

# Multiplicative Functionals

## Contents

- 8.1. Geometric growth and decay
- 8.2. Special multiplicative functionals
- 8.3. Multiplicative martingales
- 8.4. Factoring a multiplicative functional
- 8.5. Stochastic stability
- 8.6. Inferences about permanent shocks from asset prices
- 8.7. Digression on rational expectations
- 8.8. Values of stochastic cash flows
- 8.9. Bounding investor beliefs

[!\[\]\(666e09182d4cd268646ea700ea60dcdf\_img.jpg\) Download PDF here](#)

**Authors:** Jaroslav Borovicka (NYU), Lars Peter Hansen (University of Chicago), and Thomas Sargent (NYU)



Challenges in navigating long-term uncertainty.

[Chapter 4: Processes with Markovian Increments](#) described **additive functionals** of a Markov process. This chapter describes exponentials of additive functionals that we call **multiplicative functionals**. We can use them to model stochastic growth, stochastic discounting, belief distortions and their interactions. After adjusting for geometric growth or decay, a multiplicative functional contains a martingale component that turns out to be a likelihood ratio process that is itself a special type of multiplicative functional called an **exponential martingale**. By simply multiplying random variables of interest by the multiplicative martingale prior to computing conditional expectations under a baseline probability, we compute conditional expectations under an alternative probability measure. The exponential martingale thus functions as a relative density that converts the baseline probability measure into the alternative probability measure. For an early application of multiplicative functionals to asset valuation under preferences that express investors' concerns about model misspecification, see [[Anderson et al., 2003](#)]. We will explore these and other applications in discussions that follow. We will see that multiplicative functionals have several applications. We shall use them study consequences of stochastic growth and stochastic discounting that persist over long time horizons and how macroeconomic shocks can have long-term impacts on asset prices. They provide a tractable modeling tool for studying models in which investors have subjective beliefs that can be described in terms of discrepancies from a baseline probability in ways that can be measured with statistical discrimination measures. We will also describe other applications of multiplicative functionals, including models of returns and positive cash flows that compound over multiple horizons, cumulative stochastic discount factors are used to long-horizon, and risk-return tradeoffs.

## 8.1. Geometric growth and decay

To construct a multiplicative functional, we start with an underlying Markov process  $X$  that has a stationary distribution  $Q$ , and we suppose that  $Y_0$  and  $X_0$  depend only on date zero information summarized by  $\mathfrak{A}_0$ .

### Definition 8.1

Let  $Y \stackrel{\text{def}}{=} \{Y_t\}$  be an additive functional that as in [Chapter 4](#) is described by

$$Y_{t+1} - Y_t = \kappa(X_t, W_{t+1}),$$

where  $X_t$  is the time  $t$  component of a Markov state vector satisfying  $X_{t+1} = \phi(X_t, W_{t+1})$  and  $W_{t+1}$  is the time  $t + 1$  value of a martingale difference process ( $\mathbb{E}(W_{t+1} | \mathfrak{A}_t) = 0$ ) of unanticipated shocks. We say that  $M \stackrel{\text{def}}{=} \{M_t : t \geq 0\} = \{\exp(Y_t) : t \geq 0\}$  is a **multiplicative functional** parameterized by  $\kappa$ .

An additive functional grows or decays *linearly*, so the exponential of an additive functional grows or decays *geometrically*. We construct a multiplicative functional recursively by

$$M_{t+1} = N_{t+1}M_t. \tag{8.1}$$

If we adopt the normalization that  $N_t = 1$  if  $M_t = 0$ , we can solve for  $N_{t+1}$

$$N_{t+1} = \begin{cases} \frac{M_{t+1}}{M_t}, & M_t > 0 \\ 1, & M_t = 0 \end{cases}$$

In light of [\(8.1\)](#), we refer to  $N$  as the **multiplicative increment** of the process  $M$ .

[Chapter 4](#) stated a Law of Large Numbers and a Central Limit Theorem for additive functionals. In this chapter, we use other mathematical tools to analyze the limiting behavior of multiplicative functionals.

## 8.2. Special multiplicative functionals

We define the three primitive multiplicative functionals.

### Example 8.1

Suppose that  $\kappa = \eta$  is constant and that  $M_0$  is a Borel measurable function of  $X_0$ . Then

$$M_t = \exp(t\eta)M_0.$$

This process grows or decays geometrically.

### Example 8.2

Suppose that

$$\mathbb{E}[\exp[\kappa(X_t, W_{t+1})] | \mathfrak{A}_t] = 1.$$

Then

$$\mathbb{E}(M_{t+1} | \mathfrak{A}_t) = M_t \quad (8.2)$$

so that

$$\mathbb{E}(N_{t+1} | \mathfrak{A}_t) = 1$$

A multiplicative functional that satisfies [\(8.2\)](#) is called a *multiplicative martingale*. We can view this process as a likelihood ratio process, so we'll label it as  $M \stackrel{\text{def}}{=} L$ . It will often be convenient to initialize this likelihood ratio process at  $M_0 = L_0 = 1$ .

### Example 8.3

Suppose that  $M_t = \exp[f(X_t)]$  where  $h$  is a Borel measurable function. The associated additive functional satisfies

$$\begin{aligned} Y_{t+1} - Y_t &= \log M_{t+1} - \log M_t \\ &= f(X_{t+1}) - f(X_t) \\ &= h[\phi(X_t, W_{t+1})] - f(X_t) \end{aligned}$$

and is parameterized by  $\kappa(X_t, W_{t+1}) = h[\phi(X_t, W_{t+1})] - f(X_t)$  with initial condition  $Y_0 = f(X_0)$ .

When the process  $\{X_t\}$  is stationary and ergodic, multiplicative functional [Example 8.1](#) displays expected growth or decay, while multiplicative functionals [Example 8.2](#) and [Example 8.3](#) do not. Multiplicative functional [Example 8.3](#) is stationary, while [Example 8.1](#) and [Example 8.2](#) are not.

By multiplying together instances of these primitive multiplicative functionals, we can construct other multiplicative functionals. We can also on reverse that process by taking an arbitrary multiplicative functional and (multiplicatively) decomposing it into instances of our three types of multiplicative functionals. Before doing that, we explore multiplicative martingales in more depth.

## 8.3. Multiplicative martingales

We can use multiplicative martingales to represent alternative probability models. We can characterize an alternative model with a set of implied conditional expectations of all bounded random variables,  $B_{t+1}$ , that are measurable with respect to  $\mathcal{A}_{t+1}$ . The constructed conditional expectation is

$$\mathbb{E}(N_{t+1}B_{t+1} \mid \mathcal{A}_t). \quad (8.3)$$

We want multiplication of  $B_{t+1}$  by  $N_{t+1}$  to change the baseline probability to an alternative probability model. To accomplish this, the random variable  $N_{t+1}$  must satisfy:

1.  $N_{t+1} \geq 0$ ;
2.  $\mathbb{E}(N_{t+1} \mid \mathcal{A}_t) = 1$ ;
3.  $N_{t+1}$  is  $\mathcal{A}_{t+1}$  measurable.

Property 1 is satisfied because conditional expectations map positive random variables  $B_{t+1}$  into positive random variables that are  $\mathcal{A}_t$  measurable. Properties 2 and 3 are satisfied because  $N$  is the multiplicative increment of a multiplicative martingale. Where

$$L_t = \prod_{j=1}^t N_j, \quad (8.4)$$

$L = \{L_t\}_{t=0}^{\infty}$  can be viewed as a likelihood ratio or Radon-Nikodym derivative process for the alternative probability measure relative to the baseline measure.

Representing an alternative probability model in this way is restrictive. For instance, if a nonnegative random variable has conditional expectation zero under the baseline probability, it will also have zero conditional expectation under the alternative probability measure, an indication that the implied probability measure is *absolute continuous* with respect to the baseline measure. Two models that violate absolute continuity can be distinguished with probability one from only finite samples. Likelihood-based statistical discrimination tests typically exclude such “easy to distinguish” alternatives by assuming mutually absolutely continuous alternative probability models.

Representation (8.4) implies that the implied alternative conditional probability measure over  $\tau$  periods can be represented in terms of the likelihood ratio

$$\frac{L_{t+\tau}}{L_t} = \prod_{j=1}^{\tau} N_{t+j}, \quad (8.5)$$

which we can use to compute a  $\tau$  time-period-ahead conditional expectation. By using (8.5), we can compute a  $\tau$ -period conditional expectation by iterating on the one-period conditional expectations in accordance with the Law of Iterated expectations.

Multiplicative martingales provide a way to represent various models of subjective beliefs of private or government agents within dynamic, stochastic equilibrium models in terms of a statistical departure from a baseline model that the model builder uses to describe the data. [Hansen and Scheinkman, 2009] show that they also offer a way to value cumulative returns. Let  $R_t$  be a multiplicative process that measures a cumulative return between date  $t$  and date zero. Let  $S_t$  be a corresponding equilibrium discount factor between these same two dates. That  $L = RS$  is a multiplicative martingale follows from equilibrium restrictions on one-period returns. That is,

$$\mathbb{E} \left[ \left( \frac{S_{t+1}}{S_t} \right) \left( \frac{R_{t+1}}{R_t} \right) \mid \mathfrak{A}_t \right] = 1,$$

where  $S_{t+1}/S_t$  is the one-period stochastic discount factor and  $R_{t+1}/R_t$  is the one-period gross return. For the application, we construct

$$N_{t+1} = \frac{S_{t+1}R_{t+1}}{S_tR_t}$$

Here are some examples of multiplicative martingales constructed from some standard probability models.

#### Example 8.4

Consider a baseline Markov process having transition probability density  $\pi_o$  with respect to a measure  $\lambda$  over the state space  $\mathcal{X}$

$$P_o(dx^+ | x) = \pi_o(x^+ | x)\lambda(dx^+)$$

Let  $\pi$  denote some other transition density that we represent as

$$\pi(x^+ | x)\lambda(dx^+) = \left[ \frac{\pi(x^+ | x)}{\pi_o(x^+ | x)} \right] \pi_o(x^+ | x)\lambda(dx^+)$$

where we assume that  $\pi_o(x^+ | x) = 0$  implies that  $\pi(x^+ | x) = 0$  for all  $x^+$  and  $x$  in  $\mathcal{X}$ . Construct the multiplicative increment process as:

$$N_{t+1} = \frac{\pi(X_{t+1} | X_t)}{\pi_o(X_{t+1} | X_t)}.$$

### Example 8.5

Let an alternative model for a vector  $X$  be a vector autoregression:

$$X_{t+1} = \mathbb{A}X_t + \mathbb{B}W_{t+1}$$

where  $\mathbb{A}$  is a stable matrix,  $\{W_{t+1} : t \geq 0\}$  is an i.i.d. sequence of  $\mathcal{N}(0, I)$  random vectors conditioned on  $X_0$ , and  $\mathbb{B}$  is a square, nonsingular matrix. Assume that a baseline model for  $X$  has the same functional form but different settings  $(\mathbb{A}_o, \mathbb{B}_o)$  of its parameters. Construct  $N_{t+1}$  as the one-period conditional log-likelihood ratio

$$\begin{aligned} \log N_{t+1} = & -\frac{1}{2}(X_{t+1} - \mathbb{A}X_t)'(\mathbb{B}\mathbb{B}')^{-1}(X_{t+1} - \mathbb{A}X_t) \\ & + \frac{1}{2}(X_{t+1} - \mathbb{A}_oX_t)'(\mathbb{B}_o\mathbb{B}_o')^{-1}(X_{t+1} - \mathbb{A}_oX_t) \\ & - \frac{1}{2} \log \det(\mathbb{B}\mathbb{B}') + \frac{1}{2} \log \det(\mathbb{B}_o\mathbb{B}_o') \end{aligned}$$

Notice how the matrices  $(\mathbb{A}_o, \mathbb{B}_o)$  of the baseline model and parameters  $(\mathbb{A}, \mathbb{B})$  of the alternative model both appear.

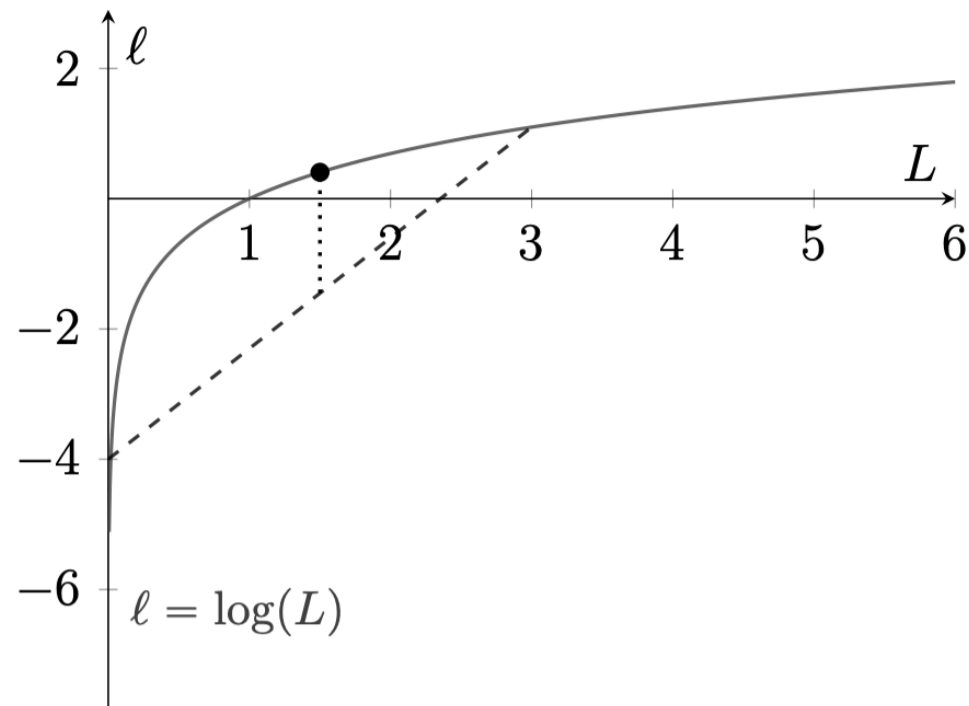
### Remark 8.1

Because  $\mathbb{B}$  is a nonsingular square matrix, model [Example 8.5](#) has the same number of shocks, i.e., entries of  $W$ , as there are components of  $X$ . A more general setting would be a hidden Markov state model like one presented in Section [Kalman Filter and Smoother](#) of Chapter [Hidden Markov Models](#) that has a time-invariant representation with an “information state vector” constructed as a way to condition on an infinite past of an observation vector.

