Modeling the Long Run:  
Valuation in Dynamic Stochastic Economies

Lars Peter Hansen*  
November 12, 2007

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.

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

*Prepared for the Fisher-Schultz Lecture at the 2006 European Meetings of the Econometric Society. I appreciate helpful comments by Jaroslav Borovicka 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.
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. (2005) 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.
2 Game plan

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

i) **Underlying Markov structure:** 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. I build stochastic discount factor processes that decline over time in a probabilistic way. The result will a multiplicative functional of the Markov process implied by an economic model designed to capture both pure discount effects and risk adjustment. The multiplicative construction reflects the effect of compounding over intervals of time.

ii) **Valuation with growth:** I study valuation in conjunction with growth by constructing families of operators indexed by the valuation horizon. The operators will map the transient components to payoffs, cash flows or Markov claims to a numeraire consumption good. As special cases I will study growth abstracting from valuation and valuation abstracting from growth.

iii) **Representation of operators with processes:** I use multiplicative functionals constructed from the underlying Markov process to represent the previously described family of operators. Transient components of this process produce transient model components.

iv) **Long-run approximation:** 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 we will these objects give us a convenient characterization of long-run behavior. They will also give us way to formally define permanent and transitory model components.

v) **Sensitivity and long-run pricing:** Of primary 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.

2.1 A revealing special case

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

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 it 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 \( n \) by \( n \) 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 \( t \) \( M_t \) 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 \( n \) by one vector \( f \), Perron-Frobenius theory characterizes limiting behavior 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 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 \( \rho \) is 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 motivate 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:

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

(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.

3 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.
3.1 Underlying Markov Process

I consider a Markov process $X$ defined on a state space $E$. 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:

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

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 $X$.

3.2 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 and hence can display arithmetic growth (or decay) even when the underlying process $X$ does not. An additive functional can be a 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 $(\beta, \gamma, \kappa)$ where:

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

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

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

$$A_t = \int_0^t \beta(X_u)du + \int_0^t \gamma(X_u) \cdot dW_u + \sum_{0 \leq u \leq t} \kappa(X_u, X_{u-})$$
While a multiplicative functional can be defined more generally, we will consider ones that are constructed as exponentials of additive functionals: $M = \exp(A)$. 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.

4 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:

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

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:

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

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

$$S_t S_\tau = S_{t+\tau}$$

for $t > 0, \tau > 0$ 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 (3) 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 (3).

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:

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

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

$$E_t E_\tau = E_{t+\tau}$$

1This latter implication gives the key ingredient of a more general definition of a multiplicative functional.
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

1) $M_0 = I$ and $M_t M_\tau = M_{t+\tau}$ 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 $\{M_t : t \geq 0\}$ constructing using (2) 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

$$C_t = G_t C_0 f(X_t)$$

displays stochastic growth when $G$ is a multiplicative functional. Since $G_0$ is normalized to be unity, $C_0$ allows us to consider other initial conditions.

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

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

The value assigned to $C_t$ is given by $C_0 P_t f(X_0)$ because $C_0$ is presumed to be in the date zero information set. Importantly, it is the product of two multiplicative functionals that we use representing the operator $P_t$, $M = GS$.

## 5 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). 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.

---

2Interestingly, Box and Tiao (1977) anticipate in the potentially important notion of long run co-movement in their method of extracting canonical components of multivariate time series.
5.1 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:

**Theorem 5.1.** Suppose that $X$ is a stationary Markov process and $Y$ is an additive functional for which:

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

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

Then $Y$ can be represented as:

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

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

**Proof.** Let $Y^*_t = Y_t - \nu t$. As a consequence of the Law of Iterated Expectations,

$$
g(X_t) + Y^*_t = E \left[ g(X_{t+\tau}) | X_t \right] + E \left[ Y^*_{t+\tau} | X_t \right]
$$

Form the additive functional:

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

Notice that

$$
E \left[ \hat{Y}_{t+\tau} - \hat{Y}_t | X_t \right] = E \left[ Y^*_{t+\tau} + g(X_{t+\tau}) | X_t = x \right] - Y^*_t - g(X_t) = 0.
$$

The holds for any $t \geq 0$ and any $\tau \geq 0$, implying that $\hat{Y}$ is a martingale.

