aversion is also enhanced and in fact it is equal to \(\gamma[1 - \lambda(\mu_x)]\) for \(\mu_x = x\). (See also appendix C.)

### 7.3.3 Long-term pricing

Consider now the long-run behavior of value. I use evolution equation (28), and the formula for the logarithmic derivative of the density for a scalar diffusion:

\[\frac{d\log q}{dx} = 2 \text{ drift diffusion} - \frac{d\log \text{ diffusion}}{dx}\]

(31)

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

$$\lim_{x \to \infty} \frac{d \log q}{dx} = -\frac{\gamma^2 \xi}{\gamma \xi - \theta} < 0.$$  (32)

As a consequence the process $X$ is stationary under the twisted probability measure and under the original probability measure as reflected by (27) and (28) 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 (32) is strictly less than $-\gamma$. In contrast, when $\theta < 0$ this expectation will be infinite. Thus when $\theta > 0$ the contribution to preferences will be transient, but not when $\theta < 0$.

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

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

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

We may now characterize the limiting risk premia:

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

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

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

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

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

but $\rho(S^*)$ differs and is given by the implied real interest rate $r^*$. Thus

**risk premium** $= \gamma \sigma_c \cdot \sigma_g + r - r^*$.  (33)
Viewing this risk premium as a function of $\sigma_g$, there is an implied “discontinuity” at $\sigma_g = 0$. This discontinuity is depicted in figure 1 in which we depict the risk premia when investors care about external habits and when they do not. Typically the risk premia converges to zero as $\sigma_g$ 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.\footnote{I make this comparison of interest rates as device to characterize risk premia for the Campbell and Cochrane (1999) model. I hold fixed parameters such as the subject rate of discount. If the aim is fit a given risk-free rate, then the parameters in the two models would be set differently. 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.} In figure 1 this discontinuity is sizable. It is the distance on the vertical axis between the circle and the dot. While this discontinuity is only present in the limit, it is indicative that risk prices are large near $\sigma_g = 0$ for valuation over long time horizons.

### 7.3.4 Two martingales

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

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

It is one of two such solutions since

$$E [S_t|X_0 = x] = \exp(-rt).$$

The multiplicative martingale

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

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

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

When $\theta = 0$ the multiplicative martingale $\tilde{M}$ is the pertinent one for discount bond pricing the martingale $\hat{M}$ for pricing growth rate risk over long horizons. The discontinuity
Figure 1: Risk premia as function of risk exposure. The horizontal axis is given in quarterly time units. The vertical axis is scaled by one hundred so the risk premia are in percent. The dot-dashed line denotes the implied premia when investors have external habits, and the solid line denotes the implied premia when investors have expected utility preferences. The parameter values for the state evolution are: $F c = 0.0054$ and $\nu c = .0056$. I set $\gamma = 2$, and for the model with investors that have external habits I set $\theta = 0$ and $\xi = .035$. 
in the long-term risk premia as a function of $\sigma_g$ as expressed in (33) 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 An example of Santos and Veronesi

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

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

In this case

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

for some positive number $b$.

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

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

where

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

and $\mu_x > 0$.\(^{15}\) The specification of local volatility is designed to keep the process $X$ above unity. As in the Campbell and Cochrane (1999) specification, the process $X$ responds negatively to a consumption shock. The stationary density is in the stable class with an algebraic tail, and the process is rho-mixing with mixing coefficients that decay exponentially.\(^{16}\)

The local risk prices are now given by

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

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

To study long-term pricing we again use the twisted evolution equation (28) but with

\(^{15}\)This process is the F process of Wong (1964).

\(^{16}\)See Hansen and Scheinkman (1995).
this new specification of $\lambda$. Formula (31) is again informative. In particular,

$$\lim_{x \to \infty} x \frac{d \log q(x)}{dx} = -2 \frac{\xi}{v^2} - 2 \frac{(-\gamma \sigma_c + \sigma_g) \cdot \sigma_c}{v |\sigma_c|} \neq -r$$

The twisted density has tails that behave like $x^{-r}$ for large $x$. For this to be a valid density $r > 1$; for $X_t$ have a finite expectation under the twisted distribution $r > 2$. Provided that $r > 2$, the long-term risk prices will agree with the power utility model. Thus the long-run behavior of the risk prices are potentially quite different from those pertinent for the Campbell-Cochrane specification.