We can elicit a limiting behavior of multiplicative martingales by applying Jensen’s inequality to the concave function  $\log L$  depicted in [Fig. 8.1](#) based on

$$\mathbb{E}(\log L_t | \mathfrak{A}_0) \leq 0$$

where we normalize  $L_0 = 1$ .



*Fig. 8.1* Jensen's Inequality. The logarithmic function is a concave function that equals zero when evaluated at unity. The line segment lies below the logarithmic function. An interior average of endpoints of the straight line lies below the logarithmic function.

Moreover, by Jensen's inequality,

$$\mathbb{E}(\log N_{t+1} \mid \mathfrak{A}_t) \leq \log \mathbb{E}(N_{t+1} \mid \mathfrak{A}_t) = 0.$$

for  $N_{t+1}$  satisfying  $L_{t+1} = N_{t+1}L_t$ . Note that

$$\mathbb{E}(\log L_{t+1} \mid \mathfrak{A}_t) = \log L_t + \mathbb{E}(\log N_{t+1} \mid \mathfrak{A}_t) \leq \log L_t.$$

This implies that under the baseline model the log-likelihood ratio process  $L$  is a *supermartingale* relative to the information sequence  $\{\mathfrak{A}_t : t \geq 0\}$ .

From the Law of Large Numbers described [Chapter 1: Laws of Large Numbers and Stochastic Processes](#), a population mean is well approximated by a sample average from a long time series. That opens the door to discriminating between two models. Under the baseline model, the log likelihood ratio process scaled by  $1/t$  converges to a negative number. If the baseline model actually generates the data, the expected log likelihood ratio constructed with data will eventually be negative except in the degenerate case in which  $N_{t+1} = 1$  with probability one. Such a calculation justifies discriminating between the two models by calculating  $\log L_t$  and checking if it is positive or negative. This procedure amounts to an application of the method of maximum likelihood. Sometimes

$$-\mathbb{E}(\log N_{t+1} \mid \mathfrak{A}_t) \geq 0$$

is used as a measure of statistical divergence.

Suppose now that we reverse the roles of the baseline and alternative probability measures.

### Definition 8.2

The **conditional relative entropy** of a martingale increment is defined to be:

$$E(N_{t+1} \log N_{t+1} \mid \mathcal{A}_t) \geq 0.$$

This entity is sometimes referred as Kullback-Leibler divergence. To understand why conditional relative entropy is nonnegative, observe that multiplication of  $\log N_{t+1}$  by  $N_{t+1}$  changes the conditional probability distribution for which the conditional expectation of  $\log N_{t+1}$  is calculated from the baseline model to the alternative model. The function  $n \log n$  is convex and equal to zero for  $n = 1$ . Therefore, Jensen's inequality implies that conditional relative entropy is nonnegative and equal to zero when  $N_{t+1} = 1$  conditioned on  $\mathcal{A}_t$ .

Notice that

$$\begin{aligned} \mathbb{E}(L_{t+1} \log L_{t+1} \mid \mathcal{A}_t) &= L_t \mathbb{E}(N_{t+1} \log N_{t+1} \mid \mathcal{A}_t) + L_t \log L_t \\ &\geq L_t \log L_t. \end{aligned}$$

Thus  $L \log L$  is a *submartingale*. The expression

$$\mathbb{E}(L_t \log L_t \mid \mathcal{A}_0) \geq 0,$$

and is a measure of relative entropy over a  $t$ -period horizon. Relative entropy is often used to analyze model misspecifications and also appears in statistical characterizations used in information theory and large deviations theory for Markov processes (see [\[Cho et al., 2002\]](#)).

### Remark 8.2

Consider the following example of likelihood-based model selection. Suppose that a decision-maker does not know whether a baseline or alternative model generates the data. Attach a subjective prior probability  $\pi_o$  to the baseline probability model and probability  $1 - \pi_o$  on the alternative. Let  $L$  be a likelihood ratio process with  $L_t$  reflecting information available at date  $t$ . Date  $t$  posterior probabilities for the baseline and alternative probability models are:

$$\frac{\pi_o}{L_t(1 - \pi_o) + \pi_o} \quad \text{and} \quad \frac{L_t(1 - \pi_o)}{L_t(1 - \pi_o) + \pi_o}.$$

When  $\frac{1}{t} \log L_t$  converges to a negative number under the baseline probability, the first probability converges to one. But when  $\frac{1}{t} \log L_t$  converges to a positive number under the

alternative probability, the second probability converges to one. When the data are generated by the baseline probability model, the Law of Large Numbers implies the former; and when the data are generated by the alternative probability model, the Law of Large Numbers implies the latter. This shows that model selection based on posterior probabilities eventually selects the model that generated the data, i.e., either the baseline model or the alternative model. This analysis can be extended to situations in which some other model generates the data.

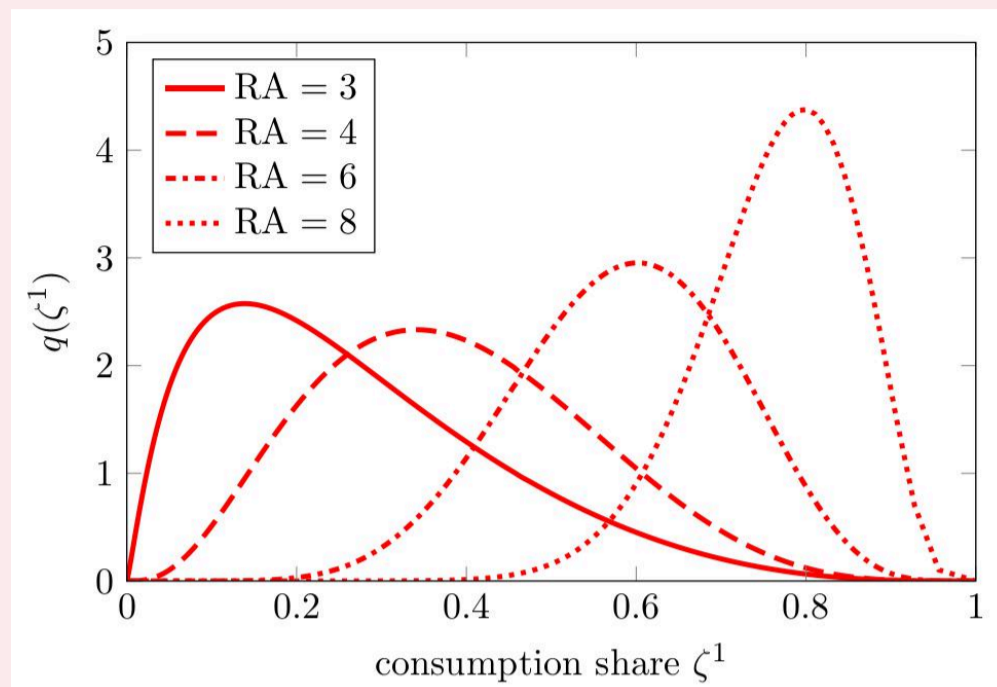
### Remark 8.3

Martingales that represent changes of measure appear in asset pricing models that posit a difference between a representative investor's "subjective" model and a distinct model that a researcher or economic agents assume actually governs the data. Parts of the behavioral finance literature appeal to insights from psychology to motivate violations of that communism of beliefs axiom. Our tools can help us interpret such models by leading us to think about sources of those model discrepancies in terms of the understandings and motivations that might lead a sophisticated investor to acknowledge possible discrepancies between a subjective model and the type of baseline approximating model commonly used in asset pricing models. Using martingales to represent belief distortions allows us to frame topics in the behavioral finance literature. For example, to be intelligible, papers about over or under reactions must tell a reader with respect to what measure these reactions are to be judged to be mistakes. Our framework invites saying more about the martingale that accounts those over or under reactions. In some settings statistical divergence measures tell us that it is very difficult for sophisticated statisticians to distinguish or select among alternative approximating models as data descriptions. Investors' belief distortions are likely to persist when statistical model-selection challenges confront both investors and the econometrician modeling them. We revisit these topics later in this chapter. See [Section 7.7](#) and [Section 7.9](#).

### Remark 8.4

These tools are also pertinent to the study of models in which investors have heterogeneous beliefs. [\[Alchian, 1950\]](#) and [\[Friedman, 1953\]](#), among others, have argued that investors with distorted beliefs will eventually be driven out of the market by investors with more accurate perceptions of the future because of the relative success of the latter type of investors. This has been one argument for imposing rational expectations in asset pricing models. [\[Kogan et al., 2006\]](#) refine this view by breaking a simple link between survival and price impact of the investors with distorted beliefs. [\[Borovička, 2020\]](#) goes further by characterizing families of investor preferences for which the investors with the distorted beliefs survive in the long run. These latter two contributions feature differences in how investors look at stochastic growth. A continuous-time counterpart to the martingale representation for belief distortions that we

describe and characterize in this chapter are featured in both the [\[Kogan et al., 2006\]](#) and the [\[Borovička, 2020\]](#) papers.



This figure is from [\[Borovička, 2020\]](#). It plots the stationary densities for the fraction of aggregate consumption allocated to the first of two types of investors. The first is mistakenly optimistic about future consumption growth. The second knows the correct evolution. The figure illustrates that both consumers “survive” in the long run and that the allocation shifts in favor of the optimistic investor when both are more risk averse.

## 8.4. Factoring a multiplicative functional

Following [\[Hansen and Scheinkman, 2009\]](#) and [\[Hansen, 2012\]](#), we factor a multiplicative functional into three multiplicative components having the primitive types [Example 8.1](#), [Example 8.2](#), [Example 8.3](#).

As in definition [Definition 8.1](#), let  $Y$  be an additive functional, and let  $M = \exp(Y)$ . Apply a one-period operator  $\mathbb{M}$  defined by

$$\begin{aligned} \mathbb{M}f(x) &\stackrel{\text{def}}{=} \mathbb{E} [\exp(Y_{t+1} - Y_t) f(X_{t+1}) \mid X_t = x] \\ &= \mathbb{E} \left[ \left( \frac{M_{t+1}}{M_t} \right) f(X_{t+1}) \mid X_t = x \right] \end{aligned} \quad (8.6)$$

to bounded Borel measurable functions  $f$  of the Markov state. By applying the Law of Iterated expectations, a two-period operator iterates  $\mathbb{M}$  twice to obtain:

$$\begin{aligned} \mathbb{M}^2 f(x) &\stackrel{\text{def}}{=} \mathbb{E} [\exp(Y_{t+2} - Y_t) f(X_{t+2}) \mid X_t = x] \\ &= \mathbb{E} \left[ \left( \frac{M_{t+2}}{M_t} \right) f(X_{t+2}) \mid X_t = x \right], \end{aligned} \quad (8.7)$$

with corresponding definitions of  $j$ -period operators  $\mathbb{M}^j$ . The family of operators is a special case of what is called a "semi-group." The domain of the semigroup can typically be extended to a larger family of functions, but this extension depends on further properties of the multiplicative process used to construct it.

We next derive a revealing representation of this semigroup by factoring the multiplicative functional  $M$  in an interesting way. We achieve this by applying what is referred to in mathematics as Perron-Frobenius theory based on the following equation:

**Eigenvalue-eigenfunction Problem:** Solve

$$\mathbb{M}\tilde{e}(x) = \exp(\tilde{\eta})\tilde{e}(x) \quad (8.8)$$

for an eigenvalue  $\exp(\tilde{\eta})$  and a positive eigenfunction  $\tilde{e}$ .

Consistent with Perron-Frobenius theory, call the positive eigenvalue, the **principal eigenvalue**, and the associated positive eigenvector, the **principal eigenfunction**, of the operator  $\mathbb{M}$ .

Use the principal eigenvalue and eigenvector, and define:

$$\tilde{N}_{t+1} \stackrel{\text{def}}{=} \exp(-\tilde{\eta}) \frac{M_{t+1}\tilde{e}(X_{t+1})}{M_t\tilde{e}(X_t)} = \exp[\tilde{\kappa}(X_t, W_{t+1})],$$

Note that we constructed  $\tilde{N}_{t+1}$  so as to have a conditional expectation equal to unity. Using  $\tilde{N}$  as a martingale increment process, build

$$\tilde{L}_{t+1} = \tilde{N}_{t+1}\tilde{L}_t, \quad \tilde{L}_0 = 1.$$

### Theorem 8.1

Let  $M$  be a multiplicative functional. Suppose that the principal eigenvalue-eigenfunction [problem](#) has a solution with principal eigenfunction  $\tilde{e}(X)$ . Then the multiplicative functional is the product of three components that are instances of the primitive functionals in examples [Example 8.1](#), [Example 8.2](#), and [Example 8.3](#):

$$\frac{M_t}{M_0} = \exp(\tilde{\eta}t)\tilde{L}_t \left[ \frac{\tilde{e}(X_0)}{\tilde{e}(X_t)} \right] \quad (8.9)$$

where  $\tilde{L}_t$  is a multiplicative martingale with  $\tilde{L}_0 = 1$ .