The martingale component of the decomposition will have stationary increments. Decomposition (4) is familiar from the construction of martingale approximations used in central limit approximation (see Gordin (1969) for an important initial contribution.) The see the relation to central limit theory, recall the representation:

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

and suppose that the function $\kappa$ is normalized so that

$$
\int \kappa(y, x) \eta(dy | x) = 0.
$$
If this were not true, we could always adjust $\kappa$ and $\beta$ to make it true. In this case

$$\hat{Y}^1_t = \int_0^t \gamma(X_u) \cdot dW_u + \sum_{0 \leq u \leq t} \kappa(X_u, X_{u-})$$

is at the very least a local martingale. Suppose in fact that it is a martingale. The continuous time version of central limit theory for stationary Markov processes produces a martingale approximation of the form

$$\int_0^t \beta(X_u) du - tE\beta(X_u) = \hat{Y}_{t}^{[2]} - g(X_t) + g(X_0)$$

where $\hat{Y}_t^2$ is a martingale. For instance, see Bhattacharya (1982) and Hansen and Scheinkman (1995). This literature characterizes functions $\beta$ for which such a decomposition can be obtained. Adding the two martingales $\hat{Y}_t = \hat{Y}^{[1]}_t + \hat{Y}^{[2]}_t$ gives the martingale in (4).

The conclusion that this decomposition is unique may seem surprising. The result is obtained in part by fixing the information structure associated with the underlying Markov process. This information structure is used in building martingales. A well known nonuniqueness emerges if we allow for alternative information structures while preserving the induced distribution of the additive functional. Another type of nonuniqueness emerges if we attempt to identify a unique set of underlying shocks that generate the conditioning information. This identification may be central to interpreting results, and require additional restrictions as is well known in the VAR literature, but it is not part of our result.

Since a conditional expectation operator is a linear operator, the decomposition may used directly to show that,

$$E(Y_t|X_0) = \nu t - E[g(X_t)|X_0] + g(X_0)$$

In particular,

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

$\rho$ is long run arithmetic growth rate. A more refined analysis leads to

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

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.

\(^3\)See Billingsley (1961) for the discrete time martingale central limit. Moreover, there are well known functional extensions of this result.
Example 5.2. Suppose that
\[ dX_t = AX_t dt + BdW_t, \]
\[ dY_t = \nu + HX_t dt + GdW_t, \]
where \( A \) has eigenvalues with strictly negative real parts and \( W \) is multivariate Brownian standard motion. In this example,
\[ \hat{Y}_t = \int_0^t (G - HA^{-1}B) dW_u \]
is the martingale component. The transient component is obtained by computing:
\[ g(X_t) = \int_0^\infty E(HX_{t+u}|X_t) du = -HA^{-1}X_t. \]

Example 5.3. Suppose that \( X \) and \( Y \) evolve according to:
\[ dX_t^{[1]} = A_1X_t^{[1]} dt + \sqrt{X_t^{[2]}}B_1dW_t, \]
\[ dX_t^{[2]} = A_2(X_t^{[2]} - \kappa) + \sqrt{X_t^{[2]}}B_2dW_t, \]
\[ dY_t = \nu + H_1X_t^{[1]} dt + H_2X_t^{[2]} + \sqrt{X_t^{[2]}}GdW_t. \]
Both \( X^{[2]} \) and \( Y \) are scalar processes. 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 \( A_2 + \frac{1}{2}|B_2|^2 < \infty \). In this example the martingale component is given by:
\[ \hat{Y}_t = \int_0^t \sqrt{X_u^{[2]}} \left[ G - H_1(A_1)^{-1}B_1 - H_2(A_2)^{-1}B_2 \right] dW_u. \]
The transient component is computed by solving:
\[ g(X_t) = \int_0^\infty E(HX_{t+u}^{[1]}|X_t) du = H(A_1)^{-1}X_t^{[1]}. \]
Then
\[ Y_t = \nu t + \hat{Y}_t - g(X_t) + g(X_0) \]
as desired.