### 7.5 Predictability in consumption growth and volatility

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

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

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

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

Consider a growth functional constructed as a martingale:

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

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

$$H_1 = -\gamma H_c \quad H_2 = -\frac{1}{2} |F_g|^2 \quad F = -\gamma F_c + F_g \quad \nu = -\delta - \gamma \mu_c - \frac{1}{2} |F|^2.$$
Then this specification of $Y$ is a special case of example 3.4. Applying formulas (24) and (21), the long-run risk price is the expected drift under the twisted measure induced by $\hat{M}$ of

$$F_{g}X_{t}^{[2]}dt - \sqrt{X_{t}^{[2]}dW_{t}}$$

where

$$dW_{t} = \sqrt{X_{t}^{[2]}}\left[F + (B_{1})'\alpha_{1}^{[b]} + (B_{2})'\alpha_{2}^{[b]}\right]dt + d\hat{W}_{t}$$

where $\hat{W}$ is a multivariate standard Brownian motion under this alternative measure and the positive eigenfunction is $\exp(\alpha_{1}^{[b]} \cdot x_{1} + \alpha_{2}^{[b]}x_{2})$. Thus the long-run risk price vector is:

$$\hat{E} \left( X_{t}^{[2]} \right) \left[ \gamma F_{c} - (B_{1})'\alpha_{1}^{[b]} - (B_{2})'\alpha_{2}^{[b]} \right]$$

local growth volatility

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

$$dX_{t}^{[2]} = A_{2}(X_{t}^{[2]} - 1)dt + B_{2}\left[F + (B_{2})'\alpha_{2}^{[b]}\right]X_{t}^{[2]}dt + \sqrt{X_{t}^{[2]}d\hat{W}_{t}}$$

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

$$dX_{t}^{[2]} = \hat{A}_{2}\left(X_{t}^{[2]} - \hat{\mu}_{2}\right)dt + \sqrt{X_{t}^{[2]}d\hat{W}_{t}}$$

where

$$\hat{A}_{2} = A_{2} + B_{2}\left[F + (B_{2})'\alpha_{2}^{[b]}\right]$$

$$\hat{\mu}_{2} = \frac{A_{2}}{\hat{A}_{2}} = \hat{E} \left( X_{t}^{[2]} \right).$$

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

$$\hat{E} \left( X_{t}^{[2]} \right) \gamma F_{1}$$

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

$$-\hat{E} \left( X_{t}^{[2]} \right) (B_{1})'\alpha_{1}^{[b]} = -\gamma \hat{E} \left( X_{t}^{[2]} \right) (B_{1})'[(A_{1})']^{-1}(Hc)'$$

(34)
reflects the temporal dependence in the growth rate of consumption, as featured in the long-
term pricing calculations by Hansen et al. (2008). The third term reflects the temporal
dependence in volatility.

7.6 Risk Sensitivity and Recursive Utility

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

As in section 7.5, we allow for predictability in consumption. Consumption is itself a mul-
tiplicative process. Consider now discounted logarithmic consumption and apply integration-
by-parts:

$$\delta \int_0^T \exp(-\delta u) \log C_u du = \int_0^T \exp(-\delta u) d \log C_u + \log C_0 - \exp(-\delta T) \log C_T.$$ 

We consider the limiting case in which the discount rate is zero, which leads us to study the
additive process

$$\log C_T = \int_0^T d \log C_u + \log C_0$$

where we can view $T$ as the time horizon of the economy that will eventually be made large.

Instead of using logarithmic utility, I follow the literature on risk-sensitive control theory
by modeling investor preferences as:

$$\log C_0 + \frac{1}{1-\gamma} \log \mathbb{E} \left[ \exp \left( (1-\gamma) \left( \int_0^T d \log C_u \right) \right) \bigg| \mathcal{F}_0 \right]$$

while maintaining a unitary elasticity of substitution. These preferences are recursive pro-
vided that the date $t$ preferences are constructed using:

$$\log C_t + \frac{1}{1-\gamma} \log \mathbb{E} \left[ \exp \left( (1-\gamma) \left( \int_t^T d \log C_u \right) \right) \bigg| \mathcal{F}_t \right]$$

Hansen and Sargent (1995) give a recursive utility counterpart that accommodates dis-
counting. Risk-sensitive control typically embraces the discounting and solves the date zero
problem of the investor (see Whittle (1990) for a discussion of the role of discounting). An
undiscounted stochastic version was studied by Runolffson (1994) without allowing for
stochastic growth. As an alternative, Kihlstrom (2007) considers a dynamic Markov game
in played by agents at different time periods as a way to confront a change of perspective of
investor preferences.\textsuperscript{17} In the $\delta = 0$ limit, these approaches coincide.