The factorization of a multiplicative functional described in [Theorem 8.1](#) is a counterpart to the [Chapter 4 Proposition 4.1](#) decomposition of an additive functional. We used the [Proposition 4.1](#) martingale to identify the permanent component of an additive functional in [Chapter 4](#). In this chapter, we shall use the multiplicative martingale isolated by [Theorem 8.1](#) to represent a change of probability measure, one that will help us understand long-term risk return tradeoffs. While  $\tilde{L}$  is a multiplicative martingale,  $\log \tilde{L}$  is a super martingale. There is a weaker connection between the two representations, however. Write the additive decomposition as:

$$\log M_t - \log M_0 = t\hat{\eta} + \hat{L}_t^+ \hat{e}(X_0) - \hat{e}(X_t)$$

where  $\hat{L}$  is an additive martingale. As [\[Hansen, 2012\]](#) argues, if  $\tilde{L}$  is not degenerate (not equal to one), then  $\hat{L}^s$  is not degenerate (not equal to zero) and conversely.

Models in which  $X$  is a finite-state Markov chain are give a direct computation in which the principal eigenvalue calculation reduces to finding an eigenvector of a matrix with all positive entries.

### Example 8.6

The stochastic process  $X$  is governed by a finite-state Markov chain on state space  $\{\mathbf{s}_1, \mathbf{s}_2, \dots, \mathbf{s}_n\}$ , where  $\mathbf{s}_i$  is the  $n \times 1$  vector whose components are all zero except for 1 in the  $i^{th}$  row. The transition matrix is  $\mathbb{P}$ , where  $\mathbf{p}_{ij} = \text{Prob}(X_{t+1} = \mathbf{s}_j | X_t = \mathbf{s}_i)$ . We represent the Markov chain as

$$X_{t+1} = \mathbb{P}' X_t + W_{t+1}$$

where  $\mathbb{E}(X_{t+1} | X_t) = \mathbb{P}' X_t$ ,  $\mathbb{P}'$  denotes the transpose of  $\mathbb{P}$ , and  $W_{t+1}$  is an  $n \times 1$  vector process that satisfies  $\mathbb{E}(W_{t+1} | X_t) = 0$ , which is therefore a martingale-difference sequence adapted to  $X_t, X_{t-1}, \dots, X_0$ . Think of  $W_{t+1}$  as the vector of errors when forecasting  $X_{t+1}$  based on current information.

Let  $\mathbb{G}$  be an  $n \times n$  matrix whose  $(i, j)$  entry  $\mathbf{g}_{ij}$  is an additive contribution to the growth that  $Y_{t+1} - Y_t$  experiences when  $X_{t+1} = \mathbf{s}_j$  and  $X_t = \mathbf{s}_i$ . The stochastic process  $Y$  is governed by the *additive functional*

$$Y_{t+1} - Y_t = (X_t)' \mathbb{G} X_{t+1} = (X_t)' \mathbb{G} \mathbb{P}' X_t + (X_t)' \mathbb{G} W_{t+1}$$

Let  $M = \exp(Y)$ . Define a matrix  $\mathbb{M}$  whose  $(i, j)^{th}$  element is  $\mathbf{m}_{ij} = \exp(\mathbf{g}_{ij})$ . Represent the stochastic process  $M$  as the *multiplicative functional*:

$$\frac{M_{t+1}}{M_t} = \exp [(X_t)'GX_{t+1}] = (X_t)'MX_{t+1}. \quad (8.10)$$

Associated with this multiplicative functional is the principal eigenvalue problem

$$\mathbb{E} \left[ \frac{M_{t+1}}{M_t} \tilde{e} \cdot X_{t+1} | X_t = x \right] = \exp(\tilde{\eta}) \tilde{e} \cdot x.$$

To convert this to a linear algebra problem, write the  $j^{\text{th}}$  entry of  $\tilde{e}$  as  $\tilde{e}_j$ . Since  $X_t$  always assumes the value of one of the coordinate vectors  $\mathbf{s}_i, i = 1, \dots, n$ ,

$$(X_t)'MX_{t+1} = \mathbf{m}_{ij}$$

when  $X_t = \mathbf{s}_i$  and  $X_{t+1} = \mathbf{s}_j$ . This allows us to rewrite the principal eigenvalue problem as

$$\sum_j \mathbf{p}_{ij} \mathbf{m}_{ij} \tilde{e}_j = \exp(\tilde{\eta}) \tilde{e}_i$$

or

$$\widetilde{\mathbf{M}}\tilde{e} = \exp(\tilde{\eta})\tilde{e} \quad (8.11)$$

where  $\tilde{\mathbf{m}}_{ij} \stackrel{\text{def}}{=} \mathbf{p}_{ij} \mathbf{m}_{ij}$  and  $\tilde{e}_i$  is entry  $i$  of  $\tilde{e}$ .

Notice that this construction incorporates the transition probabilities into the construction of the matrix  $\widetilde{\mathbf{M}}$ . We do this to reduce the problem to one of finding an eigenvalue and corresponding eigenvector of a matrix. Specifically, we want the positive eigenvalue associated with a positive eigenvector of [\(8.11\)](#).

After solving the principal eigenvalue problem, we construct an alternative transition matrix for a finite-state Markov process. Construct

$$\hat{\mathbf{n}}_{ij} = \exp(-\tilde{\eta}) \tilde{\mathbf{m}}_{ij} \frac{\tilde{e}_j}{\tilde{e}_i}, \quad (8.12)$$

and form the matrix  $\hat{\mathbf{N}} = [\hat{\mathbf{n}}_{ij}]$ . The matrix  $\hat{\mathbf{N}}$  has entries that are nonnegative, and

$$\sum_{j=1}^n \hat{\mathbf{n}}_{ij} = 1,$$

which verifies that  $\hat{\mathbb{N}}$  can be viewed as a transition matrix.

Finally, we show how to construct the factorization of the  $M$  process. To build the martingale of interest, we must undo the probability scaling that we did when forming  $\hat{\mathbb{N}}$ . With this in mind, solve for  $\tilde{n}_{ij}$  that satisfies:

$$\hat{n}_{ij} = \tilde{n}_{ij}p_{ij},$$

where we arbitrarily set  $\tilde{n}_{ij}$  equal one when  $p_{ij} = 0$  and build the matrix  $\tilde{\mathbb{N}} = [\tilde{n}_{ij}]$ . These constructions allow us to write (8.10) as

$$\frac{M_{t+1}}{M_t} = \exp(\tilde{\eta}) \left[ (X_t)' \tilde{\mathbb{N}} X_{t+1} \right] \left( \frac{\tilde{e} \cdot X_t}{\tilde{e} \cdot X_{t+1}} \right). \quad (8.13)$$

#### Remark 8.5

In the previous example, the matrix  $\tilde{\mathbb{M}}$  could be inferred from prices of one-period-ahead Arrow securities. Specifically, given prices of all of the state contingent claims for next period, we can infer the column of  $\tilde{\mathbb{M}}$  associated with the current state. By observing prices for all of the current states, all of the columns of  $\tilde{\mathbb{M}}$ . We only need to know  $\tilde{\mathbb{M}}$  to pose and solve the principal eigenvalue problem and hence to construct transition probabilities as given by (8.12), but not the  $\mathbf{p}_{ij}$ 's and the  $\mathbf{s}_{ij}$ 's. This gives us a way to *recover* a transition probabilities from asset prices in the language of [Ross, 2015]. Notice, in particular, that the  $\mathbf{p}_{ij}$ 's and  $\mathbf{s}_{ij}$ 's cannot be inferred uniquely from the  $\mathbf{m}_{ij}$ 's. In particular, the Arrow prices are insufficient to *identify* the transition probabilities. Ross imposed restrictions on  $\mathbf{s}_{ij}$ 's that implied that the recovered probabilities are actually the  $\mathbf{p}_{ij}$ 's. In general, the recovered probabilities will differ from the baseline transition probability matrix. We will have more to say about this difference and investigate why the recovered probabilities remain interesting. See [Borovička et al., 2016] for an extended discussion of this claim.

The following log-linear, log-normal specification displays relevant mechanics of the multiplicative factorization with direct connection to the corresponding additive decomposition

#### Example 8.7

Consider a stationary  $X$  process and an additive  $Y$  process described by the VAR

$$\begin{aligned} X_{t+1} &= \mathbb{A}X_t + \mathbb{B}W_{t+1} \\ Y_{t+1} - Y_t &= \nu + \mathbb{D}X_t + \mathbb{F}W_{t+1} \end{aligned}$$

where  $\mathbb{A}$  is a stable matrix and  $\{W_{t+1} : t \geq 0\}$  is a sequence of independent and identically normally distributed random vectors with mean zero and covariance matrix  $\mathbb{I}$ . In [Proposition 4.1](#) of [Chapter 4](#), we described the decomposition

$$Y_t - Y_0 = t\nu + \left[ \sum_{j=1}^t \mathbb{H}W_j \right] - g(X_{t-1}) + g(X_0) \quad (8.14)$$

where

$$\begin{aligned} \mathbb{H} &= \mathbb{F} + \mathbb{D}(\mathbb{I} - \mathbb{A}')^{-1}\mathbb{B} \\ g(x) &= \mathbb{D}(\mathbb{I} - \mathbb{A})^{-1}x. \end{aligned}$$

Let  $M_t = \exp(Y_t)$ . Use equation [\(8.14\)](#) to deduce

$$\frac{M_t}{M_0} = \exp(\tilde{\eta}t) \tilde{L}_t \begin{bmatrix} \tilde{e}(X_0) \\ \tilde{e}(X_t) \end{bmatrix}$$

where

$$\tilde{\eta} = \nu + \frac{|\mathbb{H}|^2}{2},$$

$$\tilde{N}_{t+1} = \exp\left(\mathbb{H}W_{t+1} - \frac{|\mathbb{H}|^2}{2}\right), \quad \tilde{L}_0 = 1, \quad (8.15)$$

and

$$\tilde{e}(x) = \exp[g(x)] = \exp\left[\mathbb{D}(\mathbb{I} - \mathbb{A})^{-1}x\right].$$

While the additive martingale of  $Y = \log(M)$  has a variance that grows linearly over time, this variance contributes a component to the exponential trend of the multiplicative functional  $M$  along with a simple time-horizon adjustment to the martingale component.

### Example 8.8

Consider the following stochastic volatility example. Suppose that a scalar state variable evolve as first-order autoregression:

$$\begin{aligned} X_{t+1} &= \mathbf{a}X_t + \mathbf{b}W_{t+1} \\ Y_{t+1} - Y_t &= \nu + (\mathbf{f}_0 + \mathbf{f}_1 X_t)W_{t+1} \end{aligned}$$

where  $\{W_{t+1} : t \geq 0\}$  is an iid sequence of normally distributed random variables. The state variable  $X_t$  gives the source of volatility fluctuations. Guess a principle eigenfunction of the form:

$$\tilde{e}(x) = \exp\left(\epsilon_1 x + \frac{\epsilon_2}{2} x^2\right),$$

and construct

$$e^-(x, w) \stackrel{\text{def}}{=} \epsilon_1 \mathbf{a}x + \epsilon_1 \mathbf{b}w + \frac{\epsilon_2}{2} \mathbf{a}^2 x^2 + \epsilon_2 \mathbf{a}\mathbf{b}xw + \frac{\epsilon_2}{2} \mathbf{b}^2 w^2.$$

Write the principle eigenfunction equation as:

$$\begin{aligned} \log \mathbb{E} \left( \exp \left[ \nu + (\mathbf{f}_0 + \mathbf{f}_1 X_t)W_{t+1} + e^-(X_t, W_{t+1}) \right] \mid X_t = x \right) \\ = \tilde{\eta} + \epsilon_1 x + \frac{\epsilon_2}{2} x^2, \end{aligned}$$

where we took logarithms of both sides of the equation. The computation on left side of this equation has a tractable formula for expressing the outcome as a quadratic function of the state.

Solve the equation in three steps. First, equate coefficients on  $x^2$  and deduce a quadratic equation for  $\epsilon_2$ . Next, equate coefficients on  $x$  and obtain a linear equation for  $\epsilon_1$ . Finally, equate constants obtain an equation for  $\tilde{\eta}$ . The initial quadratic equation is

$$\mathbf{b}^2(\epsilon_2)^2 + (\mathbf{a}^2 + 2\mathbf{f}_1\mathbf{a}\mathbf{b} - 1)\epsilon_2 + (\mathbf{f}_1)^2 = 0.$$

When it has a solution, there is will typically be two possible choices. For the solution to be of interest,  $1 - \epsilon_2\mathbf{b}^2$  must be positive.