This example has the same structure as example 5.2 except that the Brownian motion shocks are scaled by \( \sqrt{X_t^{[2]}} \) to induce volatility that varies over time. The process \( X^{[2]} \) is the a “Feller square root” process. While example 5.2 is fully linear, example 5.3 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, \( g \) used to measure the transient component, will not be a linear function of the Markov state.\(^4\)

\(^4\)The Markov assumption is also not necessary for such a decomposition.
Even when such nonlinearity is introduced, conveniently the sum of two additive functionals is additive and the martingale decompositions add as well provided they are constructed using a common information structure.

In what follows we will use multiplicative functionals, processes whose logarithms can be represent 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 4. 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 the assumption can be made to transform $\exp(\hat{Y})$ into a martingale. 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.

6 Limiting characterizations of stochastic growth

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.

6.1 Some interesting limiting behavior

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

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

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

$$\lim_{t \to \infty} \exp [-t\rho(M)] E[M_t f(X_t) | X_0 = x] \propto e(x) \quad (5)$$

12
which justifies calling \( \rho(M) \) a growth rate. Moreover, while the coefficient in this limit will depend on the choice of \( f \), their is a limiting form of state dependence captured by the function \( e \), which we will determine independently of \( f \). We will say more about the admissible functions \( f \) in our subsequent analysis and about the coefficient in this approximation. 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.

I will show how to represent the proportionality coefficient in the limit of (5) as as a linear functional of \( f \). The function \( e \) is is positive, and it solves:

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

for all \( t \geq 0 \). Differentiate this equation with respect to \( t \) and evaluating the derivative at \( t = 0 \) gives an equation for \( \rho(M) \):

\[
\lim_{t \downarrow 0} \frac{E[M_t e(X_t)|X_0 = x] - e(x)}{t} = \rho(M)e(x)
\]

This is an eigenvalue equation. The left-hand side of this expression can be viewed as an operator on a function space, and the function \( e \) is an eigenfunction of this operator. While there may be multiple eigenvalues, we will be interested in cases in which there is a unique strictly positive eigenfunction and associated eigenvalue. The resulting eigenvalue \( \rho(M) \) is referred to as the principal eigenvalue and the associated eigenfunction is the principal eigenfunction.

### 6.2 Products and covariation

While the product of two multiplicative functionals is multiplicative, it is not true that

\[
\rho(M_1 M_2) = \rho(M_1) + \rho(M_2).
\]

Covariation is important when characterizing even the limiting behavior of the product \( M_1 M_2 \). In fact the discrepancy:

\[
\rho(M_1 M_2) - \rho(M_1) - \rho(M_2).
\]

will of interest in some of our analysis. If \( M_1^t \) and \( M_2^t \) happen to be jointly log normal for each \( t \), then (6) is equal to the limiting covariance between the corresponding logarithms:

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

where \( M_j^t = \exp(Y_j^t) \) for \( j = 1, 2 \). While this illustrates that covariation plays a central role in \( \rho(M_1 M_2) \), we will not require log-normality in what follows.

### 6.3 Local versus global

Consider for the moment a special class of multiplicative functionals:

\[
M_t = \exp \left[ \int_0^t \nu(X_u)du \right].
\]
Such functionals are special because they are smooth, or locally riskless. The multiplicative functional has a state dependent growth rate given by $\nu(x)$. If $\nu(x)$ were constant (state independent), then the long-run growth rate $\rho(M)$ and the local growth rate would coincide. When $\nu$ 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 $\frac{1}{t} \int_0^t \nu(X_u) du$ 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 $\mathcal{E}$ of the Markov process $X$. Following Donsker and Varadhan (1976) and Dupuis and Ellis (1997) and others, a rate function $\mathbb{I}(Q)$ is constructed to measure the discrepancy between the original stationary distribution and $Q$. Equivalently, $\mathbb{I}$ is a measure of relative entropy between two probability distributions as justified in appendix A. The function $\mathbb{I}$ is convex in the probability measure $Q$, and it is used to construct what is called a Laplace principle that characterizes the limit:

$$ \rho(M) = \sup_Q \int \nu(x) dQ - \mathbb{I}(Q) \geq E[\nu(X_t)] $$

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

This optimization problem is inherently static, with the dynamics loaded into the construction of convex function $\mathbb{I}$ as we illustrate in appendix A. This construction is independent of the choice of $\nu$. Recall that $\nu$ is 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 (7) characterizes formally this difference.\(^5\)