Form the multiplicative process:

$$V_t = \exp \left[ (1 - \gamma) \left( \int_0^t d \log C_u d u \right) \right].$$

The stochastic discount factor for the risk sensitive economy is:

$$S_{t, T}^* = \left( \frac{C_0}{C_t} \right) \frac{E (V_T | \mathcal{F}_t)}{E (V_T | \mathcal{F}_0)}$$

The methods described in this paper allow us to study limits as the time horizon $T$ is extended to $\infty$. Factor $V$ as:

$$V_t = \exp \left( \rho^{[r]} t \right) \tilde{M}_t \left[ \frac{e^{[r]}(X_0)}{e^{[r]}(X_t)} \right]$$

where $\tilde{M}_t$ is a martingale that implies stable dynamics for $X$, $\rho^{[r]}$ the corresponding principal eigenvalue and $e$ the positive eigenfunction.\textsuperscript{18} Formulas for these are obtained as a special case of the formulas in example 5.2. Specifically, $e^{[r]}(x) = \exp(\alpha^{[r]} x)$ for some coefficient vector $\alpha^{[r]}$ and let $\rho^{[r]}$ denote the corresponding eigenvalue. Notice that

$$\lim_{T \to \infty} \exp(-\rho^{[r]} T) E \left( \frac{V_T}{V_t} | X_t \right) = \exp(-\rho^{[r]} t) \tilde{E} \left[ \frac{1}{e^{[r]}(X_t)} \right] e^{[r]}(X_t)$$

where $\tilde{E}$ is the expectation operator associated with the probability measure induced by $\tilde{M}$. As a consequence, the stochastic discount factor for the limit economy is:

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

The Breeden (1979) stochastic discount factor can be expressed as:

$$S_t = V_t \left( \frac{C_0}{C_t} \right)$$

and hence

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

\textsuperscript{17}Like Epstein and Zin (1989), Kihlstrom (2007) is specifically interested in models in which the intertemporal elasticity of substitution differs from unity.

\textsuperscript{18}Even though we have introduced stochastic growth in consumption, there is direct counterpart to $\eta$ and $e$ in Runolffson (1994)'s analysis of stochastic risk sensitive control in the absence of discounting.
This representation suggests that the adjustment to preferences is transient except for a constant adjustment to the interest rate. The multiplicative martingale components coincide.

In terms of local prices, $\rho[^r]$ alters the interest rate and the eigenfunction $e$ alters the local risk prices \textit{vis a vis} the Breeden (1979) model. These prices are given by

$$\sqrt{X_t[2]} \left[ \gamma F_1 - (B_1)'\alpha_1[^r] - (B_2)'\alpha_2[^r] \right].$$

The term

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

and is familiar from the analysis in Bansal and Yaron (2004), Campbell and Vuolteenaho (2004) and Hansen et al. (2008). It is a recursive utility enhancement of the local risk prices based on predictability in consumption growth rates. There is an additional local price adjustment for consumption predictability. The long-term risk price calculation given in (34) continues to apply to this model even though the local prices are different from the Breeden (1979) model.\footnote{Since Hansen et al. (2008) consider models without stochastic volatility, this distorted expectation plays no role.}