Second, equate coefficient on  $x$  to solve an equation for  $\epsilon_1$  given  $\epsilon_2$ :

$$(\mathbf{f}_1 + \epsilon_2\mathbf{a}\mathbf{b})(\mathbf{f}_0 + \epsilon_1\mathbf{b}) + \epsilon_1(\mathbf{a} - 1)(1 - \epsilon_2\mathbf{b}^2) = 0.$$

Third, we solve for  $\tilde{\eta}$  by equating the constant and plugging the solutions for  $\epsilon_1$  and  $\epsilon_2$ :

$$\tilde{\eta} = \frac{(f_0 + \epsilon_1 \mathbf{b})^2}{1 - \epsilon_2 \mathbf{b}^2} - \frac{1}{2} \log(1 - \epsilon_2 \mathbf{b}^2) + \nu.$$

For this example,

$$\tilde{N}_{t+1} \propto \exp \left[ (f_0 + f_1) X_t W_{t+1} + \epsilon_1 \mathbf{b} W_{t+1} + \frac{\epsilon_2}{2} \mathbf{b}^2 (W_{t+1})^2 + \epsilon_2 \mathbf{a} \mathbf{b} X_t W_{t+1} \right]$$

where  $\propto$  means proportional to up to a scale factor that depends on  $X_t$ . Under the implied change in measure,  $W_{t+1}$  is distributed as normal random variable with conditional mean

$$\frac{f_0 + f_1 X_t + \epsilon_1 \mathbf{b} + \epsilon_2 \mathbf{a} \mathbf{b} X_t}{1 - \epsilon_2 \mathbf{b}^2}$$

and precision  $1 - \epsilon_2 \mathbf{b}^2$ . Under the change of probability measure,  $W_{t+1}$  remains normally distributed with a different variance and state-dependent mean. Under this change in probability measure, the  $X$  remains a first-order autoregression but with different dynamics. It could, for instance, be an explosive autoregression.

#### Remark 8.6

The martingale component of the multiplicative functional has “peculiar behavior.” It has expectation one by construction. The Martingale Convergence Theorem guarantees that sample paths converge, typically to zero. This theorem is operative because the positive martingale is bounded from below. Since its date zero conditional expectation is one, for long horizons this process necessary has a fat right tail. [Fig. 8.2](#) plots probability density functions of the martingale component for different values of  $t$ .

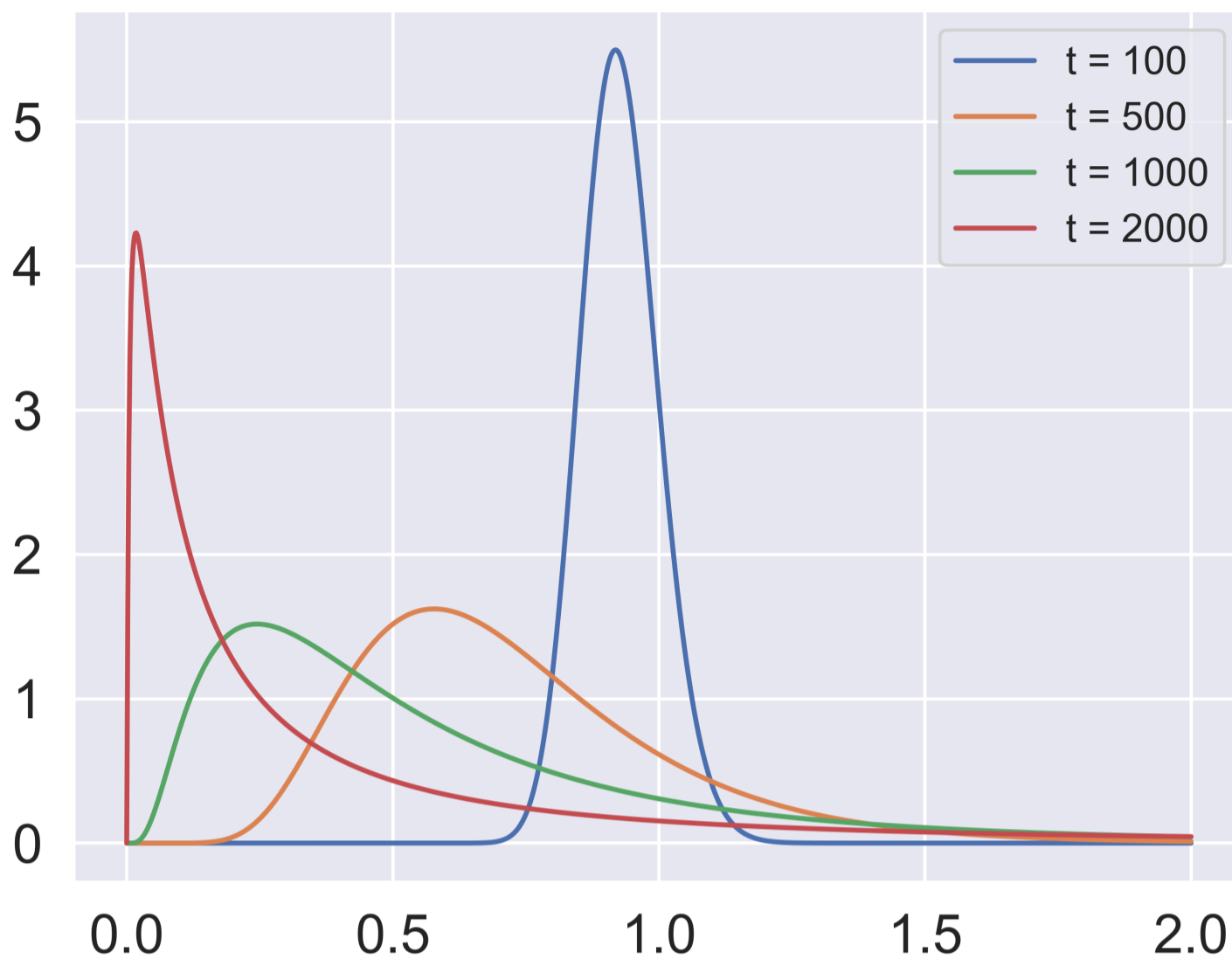


Fig. 8.2 Density of  $\tilde{L}_t$  for different values of  $t$ .

#### Remark 8.7

Consider a positive cumulative return process,  $R$ , modeled as a stochastic functional. As we noted previously, for such a process,  $R_t/R_\tau$  for  $\tau < t$  is a  $t - \tau$ -period return for any such  $t$  and  $\tau$ . Consider also the associated cumulative stochastic discount factor process,  $S$ , which we also take to be a multiplicative functional. For this process,  $S_t/S_\tau$  for  $\tau < t$  is the  $t - \tau$ -period stochastic discount factor for any such  $t$  and  $\tau$ . Normalize  $R_0 = 1$  and  $S_0 = 1$ . Repeating what we stated previously, by standard asset pricing logic, SR is a multiplicative martingale. [Martin, 2012] studies tail behavior of cumulative returns. Since its date zero conditional expectation is one, for long horizons this process necessarily has a fat right tail as illustrated by Fig. 8.2.

While the multiplicative martingale has peculiar sample path properties, we are primarily interested in this martingale component as a change of probability measure. For instance, in Example 8.7, formula (8.15) for  $\tilde{N}_t$  tells how the change in probability measure induces mean  $\mathbb{H}$  in the conditional distribution for the shock  $W_{t+1}$ . Similarly,  $\tilde{\mathbb{L}}$ , with entries given in formula (8.12), provides an alternative transition matrix in Example 8.6.

## 8.5. Stochastic stability

Our characterization of a change of probability measure as the solution of a Perron-Frobenius problem determines only transition probabilities. Since the process is Markov, it is of interest to seek an initial distribution of  $X_0$  under which the process is stationary and satisfies a stochastic stability property that we will define and explore. Stochastic stability opens the door the study of a variety of limiting behavior and hence justifies our interest in the multiplicative factorization. The [eigenfunction problem](#) can have multiple solutions, it turns out, however, that there is a unique solution for which the process  $X$  is **stochastically stable** under the implied change of measure. See [[Hansen and Scheinkman, 2009](#)] and [[Hansen, 2012](#)] for formal analyses of this problem in a continuous-time Markov setting. In what follows we investigate the discrete-time counterpart to their investigations.

### Definition 8.3

A process  $X$  is **stochastically stable** under a probability measure  $\tilde{P}_r$  if it is stationary and  $\lim_{j \rightarrow \infty} \tilde{\mathbb{E}} [f(X_j) | X_0 = x] = \tilde{\mathbb{E}} [f(X_0)]$  for any Borel measurable  $h$  satisfying  $\tilde{\mathbb{E}} |f(X_t)| < \infty$ .

As we discussed previously, a multiplicative martingale  $L$  with  $L_0 = 1$  induces a probability measure conditioned on  $X_0$ . To check for stochastic stability, we extend the probability by including a distribution over  $X_0$  in a manner that ensure the process  $X$  is stationary.

### Remark 8.8

Notice that the stationary probability distribution for  $X_t$  implied by  $\tilde{P}_r$  is used to defined the limits of conditional the conditional expectations. For the mathematical results, however we initialize the distribution over  $X_0$  is not crucial. Thus we could equivalently consider other probabilities measures for which the same limiting distribution for  $X_t$  as  $t$  gets large. Under such probability measures, the process  $X$  would be asymptotically stationary. In this sense, we impose the stronger stationarity restriction as a matter of convenience.

### Theorem 8.2

Let  $M$  be a multiplicative functional. Suppose that  $(\tilde{\eta}, \tilde{e})$  solves the [eigenfunction problem](#) and that under the change of measure  $\tilde{P}_r$  implied by the associated martingale  $\tilde{L}$  the stochastic process  $X$  is stationary and ergodic. Consider any other solution  $(\eta^*, e^*)$  to [eigenfunction problem](#) with implied martingale  $\{L_t^*\}$ . Then

1.  $\eta^* \geq \tilde{\eta}$ .

2. If  $X$  is stochastically stable under the change of measure  $Pr^*$  implied by the martingale  $L^*$ , then  $\eta^* = \tilde{\eta}$ ,  $e^*$  is proportional to  $\tilde{e}$ , and  $L^* = \tilde{L}$  for all  $t = 0, 1, \dots$

### Proof

First we show that  $\eta^* \geq \tilde{\eta}$ . Write:

$$\mathbb{M}^t e^*(x) = \exp(\tilde{\eta}t) \tilde{\mathbb{E}} \left( \left[ \frac{\tilde{e}(X_0)}{\tilde{e}(X_t)} \right] e^*(X_t) \middle| X_0 = x \right) = \exp(\eta^*t) e^*(x).$$

Thus,

$$\tilde{\mathbb{E}} \left( \left[ \frac{e^*(X_t)}{\tilde{e}(X_t)} \right] \middle| X_0 = x \right) = \exp(\eta^*t - \tilde{\eta}t) \left[ \frac{e^*(x)}{\tilde{e}(x)} \right].$$

If  $\tilde{\eta} > \eta^*$ , then

$$\lim_{t \rightarrow \infty} \tilde{\mathbb{E}} \left( \left[ \frac{e^*(X_t)}{\tilde{e}(X_t)} \right] \middle| X_0 = x \right) = 0.$$

But this equality cannot be true because under  $\tilde{Pr}$ ,  $X$  is stochastically stable and  $\frac{e^*}{\tilde{e}}$  is strictly positive. Therefore,  $\eta^* \geq \tilde{\eta}$ .

Consider next the case in which  $\eta^* > \tilde{\eta}$ . Write

$$\frac{M_t}{M_0} = \exp(\eta^*t) \left( \frac{L_t^*}{L_0^*} \right) \left( \frac{e^*(X_0)}{e^*(X_t)} \right),$$

which implies that

$$\mathbb{M}^t \tilde{e}(x) = \exp(\eta^*t) E^* \left[ \left( \frac{e^*(X_0)}{e^*(X_t)} \right) \tilde{e}(X_t) \middle| X_0 = x \right] = \exp(\tilde{\eta}t) \tilde{e}(x).$$

Thus,

$$E^* \left( \left[ \frac{\tilde{e}(X_t)}{e^*(X_t)} \right] \middle| X_0 = x \right) = \exp(\tilde{\eta}t - \eta^*t) \left[ \frac{\tilde{e}(x)}{e^*(x)} \right].$$

Suppose that  $\tilde{\eta} < \eta^*$ , then

$$\lim_{t \rightarrow \infty} \mathbb{E}^* \left( \left[ \frac{\tilde{e}(X_t)}{e^*(X_t)} \right] \middle| X_0 = x \right) = 0,$$

so that  $X$  cannot be stochastically stable under the  $Pr^*$  measure.

Finally, suppose that  $\tilde{\eta} = \eta^*$  and that  $\frac{\tilde{e}(x)}{e^*(x)}$  is not constant. Then

$$\mathbb{E}^* \left( \left[ \frac{\tilde{e}(X_t)}{e^*(X_t)} \right] \middle| X_0 = x \right) = \frac{\tilde{e}(x)}{e^*(x)}$$

and  $X$  cannot be stochastically stable under the  $Pr^*$  measure.