The local growth rate can be defined more generally via:

$$ \nu(x) = \lim_{t \downarrow 0} \frac{E(M_t | X_0 = x) - 1}{t} $$

provided that this limit exists. The multiplicative functional can be decomposed into two component multiplicative functionals:

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

\(^5\)Large deviation theory exploits problem 7 because $\rho(M)$ implies a bound of the form:

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

for large $t$. This bound is only of interest when $k > \rho(M)$. Our interest in $\rho(M)$ is different, but the probabilistic bound is also intriguing.
where $M_t^*$ is a local martingale. Both components are multiplicative functionals. When this local martingale is a martingale, as we argued previously it is associated with a distorted probability distribution for $X$. This probability distribution preserves the Markov structure. 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. The inequality:

$$\rho(M) \geq E[\nu(X_t)]$$

can be shown to hold more generally as an implication of Jensen’s Inequality since the exponential function is convex.

## 7 Multiplicative Decomposition

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

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

(9)

Component a) governs the long-term growth or decay. It is constructed from a dominant eigenvalue and eigenfunction. I will use component b), the positive martingale, to build an alternative probability measure. This alternative measures gives us a tractable framework for a formal study of approximation and transitivity. Component c) is built directly from dominant eigenfunction.

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 components is transient when the underlying Markov process is stationary. While the stochastic inputs of the martingale $\tilde{M}$ will be long lasting the same is not true of the transient component. Although positive, this martingale will typically not converge. 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

---

6In 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:

$$M_t = M_t^d M_t^d$$

where $\{M_t^d : t \geq 0\}$ 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 continuous.

7The 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).
behavior of the process. Thus I am lead to ask: Is this decomposition unique? When is this decomposition useful? The answers to these questions are intertwined.

7.1 Multiplicative Decomposition

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

The simplest case is long normal example just considered in Example 5.2.

Example 7.1. Consider again example 5.2 and construct an additive process:

\[ dY_t = \nu dt + Hx_t dt + GdW_t. \]

Form

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

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

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

is a martingale. The growth rate for \( M \) is:

\[ \rho(M) = \nu + \frac{|G - HA^{-1}B|^2}{2} \]

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

For the more general Markov diffusion case, we could still exploit the local normality and make a state-dependent volatility adjustment, but this is typically not the martingale that is of interest. There is another approach at our disposal. Suppose that we can solve an eigenvalue problem:

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

for a positive function of \( e \). Then

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

is a positive martingale. Inverting this relation:

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

Thus \( \hat{e}(X_t) = \frac{1}{e(X_t)} \) is a temporary component of \( M \) and \( \rho \) is the growth rate, or are they? Unfortunately, this eigenvalue problem, as posed, can have multiple solutions with multiple candidates for growth rates and eigenfunctions. At most one such decomposition is of interest. To defend this representation, we consider formally an approximation problem.
Example 7.2. Consider again Example 5.3, and construct an additive functional:

\[ dY_t = \nu + H_1 X_t^{[1]} dt + H_2 X_t^{[2]} + \sqrt{X_t^{[2]}} G dW_t. \]

Form

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

Guess a solution \( e(x) = \exp(\alpha \cdot x) \) where \( x = [x_1 \quad x_2] \) and \( \alpha = [\alpha_1 \quad \alpha_2] \). To compute \( \rho(M) \), solve

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

Thus the coefficients on \( x_1 \) and \( x_2 \) are zero:

\[
\begin{align*}
A_1' \alpha_1 + H_1' &= 0 \\
A_2 \alpha_2 + H_2 + \frac{1}{2} |B_1' \alpha_1 + B_2 \alpha_2 + G'|^2 &= 0.
\end{align*}
\]