In figure 2 I depict the risk-prices indexed by the investment horizon for a three shock version of the consumption dynamics. These trajectories converge to the limit prices that I just characterized. Only the first shock has a direct impact on consumption; the second shock alters the growth rate in consumption via a continuous-time scalar autoregression ($X[^1]$), and the third shock alters volatility via a Feller-square root process ($X[^2]$). A positive movement in the third shock diminishes consumption volatility. The formula for risk prices are given in appendix (D). In the Breeden (1979) model, the local risk prices are zero for the second two shocks. Investor preferences are forward looking in the recursive utility model, and this is reflected in nonzero local risk prices for the second two shocks. The enhanced local price of the growth rate shock illustrates the pricing mechanism featured by Bansal and Yaron (2004), and the similarity of the risk prices over long-horizons between the Breeden (1979) model and the recursive utility model illustrates a finding in Hansen et al. (2008). The forward-looking feature of recursive preferences leads to a flatter trajectory for the risk prices. The trajectory is literally flat for the first shock and the two models imply the same risk prices. The coincidence of the pricing trajectories for the two models of investor preferences illustrates a point made by Kocherlakota (1990). The persistence of large local risk prices for the consumption growth rate shock over long horizons is consistent with the empirical findings of Hansen et al. (2008), although their model of consumption dynamics is
Figure 2: Risk prices indexed by investment horizon. The horizontal axis is given in quarterly time units. The solid line denotes recursive utility model, and dashed line the expected utility model. The parameter values for the state evolution are: $A_1 = -0.02$, $A_2 = -0.02$, $B_1 = [0 \ 0.00047 \ 0]$, $B_2 = [0 \ 0 \ 0.038]$, $H_c = [1 \ 0]$, $F_c = [0.0054 \ 0 \ 0]$ and $\nu_c = 0.0056$. The risk prices are by localizing around $G = C$. For illustrative purposes I set $\gamma = 10$. How to “calibrate” $\gamma$ 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).
little different.\textsuperscript{20} I specified stochastic volatility to be very persistent, and this is reflected in the slow rate of convergence of the risk prices. While stochastic volatility induces variation in local risk prices, the shock to volatility like the direct shock to consumption commands a relatively small risk price at all horizons.

The preceding analysis exploits two important restrictions on investor preferences. The intertemporal elasticity of substitution is unity and the subjective rate of discount is zero. A natural extension is to compute two additional “derivatives” as a device to study the impact of changing investor preferences. For the long term risk prices, this can be done using the perturbation method described in section 6. Hansen et al. (2008) have used this method to explore changes in the intertemporal elasticity of substitution.\textsuperscript{21}

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

8 Conclusion

To conclude I want to be clear on two matters.

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

Second, while my discussion of statistical approximation has been notably brief, I do not have to remind time series econometricians of the particular measurement challenges associated with the long run. Indeed there is a substantial literature on such issues. In

\textsuperscript{20}Hansen et al. (2008) abstract from stochastic volatility and they use a discrete-time vector autoregressive model of consumption and corporate earnings to model the consumption dynamics.

\textsuperscript{21}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), expected utility as in Kihlstrom (2007), or risk-sensitivity under commitment.
part my aim is to suggest an econometric framework for the use of such measurements. But some of the measurement challenges remain. My own view is that many of the same statistical challenges that we as econometricians struggle with should be passed along to the hypothetical investors that populate our economic models. Difficulties in selecting a statistical model to use in extrapolation and associated ambiguities in inferences may well be an important component to the behavior of asset prices.
A A static max-min problem

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

$$\varrho = \sup_{Q \in \mathcal{Q}} \inf_{f \in \mathcal{D}^+} \int \left( \frac{\mathbb{B} f}{f} \right) dQ.$$  \hfill (35)

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

$$B f(x) = \beta(x) f(x) + A f(x)$$

where

$$\beta^*(x) = B 1(x)$$

$$A f(x) = B f(x) - \beta^*(x) f(x).$$

Notice that by construction $A f = 0$ when $f$ is a constant function. Suppose that $A$ 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 $\beta$ as the local growth or decay rate for the semigroup.

Consider the first the inner minimization problem of (35). Split the objective and write:

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

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

$$I^*(Q) = \sup_{f \in \mathcal{D}^+} - \int \left( \frac{A f}{f} \right) dQ$$  \hfill (36)

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

---

\textsuperscript{22}While the function 1 does not vary over states the outcome applying $\mathbb{B}$ to 1 will typically vary with $x$ and hence the notation $\mathbb{B} 1(x)$.
between probabilities when the process implied by $A$ possess a stationary distribution. The measure is relative because it depends on the generator $A$ of a Markov process and measure discrepancies from the stationary distribution of this process.

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

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

which is the problem posed in (13).