Stochastic stability under the martingale-induced change of measure provides a way to frame some interesting long-term approximations. Let

$$\mathcal{F}(\tilde{L}) \stackrel{\text{def}}{=} \left\{ f > 0 : \tilde{\mathbb{E}} [f(X_t)] < \infty \right\} \quad (8.16)$$

where  $\tilde{L}$  is the martingale component in the factorization of  $M$ . Construct the operator:

$$\tilde{\mathbb{M}}f(x) \stackrel{\text{def}}{=} \frac{1}{\tilde{e}(x)} \mathbb{M}\tilde{e}f(x) = \exp(\tilde{\eta}) \tilde{\mathbb{E}} [f(X_1) | X_0 = x]$$

where we have used the factorization of  $M$ . Form the family (semi group) of operators:

$$\tilde{\mathbb{M}}^j f(x) = \exp(j\tilde{\eta}) \tilde{\mathbb{E}} [f(X_j) | X_0 = x] = \frac{1}{\tilde{e}(x)} \mathbb{M}^j \tilde{e}f(x)$$

where we used the factorization given by [Theorem 8.1](#) and the unique selection of such a representation [Theorem 8.2](#) to obtain the right-side equality. Since  $X$  is stochastically stable under  $\tilde{Pr}$ ,

$$\lim_{j \rightarrow \infty} \frac{1}{j} \log \tilde{\mathbb{M}}^j f(x) = \tilde{\eta} + \lim_{j \rightarrow \infty} \frac{1}{j} \log \tilde{\mathbb{E}} [f(X_j) | X_0 = x] = \tilde{\eta}$$

for all  $f \in \mathcal{F}(\tilde{L})$ . In particular,

$$\begin{aligned}
\lim_{j \rightarrow \infty} \frac{1}{j} \log M^j f(x) &= \lim_{j \rightarrow \infty} \frac{1}{j} \log \left[ \exp(j\tilde{\eta}) \tilde{e}(x) \widetilde{M}^j \tilde{f}(x) \right] \\
&= \tilde{\eta} + \frac{1}{j} \log \tilde{e}(x) + \lim_{j \rightarrow \infty} \frac{1}{j} \log \widetilde{M}^j \tilde{f}(x) \\
&= \tilde{\eta}
\end{aligned} \tag{8.17}$$

for  $f = \tilde{f}\tilde{e}$  provided that  $\tilde{f} \in \mathcal{F}(\tilde{L})$ . Observe that the division by  $\tilde{e}$  when constructing  $\tilde{f}$  absorbs the transient contribution to the factorization of  $M$ :

$$\frac{M_t f(X_t)}{M_0 f(X_0)} = \exp(\tilde{\eta}t) \tilde{L}_t \left[ \frac{\tilde{f}(X_t)}{\tilde{f}(X_0)} \right] M_0$$

We use limit (8.17) to justify calling the Perron-Frobenius eigenvalue as the long-term growth or decay rate of the multiplicative functional.

Under the stochastic stability restriction, we obtain a more refined approximation by adjusting for the growth or decay in the family of operators:

$$\begin{aligned}
\lim_{j \rightarrow \infty} \exp(-\tilde{\eta}j) M^j f(x) &= \lim_{j \rightarrow \infty} \tilde{\mathbb{E}} \left[ \tilde{f}(X_j) \mid X_0 = x \right] \tilde{e}(x) \\
&= \tilde{E} \left[ \tilde{f}(X_t) \right] \tilde{e}(x),
\end{aligned}$$

where we assume that  $\tilde{f} = f/\tilde{e}$  is in  $\mathcal{F}(\tilde{L})$ . Thus once we adjust for the impact of  $\tilde{\eta}$ , the limiting function is proportional to  $\tilde{e}$ . The function  $f$  determines only a scale factor  $\tilde{\mathbb{E}} \left[ \frac{f(X_t)}{\tilde{e}(X_t)} \right] \tilde{e}(x)$ .

We will apply the limiting results in a variety of ways in this and subsequent chapters. The multiplicative factorization can help us understand implications of stochastic equilibrium models for valuations of random payout processes. In addition, such factorizations can help organize empirical evidence in ways that make contact with such stochastic equilibrium asset pricing models.

## 8.6. Inferences about permanent shocks from asset prices

Macroeconomists often study dynamic impacts of shocks to systems of variables measured in logarithms. For example, [Alvarez and Jermann, 2005] suggest looking at asset prices using a multiplicative representation of a cumulative stochastic discount factor, though without the tools provided by this chapter. The additive decomposition derived and analyzed in Chapter 4 is a convenient tool for identifying permanent shocks. A prominent multiplicative martingale component implies a prominent role for permanent shocks in the underlying economic dynamics. A formal probability model lets us link an

additive decomposition as described in this chapter and [Chapter 4](#) to the multiplicative factorization that we study in this chapter. Log normal models, at least as approximations, are often used by applied macroeconomists. [Example 8.7](#) provides an example with explicit formulas linking the two representations. While distinct, the two martingales, in this special case, are closely linked. In general, the connection is more subtle. In what follows we describe how a multiplicative martingale component to a stochastic discount factor is reflected in asset prices.

Consider stochastic discount factorization

$$S_t = \exp(\eta^s t) L_t^s \left[ \frac{e^s(X_0)}{e^s(X_t)} \right]$$

where  $S_0 = L_0^s = 1$ . A date zero price of a long-term bond is:

$$\mathbb{E}(S_t | \mathfrak{A}_0) = \exp(\eta^s t) \tilde{\mathbb{E}} \left[ \frac{1}{e^s(X_t)} \middle| X_0 \right] e^s(X_0).$$

Compute the corresponding yield by taking  $1/t$  times minus the logarithm:

$$-\eta^s - \frac{1}{t} \log \tilde{\mathbb{E}} \left[ \frac{1}{e^s(X_t)} \middle| X_0 \right] + \frac{1}{t} \log e^s(X_0).$$

Provided that

$$\tilde{\mathbb{E}} \left[ \frac{1}{e^s(X_t)} \right] < \infty,$$

the limiting yield on a discount bond is  $-\eta^s$ .

Next consider a one-period holding period return on a  $t$  period discount bond:

$$\frac{\exp[\eta^s(t-1)] \tilde{\mathbb{E}} \left[ \frac{1}{e^s(X_t)} \middle| X_1 \right] e^s(X_1)}{\exp(\eta^s t) \tilde{\mathbb{E}} \left[ \frac{1}{e^s(X_t)} \middle| X_0 \right] e^s(X_0)}$$

Using stochastic stability and taking limits as  $t$  tends to  $\infty$  gives the limiting holding-period return:

$$R_1^\infty \stackrel{\text{def}}{=} \exp(-\eta^s) \frac{e^s(X_1)}{e^s(X_0)}. \quad (8.18)$$

A simple calculation shows that  $R_1^\infty$  satisfies the following equilibrium pricing restriction on a one-period return:

$$\mathbb{E} \left[ \left( \frac{S_1}{S_0} \right) R_1^\infty \middle| \mathfrak{A}_0 \right] = \tilde{\mathbb{E}} \left( \exp(\eta^s) \left[ \frac{e^s(X_0)}{e^s(X_1)} \right] \exp(-\eta^s) \left[ \frac{e^s(X_1)}{e^s(X_0)} \right] \middle| X_0 \right) = 1.$$

These long-horizon limits provide approximations to the eigenvalue for the stochastic discount factor and the ratio of the eigenfunctions.

In a model without a martingale component, [Kazemi, 1992] observed that the inverse of this holding-period return is the one-period stochastic discount factor. Within the [Kazemi, 1992] setup,  $(R_1^\infty)^{-1}$  equals the one-period stochastic factor. Let  $Y_1$  be a vector of asset payoffs and  $Q_0$  be the corresponding vector of prices. Then standard asset pricing theory implies the conditional moments restriction

$$\mathbb{E} \left( (R_1^\infty)^{-1} Y_1 \middle| \mathfrak{A}_0 \right) = Q_0. \quad (8.19)$$

[Alvarez and Jermann, 2005] extend this insight by showing that the reciprocal reveals the component of one-period stochastic discount factor net of its martingale component.

#### Remark 8.9

In practice, we have only have bond data with a finite payoff horizon, whereas the characterizations [Kazemi, 1992] and [Alvarez and Jermann, 2005] use bond prices with a limiting payoff horizon. Empirical implementations using such characterizations assume that the observed term structure data have a sufficiently long duration component to provide plausible proxy for the limiting counterpart.

## 8.7. Digression on rational expectations

Under a rational expectations assumption, a researcher equates the baseline probability measure to be the one implied by the equilibrium of the model. The model of [Kazemi, 1992] provides an example. Kazemi's underlying baseline model is Markov. Kazemi assumes that economic agents living inside his model are completely confident that they know those Markov dynamics. He thus adopts a rational expectations assumption. As noted by [Lucas, 1987] (p.13)

The term 'rational expectations', as Muth used it, refers to a consistency axiom for economic models, so it can be given precise meaning only in the context of specific models. I think this is why attempts to define rational expectations in a model-free way tend to come out either vacuous

(‘People do the best they can with the information they have’) or silly (‘People know the true structure of the world they live in’).

Once we add the assumption that our model-implied expectations actually generates the data we are studying, the rational expectations formulation of Muth and Lucas opens the door to a comprehensive econometric approach. Consider an econometrician who does not know some or all of the parameters of the underlying economic model and who proceeds to infer those parameters using formal statistical methods that have come to be called *rational expectations econometrics*. To elaborate, let’s call probabilities that are model-consistent in the sense of Muth and Lucas ‘‘model-implied probabilities’’. These probabilities can in principle be expressed as a function of the model’s parameters. The outcome is typically a set of cross-equation restrictions that link formally the beliefs of the economic agents within the model to the data generation. Computing these restrictions involves solving a set of nonlinear equations that include Euler equations like the asset pricing equation (8.19) above as well as other equations that help determine the complete Markov dynamics. Being able to compute these functions is what opens the door to the application of formal statistical methods such as maximum likelihood. Importantly, this approach assumes that at least one of the members of the parameterized set of statistical models is consistent with the data generating process and that it is this data generation that the economic agents within the model use in their decision making. An application of rational expectations econometrics thus relies heavily on the economist’s model being *complete* in the sense of [Geweke, 2010]. [Sargent, 1978], [Saracoglu and Sargent, 1978], [Sargent, 1979], [Hansen and Sargent, 1980], [Wallis, 1980], [Hansen and Sargent, 2019] and other contributions in [Lucas and Sargent, 1981] present early applications of this complete-specification approach. [Hansen and Sargent, 2013], ch. 9 describes time-domain and frequency domain maximum likelihood estimation for a class of linear-quadratic-Gaussian dynamic stochastic general equilibrium models. These procedures are widely used throughout applied macro and monetary economics, due in part to the convenient software provided by [Adjemian et al., 2024] and [Herbst and Schorfheide, 2016].

[Hansen and Singleton, 1982] and [Hansen and Richard, 1987] suggest a different approach. They assume that an econometrician only has *partial information* and does not impose or fully parameterize the baseline probability model. Implicitly, they allow for the model components that are not fully parameterized to be revealed imperfectly (except in infinite sample limit) to the econometrician by the actual data generating process. Thus, [Hansen and Singleton, 1982] and [Hansen and Richard, 1987] replace Muth’s consistency axiom operational in a fully specified econometric model with another notion of consistency that is implied by a Law of Large Numbers limit in a statistical setting that assumes that process described by the baseline model is stationary and ergodic. Formally, this approach is implemented with a vector of conditional moment restrictions. [Hansen and Singleton, 1982] and [Hansen and Richard, 1987] thereby make possible a limited-information alternative to imposing rational-expectations with a fully-specified baseline model. Such an approach continues to adopt the simplification that economics agents, in contrast to an econometrician, know the underlying data generating process. Relieved of having to specify the full economic dynamic system, this limited information approach gains

further flexibility by applying the law of iterated expectations to derive implications that the econometrician can use for estimation and testing. Applying such an approach in the context of the [\[Kazemi, 1992\]](#) setup converts the degenerate martingale restriction imposed by Kazemi into a moment restriction that accommodates estimation and statistical testing by using the generalized method of moments. See [\[Hansen, 1982\]](#).

There are now many published papers in the applied macro, industrial organization, and asset pricing literatures that apply either the “rational expectations econometrics” approach that assumes that the econometrician knows a parameterized family models that includes the rational expectations equilibrium, while only the agents know the correct member of this parameterized family; or the alternative approach that assumes that the econometrician has only a partial *a priori* specification and that leverages the Law of Large Numbers implicitly to fill in the rest of the specification.

The rational expectations econometrics assumptions that a correctly-specified model is among the parameterized family and that the economic agents inside the model know key parameters that the econometrician does not are heroic, to put it politely. Even Lucas’s statement above expresses skepticism when he dismisses the statement that ‘People know the true structure of the world they live in’ as being silly. Sophisticated applied economists acknowledge up front that their models are approximations, meaning that they are misspecified. Instead of formally integrating their concerns about possible misspecifications into statistical implementations, for reasons of tractability users of rational expectations econometrics routinely ignore such approximation concerns. For skeptics, the rational expectations econometrics straitjacket is simply too confining, leading them to adopt less formal methods of model matching and validation. Relatedly, without ever claiming or providing evidence that their model-implied expectations agree with an actual data generation process, many papers in applied theory routinely impose *rational expectations* in the process of constructing models of so-called *stylized facts*.

As we argued previously, a subjective belief specification is often expressible as a multiplicative martingale relative to a baseline probability specification, say the rational expectations counterpart or the probability that captures the actual data generation. Thus subjective beliefs offer one candidate for a multiplicative martingale component to the cumulative stochastic discount factor within the [\[Kazemi, 1992\]](#) setup. Statistical divergences provide a way to measure the magnitude of the belief-induced misspecification needed to capture empirical failures that arise from using one of the rational expectations econometrics approaches just described. This opens the door to using a statistical perspective to assess the role of subjective probabilities in models without rational expectations, in contrast to much of the existing research on subjective beliefs in asset pricing. See [Section 7.9](#) for an elaboration. To its credit, rational expectations provides a framework for conducting counterfactuals using underlying dynamic economic model, along the lines of [\[Marschak, 1953\]](#) and [\[Hurwicz, 1966\]](#). More general subjective belief formulations including ones represented by multiplicative martingales, unfortunately, struggle to do something comparable. Deciding when to use misspecified models and for what purposes are important challenges for applied researchers.

## 8.8. Values of stochastic cash flows

In a fundamental paper, [\[Rubinstein, 1976\]](#) featured the importance of cash-flow pricing in contrast to the often studied return-based analyses.