The first equation can be solved for \( \alpha_1 \) and the second one for \( \alpha_2 \) given \( \alpha_1 \). The second equation is quadratic in \( \alpha_2 \), so there may be two solutions. We will have cause to select one of these solutions as the interesting one. Finally,

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

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

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

For each time horizon \( t \), we define an alternative conditional expectation operator. The martingale property is needed so the 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 7.3. The process \( X \) is stochastically stable under the measure \( \hat{\cdot} \) if

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

for any \( f \) for which \( \hat{E}(f) \) is well defined and finite.\(^8\)

---

\(^8\)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.
Consider now the construction,

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

It follows that

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

for large \( t \) provided that \( X \) is stochastically stable and \( \hat{E} \left[ f(X_t) \hat{e}(X_t) \right] \) where is finite and \( \hat{e} = 1/e \).

Notice that this stability condition is presumed to hold under the distorted probability distribution. Establishing this property allows us to ensure that the dependence between the martingale and transient components is limited sufficiently so that the we may think of \( \rho \) as the exponential growth rate. In other words, this is necessary for

\[
\rho = \rho(M)
\]

defined previously.

Once we scale by this growth rate, 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 \( e(x) \). 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 density 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:

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

**Definition 7.4.** 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 dominant 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 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 alterantive approaches based on mean-square approximation. While there may be multiple representations of the form (9), Hansen and Scheinkman (2007) show that there is at most one such representation for which the process \( X \) is stochastically stable. For instance, in example 7.2 at most one of the solutions will result in a stochastically stable process for \( X^{[2]} \). At most one of the solutions of the quadratic equation for \( \alpha_2 \) will imply this stability.

18
8 Transient model components

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

Consider a benchmark valuation model represented by a stochastic discount factor or a benchmark growth process or the product of the two components. I now ask what modifications are transient? The tools in describe in section 7 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 stability condition 7.3. 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 processes is altered because $\{f(X_t)\}$ will be transitory in this alternative representation when:

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

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

9 Perturbation calculation

In applications the multiplicative functionals used in constructing the semigroup depend on model parameters. Thus I consider $M(\alpha)$ 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. (2005), or it could a parameter that governs the exposure to a source of long run 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 $M(\alpha)$ by changing $\alpha$ is equivalent to perturbing the operators associated with this process. My choice of scalar parameterization is made for notational convenience.

In the Hansen et al. (2005) application, $\alpha + 1$ is a common intertemporal substitution parameter across investors. The aim is to study how long-run risk premia change with $\alpha$.

Thus the stochastic discount factor process is depicted as $S(\alpha)$ and $M(\alpha) = S(\alpha)G$ where $G$ is the stochastic growth component of a hypothetical or real cash flow. As an alternative, $\alpha$ could parameterize the long run risk exposure of a hypothetical cash flow. In this case $M(\alpha) = SG(\alpha)$ and $\alpha$ could parameterize the risk exposure of the cash flow. While long-risk prices will often not be linear in $\alpha$, the derivative would characterize their behavior. By exploring alternative directions of risk, we infer which directions are of most concern to investors as reflected by an underlying economic model.

9.1 Finite horizon

Consider first the finite horizon calculation. Let $\{M_t(\alpha) : t \geq 0\}$ be a parameterized family of multiplicative martingales. There is an associated parameterized family of valuation functionals:

$$M_t(\alpha)f(x) = E[M_t(\alpha)f(x_t)|x_0 = x].$$

Then

$$\frac{d}{d\alpha} M_t(\alpha)f(x)|_{\alpha=0} = E\left[\frac{d}{d\alpha} M_t(\alpha) \log M_t(\alpha)f(x_t)|x_0 = x\right]_{\alpha=0}.$$