Suppose that the solution to the max-min problem is attained with probability measure $Q^*$. Consider again the inner optimization problem (36) and suppose that the supremum is attained at $f^*$. Let $g$ be any other function in the domain of $B$ such that $f^* + rg$ is strictly positive for some $r$. For instance, if $f^*$ is strictly positive and continuous, then it suffices that $f$ be continuous and have compact support in the interior of the state space. The first-order conditions are:

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

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

$$
\int \left[ \frac{A(f^* f)}{f^*} - \frac{f(A f^*)}{f^*} \right] dQ^* = 0. \tag{37}
$$

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

$$
A^* f = \frac{A(f^* f)}{f^*} - \frac{f(A f^*)}{f^*} = \frac{B(f^* f)}{f^*} - \frac{f(B f^*)}{f^*}, \tag{38}
$$

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

$$
\rho e = B e
$$

for $e$ in $D^+$. Construct a twisted generator using algorithm (38) with $f^* = e$, and suppose this generates a stochastically stable Markov process with stationary distribution $Q^*$. In particular, it satisfies (37). Notice that

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

51
because $e$ is in $\mathcal{D}^+$ and $\rho$ 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.

\section*{B Derivative of a principal eigenfunction}

In section 6 I showed how to compute the derivative of the principal eigenvalue. I now add to this discussion by providing the formula for the derivative of the principal eigenfunction. This derivative is useful in obtaining a more refined calculation. Recall that the principal eigenfunction is only defined up to scale. This leads me to study the derivative of logarithm of the principal eigenfunction evaluated at $a = 0$, denoted $D \log e$, which is well defined. I use the formula:

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

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

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

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

\section*{C Reconsidering the Campbell-Cochrane Model}

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

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

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

$$\text{risk aversion} = \gamma \exp(X_t + b).$$
The local risk exposure for $C^*_t$ is

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

Thus we require that

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


or

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

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

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

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 $x^*$:

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


or

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

Thus

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

which is restriction on the underlying parameters. Specifically,

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

Notice that we may now express $\lambda$ as:

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

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

$$= -\left(\frac{\gamma\xi - \theta}{\gamma^2|\sigma_c|^2}\right)^{1/2}[1 + 2(x - \mu_x)]^{1/2}.$$
as derived in Campbell and Cochrane (1999).

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

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

It implies an alternative distorted evolution:

\[
\begin{align*}
\tilde{W}_t &= \gamma \lambda (X_t)^2 \sigma_c dt + \tilde{W}_t \\
\end{align*}
\]

and \( \tilde{W}_t \) is a standard Brownian increment under the probability measure implied by \( \tilde{M} \).

Given the strong pull of the drift to the right for large \( X_t \), this evolution results in unstable stochastic dynamics.

## D Finite Horizons

Consider Example 3.4 and continued in Example 5.2. The additive functional is:

\[
dY_t = \nu dt + H_1 X_t^{[1]} dt + H_2 (X_t^{[2]} - 1) dt + \sqrt{X_t^{[2]}} F d\tilde{W}_t.
\]

Form

\[ M_t = \exp(Y_t) \]

My aim is to compute

\[ M_t 1(x) = E[M_t | X_0 = x] \]

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.

\[
\mathbb{B}M_t f = \frac{d}{dt} [M_t f(x)]
\]

Guess a solution

\[ M_t 1(x) = E[M_t | X_0 = x] = \exp[\alpha(t) \cdot x + \varphi(t)] \]
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

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

Moreover,

$$\frac{B \exp[\alpha(t) \cdot x + \varrho(t)]}{\exp[\alpha(t) \cdot x + \varrho(t)]} = \nu + H_1 x_1 + H_2 (x_2 - 1) + x_1' A_1' \alpha_1(t) + (x_2 - 1) A_2 \alpha_2(t) + \frac{x_2}{2} |F + \alpha_1(t)' B_1 + \alpha_2(t) B_2|^2.$$

First use (40) to produce a differential equation for $\alpha_1(t)$:

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

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

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

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

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

This is a differential equation for $\varrho(t)$ given the solution for $\alpha_2(t)$. The initial condition is $\varrho(0) = 0$. Example 5.2 has formulas for the limiting values of $\alpha_1$ and $\alpha_2$ as $t$ becomes large. The function $\varrho$ will eventually grow linearly.
References