### 8.8.1. Long-term risk-return tradeoff for cash flows

Following [\[Hansen and Scheinkman, 2009\]](#) and [\[Hansen et al., 2008\]](#), we consider the valuation of stochastic cash flows,  $G$ , that are multiplicative functionals. Such cash flows are determinants of prices of both equities and bonds.

We now study valuation from the perspective of the long-term limits measured by the logarithms of the eigenvalues, the  $\tilde{\eta}$ 's. long-term limits of prices of such cash flows. Consider a cash-flow growth process  $G$  measured as a multiplicative functional. The corresponding cash-flow return over horizon  $t$ :

$$\frac{G_t}{\mathbb{E} [ S_t G_t | \mathcal{A}_0 ]} = \frac{\frac{G_t}{G_0}}{\mathbb{E} \left( \frac{S_t G_t}{G_0} \middle| X_0 \right)}$$

where we have normalized  $S_0 = 1$ . Note that as a special case, the cash-flow return on a unit date  $t$  cash-flow is:

$$\frac{1}{\mathbb{E} (S_t | X_0)}$$

Define the *proportional risk premium* on the initial cash-flow return as the ratio of the expected return divided by the riskless counterpart for the same horizon. Taking logarithms and adjusting for the time horizon gives:

$$\begin{aligned} & \frac{1}{t} \log \mathbb{E} (G_t | \mathcal{A}_0) - \frac{1}{t} \log \mathbb{E} (S_t G_t | \mathcal{A}_0) + \frac{1}{t} \log \mathbb{E} (S_t | \mathcal{A}_0) \\ & \frac{1}{t} \left[ \log \mathbb{E} \left( \frac{G_t}{G_0} \middle| X_0 \right) + \log G_0 \right] - \frac{1}{t} \left[ \log \mathbb{E} \left( \frac{S_t G_t}{G_0} \middle| X_0 \right) + \log G_0 \right] + \frac{1}{t} \log \mathbb{E} (S_t | X_0) \quad (8.20) \\ & \frac{1}{t} \log \mathbb{E} \left( \frac{G_t}{G_0} \middle| X_0 \right) - \frac{1}{t} \log \mathbb{E} \left( \frac{S_t G_t}{G_0} \middle| X_0 \right) + \frac{1}{t} \log \mathbb{E} (S_t | X_0) \end{aligned}$$

where in the first expression the first term is the logarithm of the expected payoff divided by, the second term is minus the logarithm of the price, and the third term is minus the logarithm of the riskless cash-flow return for horizon  $t$ . Scaling by  $1/t$  adjusts for the investment horizon.

The product  $SG$  is itself a multiplicative functional. Let  $\eta^{sg}$  denote its geometric growth component. Then from [\(8.20\)](#), the limiting cash-flow risk compensation is:

$$\eta^g + \eta^s - \eta^{gs}.$$

under the moment restrictions imposed in limit [\(8.17\)](#). This expression resembles the negative of covariance as is often found in asset pricing, but it differs from a covariance because we are working with proportional measures of the risk compensations. When the cash flow process,  $G$ , is a cumulative return, the asset pricing restrictions informs us that  $SG$  is a martingale. In this case, the proportional risk premium is just:

$$\eta^g + \eta^s,$$

where, as we noted previously,  $-\eta^s$  is the long-term counterpart to a riskless rate of return.

## 8.8.2. Finite versus infinite-valued claims

In a recent paper, [\[Panageas, 2025\]](#) reminds us that several important issues that bear on finite values of claims to infinite cash flows, including macro processes like aggregate consumption and aggregate output. These issues include speculative bubbles, dynamic efficiency, and equity claims linked to macroeconomic outcomes. Importantly, he argues that the adjustment for uncertainty can play an important and sometimes subtle role in such evaluations.<sup>[1]</sup> We now explore in more detail when such valuations are finite taking account of uncertainty adjustments.

We start by building stochastic counterparts to the well known and pedagogically revealing Gordon growth model posed in discrete time. By using the framework developed in this chapter, we characterize the uncertainty adjustments necessary for such valuations. In so doing, we show how tools featured in Black-Scholes style option-pricing type analyses must be altered in important ways to understand these valuations.

We build a multiplicative functional using the product,  $M = GS$ , where  $G$  captures stochastic growth in the cash flow and  $S$  stochastic discounting. The factorization in [Theorem 8.1](#) and the selection we featured in [Theorem 8.2](#) provide a defense for using  $\tilde{\eta}$  as an asymptotic growth or decay rate of  $M$  provided that we select the solution that is stochastically stable. To justify finite values of equity-based assets for a family of cash flows, we restrict  $\tilde{\eta} < 0$  so that stochastic discounting dominates stochastic growth. We now revisit that analysis, deconstruct the outcome, and explore the limitations.

### 8.8.2.1. Constant discounting

As a warm-up, consider a discrete-time Gordon growth model where:

$$\begin{aligned} S_t &= \exp(\eta_s t) \\ G_t &= \exp(\eta_g t) G_0 \end{aligned}$$

with  $\eta_s < 0$  reflecting discounting and  $\eta_g$  capturing growth. The well known and simple to compute outcome gives:

$$\sum_{t=0}^{\infty} M_t = \frac{1}{1 - \exp(\eta_s + \eta_g)} G_0,$$

for  $M = SG$ , provided that  $\tilde{\eta} = \eta_s + \eta_g < 0$ . Thus we have the familiar restriction that discounting has to dominate growth in a deterministic setting.

As is also well-known, the separation is not quite so simple because growth contributes to the rate of return used for discounting. The term  $\tilde{\eta}_s$  could reflect both subjective discounting in preferences and an adjustment contributed by economic growth in consumption that can either enhance or diminish the discount rate depending on investor preferences for intertemporal substitution.

### 8.8.2.2. Expected discounted cash flows

Next consider discounted expected values for a  $\tilde{\eta} < 0$  and a stochastically stable Markov process  $X$ . Recall that stochastic stability implies that conditional expectations converge to their unconditional counterpart. Thus we, temporarily, abstract from uncertainty adjustments in valuation not reflected in conditional expectations or encoded in the discount rate. (We will be agnostic for how we come up with  $\tilde{\eta}$ .) Consider the valuation of a positive claim on the Markov state say  $f(X)$ ,

$$\sum_{t=0}^{\infty} \exp(\tilde{\eta} t) \mathbb{E} [f(X_t) \mid X_0 = x]. \quad (8.21)$$

Under the presumed stochastic stability of the Markov process, there are two possibilities are of interest:

$$\begin{aligned} \mathbb{E} [f(X_t) \mid X_0 = x] &\rightarrow \mathbb{E} [f(X_t)] < \infty \\ \mathbb{E} [f(X_t) \mid X_0 = x] &\uparrow \mathbb{E} [f(X_t)] = +\infty. \end{aligned} \quad (8.22)$$

While it is common to feature the finite-moment case, here we proceed more generally. Except for the finite-state Markov case, we may easily construct payoffs that are positive functions of the Markov state that have infinite expectations. When the unconditional expectation is infinite, the value of the equity-type claim may or may not be finite. In effect, there is a race between how fast the expectation diverges versus the magnitude of the discount rate  $-\tilde{\eta}$ . When the former grows faster than  $-\tilde{\eta}$ , the valuation in [\(8.21\)](#) will be infinite, even though the underlying Markov is stochastically stable.

Our more general approach required that we address a multiplicity that emerges when we factor  $M$ . Take  $M_t = \exp(\tilde{\eta}t)\phi(X_t)$ . For sake of discussion, take  $S_t = \exp(\tilde{\eta}t)$ , and  $G_t = \phi(X_t)$ . As a special case for such an  $M$  process, consider the solutions to the following equation:

$$\mathbb{E} [\hat{e}(X_{t+1}) | X_t = x] = \exp(\hat{\eta})\hat{e}(x).$$

Setting  $\hat{e} = 1$  and  $\hat{\eta} = 0$  solves this equation, but there can be other solutions. For such solutions,  $\hat{\eta} > 0$ , the conditional expectations will diverge geometrically, and they must have infinite unconditional moments. More generally, the potential divergence of the conditional moments is precisely the source of multiple factorizations for  $M$ , and what drives our choice of selecting one. The associated cash flow with a date  $t$  contribution,  $\phi(X_t) = \hat{e}(X_t)$ , will have Infinite value when  $\hat{\eta} > -\tilde{\eta}$ .

### Example 8.9

Consider a special case of [Example 8.8](#) with  $\nu = f_1 = f_2 = 0$ . This makes  $M = 1$ . We look for solutions of the form:

$$\hat{e}(x) = \exp\left(\epsilon_1 x + \frac{\epsilon_2}{2} x^2\right).$$

The quadratic equation of interest simplifies to:

$$b^2(\epsilon_2)^2 + (a^2 - 1)\epsilon_2 = 0.$$

This has two solutions. One is  $\epsilon_2 = 0$ . For this one,  $\hat{\eta} = 0$  and  $\hat{e} = 1$  and gives rise to the solution that we expect. For the other one,

$$\epsilon_2 = \frac{1 - a^2}{b^2},$$

$\epsilon_1 = 0$ , and

$$\hat{\eta} = -\frac{1}{2} \log(a^2) > 0.$$

since  $a^2 < 1$ .

## 8.8.2.3. A simple stochastic extension of the Gordon growth

## model

To model stochastic growth and discounting, consider the very special case of two correlated geometric random walks with constant expected discounting or growth rates. Suppose

$$\begin{aligned}\log S_{t+1} - \log S_t &= \mu_s + \sigma_s W_{t+1}, S_0 = 1 \\ \log G_{t+1} - \log G_t &= \mu_g + \sigma_g W_{t+1}.\end{aligned}$$

where  $\mu_s$  and  $\mu_g$  are constants with  $\mu_s < 0$ ,  $\mu_g > 0$ , and  $\sigma_s$  and  $\sigma_g$  are constant row vectors and the  $W$  process is iid with  $W_{t+1}$  being a multivariate standard normally distributed random vector. Form the product process,  $M = SG$ , which will have the same mathematical structure:

$$\log M_{t+1} = \log M_t + \mu_s + \mu_g + (\sigma_s + \sigma_g)W_{t+1}.$$

Construct

$$\begin{aligned}\tilde{\eta} &\stackrel{\text{def}}{=} \mu_s + \mu_g + \frac{|\sigma_s + \sigma_g|^2}{2} \\ &= \left[ \mu_s + \frac{|\sigma_s|^2}{2} \right] + \left[ \mu_g + \frac{|\sigma_g|^2}{2} \right] + \sigma_s \cdot \sigma_g.\end{aligned}$$

The first term in the second line captures the implied risk-free rate of the return, the second term the expected growth rate, and the third term,  $\sigma_s \cdot \sigma_g$ , includes an adjustment for uncertainty. The security with cash flow  $G$  has finite value when  $\tilde{\eta} < 0$  with the familiar geometric summation formula:

$$\sum_{t=0}^{\infty} \mathbb{E}(M_t | G_0, X_0 = x) = \left[ \frac{1}{1 - \exp(\tilde{\eta})} \right] G_0.$$

This computation illustrates a way to build a stochastic version the Gordon growth model, but it has important limitations from both a conceptual and empirical perspective.

### 8.8.2.4. Adjusting for uncertainty II

The stochastic example in the previous subsection has pedagogical value, but the interest rates are constant as are the cash-flow growth rates. The same is true for the one-period cash-flow exposures to the shock vector  $W_{t+1}$ . We escape this straight jacket by introducing state dependence expressed as follows:

$$\begin{aligned}\log S_{t+1} - \log S_t &= \mu_s(X_t) + \sigma_s(X_t)W_{t+1}, S_0 = 1 \\ \log G_{t+1} - \log G_t &= \mu_g(X_t) + \sigma_g(X_t)W_{t+1}.\end{aligned}$$

By using the factorization of [Theorem 8.1](#) and the unique [Theorem 8.2](#) selection of such a representation, we return to our initial present-value analysis in [Subsection 7.8.1.2]:[Expected discounted cash flows](#).

Write

$$\frac{M_t}{M_0} = \exp(\tilde{\eta}t) \tilde{L}_t \left[ \frac{\tilde{e}(X_0)}{\tilde{e}(X_t)} \right]$$

where  $M = SG$ . Consider a family of cash flows of the form

$$G_t f(X_t) \tag{8.23}$$

for different choices of  $f = \tilde{f}\tilde{e}$  and  $\tilde{f} \in \mathcal{F}(\tilde{L})$  given by [\(8.16\)](#). Stated in terms of logarithms, the cash-flows are all co-integrated (with co-integration coefficients one and minus one). The process  $G$  reflects a common stochastic growth component. [\[Hansen, 2012\]](#) explores in detail relations between co-integration as it is studied in macroeconomics and long-term valuation as it is of interest in this chapter. In particular, the set  $\mathcal{F}(\tilde{L})$  imposes moment restriction for the valuation limits. Distinct moment implications are made other applications of co-integration.

The uncertainty-adjusted valuation of the discounted cash flows of the form given in [\(8.23\)](#) is given by:

$$\sum_{t=0}^{\infty} \exp(\tilde{\eta}t) \tilde{\mathbb{E}} \left[ \tilde{f}(X_t) \mid X_0 = x \right] G(X_0) \tilde{e}(X_0)$$

where the expectation uses the probability distribution induced by  $\tilde{E}$ . Analogous to our initial computation, if  $\tilde{\eta} < 0$ , the valuation will be finite. Infinite values will arise whenever  $\tilde{\eta} \geq 0$ . Thus,  $\tilde{\eta}$  as a measure of the asymptotic growth or decay in  $SG$ , plays a central role in determining whether the discounted expected value of the cash-flow  $G$  is finite or not. But like the analysis in [Subsection 7.8.1.2]:[Expected discounted cash flows](#), there is a nuanced case in which

$$\tilde{E} \left[ \tilde{f}(X_t) \right] = +\infty.$$

For this infinite-moment case, whether the discounted expected value is finite or not, depends on the rate of divergence in comparison to  $-\tilde{\eta}$ .