9.2 Limiting behavior

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

$$\frac{d}{d\alpha} \rho(\alpha). \quad (11)$$

Calculation (11) turns out to be straightforward as we will now see. First solve the principal eigenvalue problem for $\alpha = 1$ and use the solution to construct a probability measure $\hat{\cdot}$ as we described previously. The formula for the derivative is:

$$\frac{d}{d\alpha} \rho(\alpha)|_{\alpha=0} = \frac{1}{t} \hat{E} \left[ \frac{\partial \log M_t(\alpha)}{\partial \alpha} \right]_{\alpha=0} \quad (12)$$

which can be evaluated for any choice of $t$. The functional $\log M_t(\alpha)$ is an additive functional, and its derivative is as well. For the continuous time models we described, it is most convenient to take limits as $t \to 0$. This entails computing an average local mean under the distorted distribution:

$$\frac{d}{d\alpha} \rho(\alpha)|_{\alpha=0} = \hat{E} \left( \frac{d}{d\alpha} \left[ \beta(X_t;\alpha) + \gamma(X_t;\alpha) \cdot \gamma(x) \right]_{\alpha=0} \right)$$
\[ + \hat{E} \left[ \int \frac{d}{d\alpha} \kappa(y|X_t;\alpha) \bigg|_{\alpha=0} \exp[\hat{\kappa}(y,X_t)]\eta(dy|X_t) \right] \]

where we have used the fact that the Brownian motion has \( \hat{\gamma}(X_t)dt \) as the drift under the \( \hat{\cdot} \) distribution and uses the conditional measure \( \exp[\hat{\kappa}(y,X_t)]\eta(dy|X_t) \) 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 our decomposition:

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

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

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

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

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

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

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

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

Take expectations and use the fact that \( X \) is stationary under the \( \hat{\cdot} \) probability measure to obtain derivative formula (12).

Consider next the derivative of the principal eigenfunction with respect to \( \alpha \). The principal eigenfunction is only defined up to scale. This leads us to study the derivative of logarithm of the principal eigenfunction, denoted \( D \log e \), which is well defined. In the appendix we derive the following implicit formula for the derivative of the eigenfunction:

\[ \hat{B} \left( D \log e \right)(x) = \frac{d}{d\alpha} \rho \left[ M(\alpha) \right] \bigg|_{\alpha=0} - \frac{1}{e(x;0)} \left[ \frac{d}{d\alpha} \hat{B}(\alpha) \bigg|_{\alpha=0} \right] e(x;0). \]

10 Applications to Asset Pricing

10.1 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:

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

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.

### 10.2 Power utility without predictability

Suppose that consumption is a geometric Brownian motion:

\[ dc_t = \mu_c dt + \sigma_c dW_t, \]

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:

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

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:

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

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:

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

In particular, it follows that

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

The long-term risk prices can be computed by differentiating \(-\nu(SG)\) with respect to the risk exposure vector \( \sigma_g \), and are thus equal to: \( \gamma \sigma_c \). The dynamics of pricing for this example are degenerate, and in particular the local risk price vector is also equal to \( \gamma \sigma_c \). The local and long-term prices are the same. The instantaneous risk-free interest rate is the constant and identical to the long-term counterpart: \(-\nu(S) = \delta + \gamma \mu_c - \frac{\gamma^2}{2} |\sigma_c|^2\).

The multiplicative martingale \( \hat{M} \) implies 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.
10.3 An example of Campbell and Cochrane

Campbell and Cochrane (1999) modify standard asset pricing model with power utility by introducing a stochastic subsistence point $B$ that shares the same stochastic growth properties as consumption. In language of time series, this process is cointegrated with consumption. The process is necessarily nonnegative because it must always be below consumption. The process $B$ 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:

$$S^*_t = \exp(-\delta t) \left[ \frac{(C_t - B_t)^{-\gamma}}{(C_0 - B_0)^{-\gamma}} \right]$$

We may rewrite this as:

$$S^*_t = S_t \left[ \frac{(1 - B_t/C_t)^{-\gamma}}{(1 - B_0/C_0)^{-\gamma}} \right].$$

In what follows let

$$X_t = -\log(1 - B_t/C_t),$$

which is process that should exceed zero. Using this notation, write:

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

Following Campbell and Cochrane (1999), assume that

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

where we restrict $\lambda(0) = 0$ in hopes that the zero boundary will not be attainable. Squashing the variability at zero makes it less likely to get close to this boundary. After the probability distortion, the law of motion for this equation is:

$$dX_t = -\xi(X_t - \mu_x)dt + \frac{(-\gamma \sigma_c + \sigma_g) \cdot \sigma_c}{|\sigma_c|} \lambda(X_t)dt + \frac{\lambda(X_t)}{|\sigma_c|} \sigma_c d\tilde{W}_t.$$  \hspace{1cm} (13)

10.3.1 Interest rates

The function $\lambda$ is specified to guarantee that interest rates are constant. This is done for simplicity. The interest rate for the power utility $S$ model is $r = \delta + \gamma \mu_c + \frac{\gamma^2}{2} |\sigma_c|^2$. Set $\gamma_g = 0$ and hence $G_t = 1$. Let $r^*$ be the constant interest rate in the Campbell-Cochrane model. Then

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

The local counterpart to this equation is:

$$-r - \gamma \xi(x - \mu_x) + \gamma^2 \frac{\lambda(x)^2}{2} - \gamma^2 |\sigma_c|^2 \lambda(x) = -r^*.$$
The interest rate $r^*$ can be inferred by imposing $\lambda(x) = 0$:

$$r^* = r - \gamma \xi \mu_x.$$ 

Since the instantaneous interest rate is constant, the real term structure is flat and the limiting rate of return on bonds is the same as the instantaneous rate of return. The modification in preferences changes this term structure, but this change can be offset altering $\delta$.

The terms in $x$ should be zero, implying the quadratic equation:

$$-\gamma \xi - \gamma^2 |\sigma_c| \lambda(x) + \frac{\gamma^2 \lambda(x)^2}{2} = 0.$$ 

Solving for $\lambda$ gives:

$$\lambda(x) = \frac{\gamma |\sigma_c| \pm \sqrt{\gamma^2 |\sigma_c|^2 + 2 \xi x \gamma}}{\gamma}.$$  

Imposing $\lambda(0) = 0$ leads me to take the "−" branch and thus

$$\lambda(x) = |\sigma_c| - \sqrt{|\sigma_c|^2 + 2 \xi x / \gamma}.$$  

Since the instantaneous interest rate is constant, the real term structure is flat and the limiting rate of return on bonds is the same as the instantaneous rate of return. The modification in preferences changes this term structure, but this change can be offset altering $\delta$.

### 10.3.2 Local risk prices

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

$$\gamma \sigma_c - \frac{\gamma \lambda(x)}{|\sigma_c|} \sigma_c = \gamma \frac{\sqrt{|\sigma_c|^2 + 2 \xi x / \gamma}}{|\sigma_c|} \sigma_c$$ 

which by design are state dependent and are larger than in the power utility model. Moreover, the state variable $X_t$ responds negatively to consumption growth shocks because $\lambda(x) < 0$. Risk premia are larger in bad times as reflected by unexpectedly low realizations of consumption growth.

### 10.3.3 Long-term pricing

Consider now the long-run behavior of value. Here we need to investigate behavior as $x$ becomes large. Here we use evolution equation (14) and the formula for the logarithmic derivative of the density for a scalar diffusion:

$$\frac{d \log q}{dx} = 2 \frac{\text{drift}}{\text{diffusion}} + \frac{d \log \text{diffusion}}{dx}$$  

The limiting behavior is dominated by the constant term:

$$\lim_{x \to \infty} \frac{d \log q}{dx} = -\gamma.$$
Thus log $q$ behaves like a positive scalar multiple of $-\gamma x$ for large $x$. Thus to study expectations of $\exp(\gamma X_t)$ under the distorted law of motion requires a more refined calculation.