Importantly, this analysis uses a change in probability measure for the computation. The change of measure mathematically resembles a standard risk-neutral probability, but it is conceptually distinct. It depends on the stochastic behavior of  $SG$  and not just  $S$ . Moreover, state-dependent short-term interest

rates alone make the risk-neutral probabilities not the relevant one for valuing long-term cash flows. With state-dependence in stochastic growth, it becomes valuable to move from  $S$  to  $SG$ , which builds on the Gordon-growth insight. The uncertainty adjustments for valuation depend on the implied asymptotic rate of growth or decay,  $-\tilde{\eta}$ , analogous to what happens in deterministic environments. It could, however, also depend on the corresponding behavior of the conditional expectations,

$$\tilde{E} \left[ \tilde{f}(X_t) \mid X_0 = x \right]$$

of the cash flows as the payoff horizon becomes large.

Although stated in the context of a very special example, the analysis in [Panageas, 2025] reminds us that in a stochastic setting, unconditional moments can be infinite even when there is stochastic stability in the Markov process and the horizon-dependent conditional moments are finite. This means the value comparison for stochastic cash-flows with different specifications of  $f$  could be dramatically different because of the tail behavior of the stationary distributions of random variables  $f(X_t)/\tilde{e}(X_t)$  for the different  $f$ 's. More than the rate  $\tilde{\eta}$  may come into play. It would appear to be hard to tell how restrictive the finite moment condition is empirically, but that does not mean we should dismiss it as a possibility. It sometimes matters when we build explicit models of valuation as illustrated by [Panageas, 2025].

### 8.8.3. Partitioning long-term return implications

We study a family of cash-flow and holding period returns where the cash flows are of the form (8.23) under some moment restrictions on  $f(X_t)$ . Specifically, we study limiting returns as the payoff horizon becomes long are the same for alternative choices of  $f > 0$ . [Jiang et al., 2024] use valuation impact of cointegration in the relation between government spending and revenue to assess the implied value of government debt in comparison to actual direct measurements.

#### 8.8.3.1. Cash-flow returns

Consider two factorizations:

$$\begin{aligned} \frac{G_t}{G_0} &= \exp(\eta^g t) L_t^g \frac{e^g(X_0)}{e^g(X_t)} \\ \frac{S_t G_t}{G_0} &= \exp(\eta^{sg} t) L_t^{sg} \frac{e^{sg}(X_0)}{e^{sg}(X_t)} \end{aligned}$$

Given  $G$ , entertain a family of cash flows that can be represented as in (8.23) for  $f > 0$ 's such that

$$f/e^g \in \mathcal{F}(L^g) \quad \text{and} \quad f/e^{sg} \in \mathcal{F}(L^{sg}). \quad (8.24)$$

The two factorizations of interest for cash flows [\(8.23\)](#) are:

$$\begin{aligned} \frac{G_t f(X_t)}{G_0 f(X_0)} &= \exp(\eta^{sg} t) L_t^g \frac{f(X_t) e^g(X_0)}{f(X_0) e^g(X_t)} \\ \frac{S_t G_t f(X_t)}{G_0 f(X_0)} &= \exp(\eta^{sg} t) L_t^{sg} \frac{f(X_t) e^{sg}(X_0)}{f(X_0) e^{sg}(X_t)} \end{aligned}$$

Implicit in these factorizations that the introduction of  $f$ , and no impact on the Perron-Frobenius eigenvalues nor does it alter the multiplicative martingales. It does change the eigenfunction by replacing  $e^g$  with  $e^g/f$ . With these extensions along with moment restrictions [\(8.24\)](#), the limiting expected growth and valuation rates are:

$$\begin{aligned} \lim_{t \rightarrow \infty} \frac{1}{t} \log \mathbb{E} [G_t f(X_t) \mid \mathfrak{A}_0] &= \eta^g \\ \lim_{t \rightarrow \infty} \frac{1}{t} \log \mathbb{E} [S_t G_t f(X_t) \mid \mathfrak{A}_0] &= \eta^{sg}. \end{aligned}$$

Observe that the limits remain the same for alternative choices of  $f > 0$  as is the expected rate of return given by:

$$\eta^g - \eta^{sg}.$$

### 8.8.3.2. One-period holding-period returns

We also investigate the limiting behavior of one-period holding-period returns and obtain an even stronger connection among the family of returns for cash flows satisfying To achieve this, we revisit and extend a computation we performed previously. Consider the multiplicative factorization again and pricing future cash flows in single periods. (These are sometimes referred to as “strips.”) An empirical asset pricing literature has explored these returns starting with [\[van Binsbergen et al., 2012\]](#). See [\[Golez and Jackwerth, 2024\]](#) for a recent update of this evidence.

The one-period holding period return on a  $t$ -period asset with payout  $G_t f(X_t)$  at date  $t$  is:

$$\begin{aligned}
R_1^t[Gf(X)] &= \frac{\tilde{\mathbb{E}} \left[ \frac{f(X_t)}{e^{sg}(X_t)} \mid X_1 \right]}{\tilde{\mathbb{E}} \left[ \frac{f(X_t)}{e^{sg}(X_t)} \mid X_0 \right]} \left[ \frac{G_1 f(X_1)}{G_0 f(X_0)} \right] \left[ \frac{e^{sg}(X_1) f(X_0)}{e^{sg}(X_0) f(X_1)} \right] \left[ \frac{\exp[\eta(t-1)]}{\exp(\eta t)} \right] \\
&= \frac{\tilde{\mathbb{E}} \left[ \frac{f(X_t)}{e^{sg}(X_t)} \mid X_1 \right]}{\tilde{\mathbb{E}} \left[ \frac{f(X_t)}{e^{sg}(X_t)} \mid X_0 \right]} \left[ \frac{G_1 e^{sg}(X_1)}{G_0 e^{sg}(X_0)} \right] \exp(-\eta)
\end{aligned}$$

where the  $\tilde{\mathbb{E}}$  conditional expectation uses the probability measure induced by the positive martingale  $L^{sg}$ . Using stochastic stability and taking limits as  $t$  tends to  $\infty$  gives the limiting holding-period return:

$$R_1^\infty[Gf(X)] = \exp(-\eta^{sg}) \left[ \frac{G_1 e^{sg}(X_1)}{G_0 e^{sg}(X_0)} \right] = \left( \frac{1}{S_1} \right) L_1^{sg} \quad (8.25)$$

provided that  $f/e^{sg} \in \mathcal{F}(L^{sg})$ . The limiting holding-period return does not depend on  $f$ .

Much of this chapter has featured factorizations that allow us to deconstruct economically interesting processes including growth processes. We now consider a way to parametrize cash-flow growth processes in terms of holding period returns and decay rates. We do this to reveal common long-term risk exposures positive cash-flow growth processes. Formally, we build partitions of such processes. For a given a cumulative stochastic discount factor process  $S$  with  $S_0 = 1$ , parameterize a baseline stochastic cash flows as

$$\frac{G_{t+1}}{G_t} = \exp(-\eta) R_{t+1}$$

where

- $\eta$  is a positive decay rate  $\eta$ ;
- $R$  is a one-period return process constructed as

$$\log R_{t+1} = \kappa(X_t, W_{t+1})$$

for some  $\kappa$ ;

- the implied multiplicative martingale with increment

$$\frac{L_{t+1}}{L_t} = \left( \frac{S_{t+1}}{S_t} \right) R_{t+1}$$

and  $L_0 = 1$  induces a stochastically stable process for  $X$ .

Notice that for a given cumulative stochastic discount factor process  $L$  reveals  $R$  and conversely. Since we restrict  $R$  via its implications for the martingale  $L$ , we use the notation  $G(\eta, L)$  to denote the baseline process by  $G(\eta, L)$ . The initial condition  $G_0(\eta, L) > 0$  is inconsequential at this juncture.

For each baseline process, introduce a family of cash-flow growth processes within a partition:

$$\mathcal{G}(\eta, L) \stackrel{\text{def}}{=} \{G(\eta, L)f(X) : f \in \mathcal{F}(L)\}.$$

Notice that for each  $G$  in a partition  $\mathcal{G}(\eta, L)$ , we have the factorization:

$$\frac{S_t G_t}{G_0} = \exp(-\eta t) L_t \left[ \frac{f(X_t)}{f(X_0)} \right]$$

for some  $f \in \mathcal{F}(L)$ . Each such member has the same one-period holding period return,  $R$ .

For a given cash-flow process,  $G$ , reveal the corresponding  $(\eta, L)$  and the relevant partition to which  $G$  belongs. Consider two cash flow processes  $G^1$  and  $G^2$  within the universe of all of the partitions that are related by:

$$\log G^1 - \log G^2 = \log \tilde{f}(X).$$

Thus they are co-integrated in logarithms with vector  $(1, -1)$ . These processes are necessary within the same partition with a common  $(\eta, L)$ .

#### Remark 8.10

[[Ai and Fu, 2025](#)] extend the holding-return limits derived here to study what can be revealed by announcement impacts on financial markets. They use a continuous-time formulation with hidden states and preset announcement dates at which information is revealed about monetary policy. They study instantaneous returns at the times of the announcements. They deduce counterpart representations applicable to such holding-period returns. They use this setup to estimate long-term impacts of money policy announcements on the macro economy as revealed by such returns. Their analysis is facilitated by their use of log-linear models with Brownian motion increments, which opens the door to a tractable link between the multiplicative martingale increments and permanent shocks revealed by additive decompositions applied to logarithms.

## 8.9. Bounding investor beliefs

We use the cumulative stochastic discount factorization to analyze two distance approaches to drawing inferences about investor beliefs.

### 8.9.1. Subjective beliefs in the absence of long-term risk

Suppose that we have data on prices of one-period state-contingent claims. We can use these data to infer the one-period operator,  $\mathbb{M}$ . Recall that we represent this operator using a baseline specification of the one-period transition probabilities. One possibility is that the one-period baseline transition probabilities agree with the data generation. Rational expectations models equate transition probabilities to those used by investors. More generally, investors could have subjective beliefs that can differ from the baseline specification:

- investors think there are no permanent macroeconomic shocks;
- investors don't have risk-based preferences that can induce a multiplicative martingale in a cumulative stochastic discount factor process.<sup>[2]</sup>

Under these two restrictions, we could identify the  $L^s$  as the likelihood ratio for investor beliefs relative to the baseline probability distribution. Thus, the implied martingale component in the cumulative stochastic discount factor identifies the subjective beliefs of investors. Using this change of measure, the limiting long-term risk compensations derived in the previous section are zero. These assumptions facilitate “Ross recovery” of investor beliefs.<sup>[3]</sup>

### 8.9.2. Restricting the martingale increment with limited asset market data

Initially, suppose we impose rational expectations by endowing investors with knowledge of the data generating process. With limited asset market data we cannot identify the martingale component to cumulative stochastic discount factor process without additional model restrictions. We can, however, obtain potentially useful bounds on the martingale increment. For some applications, in addition to those of [\[Alvarez and Jermann, 2005\]](#), [\[Bakshi and Chabi-Yo, 2012\]](#), see [\[Kojien et al., 2010\]](#) and [\[Lustig et al., 2019\]](#), among others. We know that the implied martingale has some peculiar limiting behavior and that martingale and transient components can be correlated. Nevertheless, the implied probability measure can be well behaved. Consequently, in contrast to the references just cited, we use the increment as a device to represent conditional probabilities instead of just as a random variable. Extensions of these same methods can be used to study restrictions on subjective beliefs implied by asset prices.

There is a substantial literature on divergence measures for probability densities. Relative entropy is an important example. More generally, consider a convex function  $\phi$  that is zero when evaluated at one.

Jensen's inequality implies that

$$\mathbb{E} [\phi(N_1) | X_0] \geq 0,$$

and equal to zero when  $N_1$  is one, provided that  $N_1$  is a multiplicative martingale increment (has conditional expectation one). This gives rise to a family of  $\phi$  divergences that can be used to assess departures from baseline probabilities. Relative entropy,  $\phi(n) = n \log n$  is an example that is particularly tractable and has been used often. Both  $n \log n$  and  $-\log n$  can be interpreted as expected log-likelihood ratios.

One way of assessing the magnitude of the martingale increment to the stochastic discount factor is to solve:

### Minimum divergence Problem

$$\min_{N_1 \geq 0} \mathbb{E} [\phi(N_1) | X_0] \tag{8.26}$$

subject to:

$$\begin{aligned} \mathbb{E}(N_1 | X_0) &= 1 \\ \mathbb{E} \left[ N_1 \left( \frac{\widehat{S}_1}{\widehat{S}_0} \right) Y_1 \middle| X_0 \right] &= Q_0 \end{aligned}$$

where  $Y_1$  is a vector of asset payoffs and  $Q_0$  is a vector of corresponding prices, and where  $\widehat{S}_1/\widehat{S}_0$  is equal to one of the two possibilities:

- i)  $\exp(\eta^s) \frac{e^s(X_0)}{e^s(X_1)}$  under the data generating process;

or

- ii) an imposed stochastic discount factor allowing for subjective beliefs to differ from the data generating process.