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

We would like this term be negative because it is the coefficient on $x^{-1/2}$ in the large $x$ approximation of the derivative of the log density. This term, however, will often be positive preventing $\exp(\gamma X_t)$ from having a finite moment under the twisted stationary distribution associated with (14). For instance, when $\sigma_g = 0$ it is positive as we expect because the adjustment proposed by Campbell and Cochrane (1999) changes interest rates including in the limit has the payoff horizon is extended to $\infty$. When $\sigma_g = \sigma_c$, this term will be positive provided that $\gamma > 1$. In this sense the modification proposed by Campbell and Cochrane (1999) is not transient even though on the surface it looks like it might be. There will be a class of payoffs with limiting values that agree with those of the power utility model, but this class is small (excluding for instance constant payoffs) unless

$$(-\gamma \sigma_c + \sigma_g) \cdot \sigma_c > 0.$$

10.4 Santos and Veronesi

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

$$S_t^* = S_t \frac{X_t}{X_0}.$$

The evolution for $X_t$ is

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

where

$$\lambda(X_t) = -\kappa(X_t - 1).$$

The specification of local volatility is designed to keep $\{X_t : t \geq 0\}$ above unity. As in the Campbell-Cochrane specification, the process $X_t$ responds negatively to a consumption shock.

The local risk prices are now given by

$$\gamma \sigma_c + \frac{\kappa(x - 1)}{x|\sigma_c|} \sigma_c.$$

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 (14) but with this new specification of $\kappa$. Formula (15) is again informative. In particular,

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

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.
A Relative Entropy for Stationary Processes

Following Donsker and Varadhan (1976), let $A$ be the generator for the continuous time Markov process $X$ with domain $\mathcal{D}$. Let $\mathcal{D}_+$ denote the strictly positive functions in $\mathcal{D}$. For any probability measure $Q$ on the state space $\mathcal{E}$ of the Markov process, let

$$I(Q) = -\inf_{f \in \mathcal{D}_+} \int \left( \frac{Af}{f} \right) dQ$$

which may be infinite. This function is convex in $Q$ since it can be expressed as the maximum of convex (in fact linear) functions of $Q$.

To understand this calculation, suppose that the infimum is attained at $e$. Then

$$\frac{A(ef)}{e} - f \frac{Ae}{e}$$

is generator of an alternative twisted Markov process. The first-order conditions for $\log e$ in the direction $f$ imply that:

$$\int \left[ \frac{A(ef)}{e} - f \frac{Ae}{e} \right] dQ = 0. \quad (16)$$

These first-order conditions imply that $Q$ is a stationary distribution for this twisted Markov process. The instantaneous conditional relative entropy of this twisted Markov process is:

$$\frac{A(e \log e)}{e} - \log e \frac{A(e)}{e} - \frac{A(e)}{e}.$$  

(See Dupuis and Ellis (1997) chapter 4.6.)

Integrating this measure with respect to the stationary distribution $Q$ and using relation (16) for $f = \log e$, we obtain:

$$-\int \frac{A(e)}{e} dQ$$

as a measure of average conditional relative entropy between the probability measure induced by the original Markov process and the measure induced by the twisted Markov process associated with $e$. This second measure has $Q$ as its stationary distribution.

To show the relation between optimization problem (7) and principal eigenvalue problem, consider any signed measure $R$ with total measure zero $R(\mathcal{E}) = 0$ such that $Q + rR$ is a probability measure at least for small $r$. Then the derivative of $I$ in direction $R$ is

$$-\int \frac{A(e)}{e} dR = 0$$

since $e$ is chosen optimally. Similarly, the directional derivative for the objective: $\int \nu dQ - I(Q)$ is:

$$\int \left[ \nu + \frac{A(e)}{e} \right] dR.$$  

This directional derivative should be zero at an optimum which will be satisfied for any admissible $R$ provided that

$$\nu + \frac{A(e)}{e} = 0.$$
for some real number \( \varrho \). Multiplying both sides by \( e \) gives the principal eigenvalue problem for the operator \( \mathbb{B}e = \mathbb{A}e + \nu e \) where \( \varrho \) is the eigenvalue and \( e \) is the eigenfunction.

The link between this optimization problem and the eigenvalue problem is well known in the literature on large deviations, for instance see Donsker and Varadhan (1976), Balaji and Meyn (2000) and Kontoyiannis and Meyn (2003).
References


Run Risk. University of Chicago.

Journal of Mathematical Economics 20:381–408.

Ito, K. and S. Watanabe. 1965. Transformation of Markov Processes by Multiplicative


Springer-Verlag.