In regards to i), recall that the term

$$\left[ \exp(\eta^s) \frac{e^s(X_0)}{e^s(X_1)} \right] = (R_1^\infty)^{-1}$$

can be approximated by the reciprocal of the one-period holding-period return on a long-term bond. In regards to ii), for a model that is *misspecified* under the data generating process, we can ask that the

distortion induced by the subjective beliefs correct the misspecification under the data generating process.

#### Remark 8.11

It is enlightening to extend problem (8.26) to include an inequality constraint requiring that the conditional expectation of any random variable of interest be less than or equal to some threshold. Adding a constraint, when it binds, increases the objective. A given increase in the divergence is achieved by reducing this threshold sufficiently far. By following this procedure, we can find a lower bound on the moment of interest that increases the objective by some pre-specified percentage. Extending this approach to any bounded random variable reveals an ambiguity-induced *nonlinear expectation* operator, where the nonlinearity follows from our construction of a lower bounds on conditional expectations. To obtain an *ambiguity interval* for a given random variable follows from also constructing the lower bound for the negative of this random variable. See [Chen et al., 2020] for a more extensive discussion. Such a constructions are an example of what econometricians describe as *partial identification* because the vector  $Y_1$  of asset payoffs used in the analysis may not be sufficient to reconstruct all potential one-period asset payoffs and prices. In effect, the econometrician is often confronted with incomplete data on financial markets.

#### Remark 8.12

Many researchers use  $\phi(n) = -\log n$  as a divergence measure. [Chauduri et al., 2023] and [Chen et al., 2024], however, isolate a potentially problematic aspect of monotone decreasing divergences such as  $-\log n$  because they can fail to detect certain limiting forms of deviations from baseline probabilities.

#### Remark 8.13

[Ghosh and Roussellet, 2023] suggest a formal econometric procedure for estimating the minimum divergence altered probability measure using conditioning information. Presumably their methods could be extended to produce conditional ambiguity intervals as well.

#### Remark 8.14

To avoid having to estimate conditional expectations, applications often study unconditional expectations. In such situations, conditioning can be brought in through the “back door” by scaling payoffs and prices with variables in the conditioning information set; for example, see [Hansen and Singleton, 1982] and [Hansen and Richard, 1987]. See [Bakshi and Chabi-Yo, 2012] and [Bakshi et al., 2017] for some related implementations.

### Remark 8.15

[[Alvarez and Jermann, 2005](#)] use  $-\mathbb{E}(\log S_1 + \log S_0)$  as the objective to be minimized.

Notice that

$$\log N_1^s = \log S_1 - \log S_0 + [\eta^s + \log e^s(X_1) - \log e^s(X_0)],$$

where the term in square brackets is the logarithm of the reciprocal of the limiting holding-period bond return. The criterion thus equals that in the minimum divergence problem, but with an additive translation. Rewrite the constraints in the minimum divergence problem as:

$$\begin{aligned}\mathbb{E}\left(\frac{S_1}{S_0} R_1^\infty\right) &= 1 \\ \mathbb{E}\left(\frac{S_1}{S_0} Y_1\right) &= Q_0.\end{aligned}$$

This leaves us with an equivalent minimization problem in which the translation term is subtracted off to obtain the bound of interest.

Applied researchers have sometimes omitted the first constraint, which weakens the bound.

## 8.9.2.1. Intertemporal divergence measures

[[Chen et al., 2020](#)] propose extensions of the one-period divergence measures to multi-period counterparts that remain tractable and enlightening. By looking over time, the dynamic formulation essentially averages over the conditioning information; but it allows for subjectivity in the transition probabilities. Their dynamic measure of divergence is constructed as follows. For a given  $N$  process, let

$$L_T = \prod_{t=1}^T N_\tau$$

As we have seen, this construction provides the implied transition probability adjustments for the alternative time horizons. Find a function  $\psi$  such that

$$n\psi\left(\frac{1}{n}\right) = \phi(n).$$

With these inputs form an alternative equivalent way to represent the one-period divergence by constructing a function  $\psi$  such that

$$n\psi\left(\frac{1}{n}\right) = \phi(n).$$

The function  $\psi$  plays a role analogous to  $\phi$  when the roles between the  $N$ -implied probability and the baseline probability are interchanged. It may be shown that  $\psi$  is also strictly convex. Note also that  $\psi(1) = \phi(1) = 1$ . By design,

$$\mathbb{E}[\phi(N_t) \mid \mathfrak{A}_{t-1}] = \mathbb{E}[N_t\psi(N_t) \mid \mathfrak{A}_{t-1}]$$

With these building blocks, [\[Chen et al., 2020\]](#) suggest the divergence measure:

$$\lim_{T \rightarrow \infty} \frac{1}{T} \sum_{t=1}^T \mathbb{E}[L_{t-1}\phi(N_t) \mid \mathfrak{A}_0] = \lim_{T \rightarrow \infty} \frac{1}{T} \sum_{t=1}^T \mathbb{E}\left[L_t\psi\left(\frac{1}{N_t}\right) \mid \mathfrak{A}_0\right].$$

An alternative equivalent representation may be obtained by taking a Law of Large Numbers limit under the altered probability measure. When  $\phi(n) = n \log n$ , this divergence measure coincides with the one that is pertinent for the study of Donsker-Varadhan formulations of large deviations in Markov environments. This dynamic formulation of divergences opens the door to recursions for the bounds in the divergence minimization described in [Remark 8.11](#).

### 8.9.2.2. Subjective belief bounds on proportional risk premia

The next two figures report proportional risk compensations for the market return using an illustration from [\[Chen et al., 2020\]](#) under a presumed data generation and when distorted investor beliefs are permitted and risk aversion is constrained to be one. The compensations condition on dividend/price ratios. The figures include upper and lower endpoints of conditional expectations computed by extending the divergence minimization problem along the lines suggested [Remark 8.11](#) given by the upper and lower edges of shaded rectangles. The horizontal lines within the shaded rectangles give the conditional moments implied by the minimum divergence change in probability measure. By comparing the two figures, we see how ambiguity intervals decrease when we weaken the constraint on the divergence. Since the divergence measure used in the computation is explicitly dynamic, the associated subjective probabilities for the transition probabilities are also substantially altered. The implied stationary probabilities for the three states are

high D/P	middle D/P	low D/P	
.42	.31	.27	baseline probability
.76	.20	.204	minimum entropy probability

Notice that the altered probabilities reflect a substantial reduction of being in the high dividend-price state, making the big conditional divergence for the high dividend-price state seem to have a smaller effect on

the dynamic divergence measure.

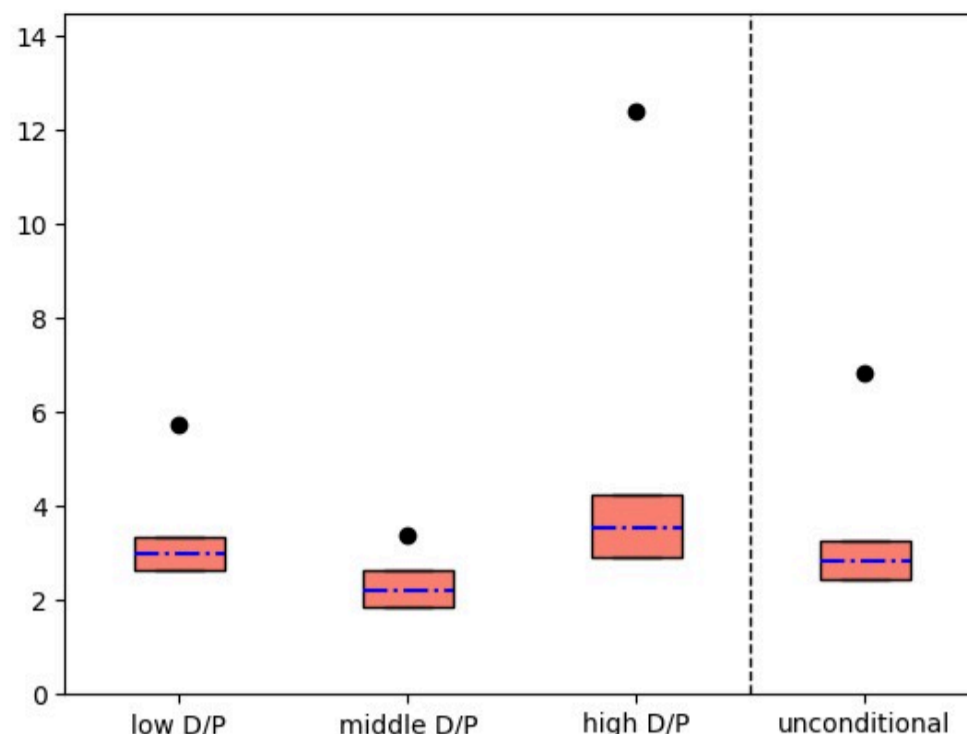


Fig. 8.3 Proportional risk compensations computed as  $\log \mathbb{E}R^m - \log \mathbb{E}R^f$  scaled to annualized percentages. The ●s are the empirical averages and the boxes give the imputed bounds when we inflated the minimum relative entropy by 20%.

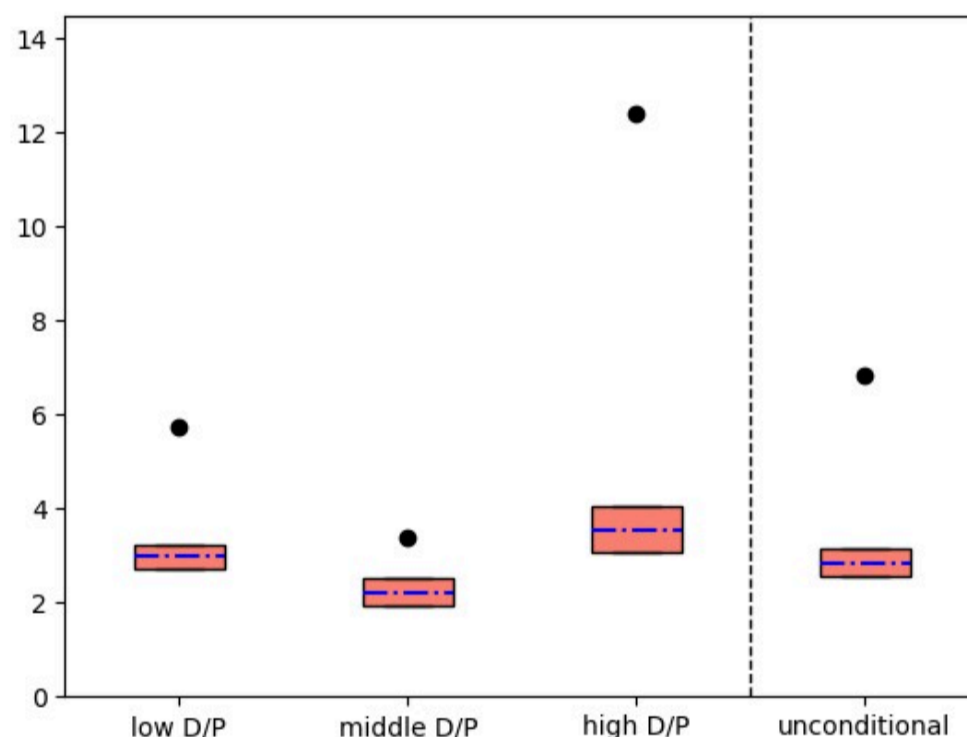


Fig. 8.4 Proportional risk compensations computed as  $\log \mathbb{E}R^m - \log \mathbb{E}R^f$  scaled to an annualized percentage. The ●s are the empirical averages and the boxes give the imputed bounds when we inflated the minimum relative entropy by 10%.

### 8.9.2.3. Subjective belief factor

[Cui et al., 2025] use a complementary approach to study subjective beliefs and asset prices. They use a convenient mathematical representation of linear functionals. Consistent with our previous discussion, distorted expectations are viewed as a (conditional) linear mapping from the random variables being forecast to associated conditional predictions. A belief factor is *identified* as the random variable that

appears in this representation of a conditional linear functional in a mathematical formulation in [\[Hansen and Richard, 1987\]](#) and [\[Cerrei-Vioglio et al., 2016\]](#). [\[Cui et al., 2025\]](#) use this belief factor to organize useful descriptions of a various data sets containing reports of subjective beliefs. In this setting, the subjective belief factor is restricted to be a (conditional) linear combination of the variables being forecast in the data set of interest, a restriction central to econometric identification. As [\[Cui et al., 2025\]](#) point out, extending the subjective belief functional to include forecasts of other variables omitted from a data set under study generates subjective belief factors whose best linear (conditional) predictors equal the subjective belief factor that they construct from their smaller data set. While identification remains partial in this, their framework has provides convenient characterizations of what can be learned and why it is informative.

[\[Cui et al., 2025\]](#) correctly remind us that they do not restrict the subjective belief factor to be positive. They go on to suggest a way to impose positivity under some particular distributional assumptions. As we have seen, by using statistical divergences, the [\[Chen et al., 2020\]](#) analysis avoids these auxiliary assumptions while still imposing positivity. It also provides an entirely different characterization of partial identification than does the approach of [\[Cui et al., 2025\]](#). There are advantages and disadvantages to both approaches.

- 
- [1]** On the third point, [\[Panageas, 2025\]](#) describes an interesting historical episode.
  - [2]** This second restriction is violated by recursive utility models of investor preferences that we study in [Chapter 10:Using Recursive Utility to Assess Uncertainty](#). Also, see discussions in [\[Alvarez and Jermann, 2005\]](#), [\[Hansen and Scheinkman, 2009\]](#), and [\[Borovička et al., 2016\]](#).
  - [3]** Our presentation is consistent with the formal analysis in [\[Ross, 2015\]](#). Ross motivates and conducts his analysis somewhat differently.

