Chapter 4

Additive Functionals

Many interesting economic time series display persistent growth that renders the assumptions of stationarity and ergodicity untenable. But stationarity and ergodicity are the theoretical foundations that activate the Laws of Large Numbers and Central Limit theorems that make statistical learning possible. In this chapter and in chapters 5, 6, and 7, we describe alternative sets of assumptions that are sufficient to render particular components of growing time series stationary and ergodic in ways that enable statistical learning. In this chapter, we describe processes that grow but have *Markov increments*, while in chapter 5 we describe a generaliztion, namely, processes with *stationary increments*. Chapter 6 describes transformations designed to render nonstationary processes stationary and ergodic. Chapter 7 describes processes whose logarithms are processes with Markov increments. All four chapters describe decompositions of a time series into stationary and nonstationary components that allow us to apply Laws of Large Numbers and Central Limit approximations to stationary components.

In this chapter, we use a stationary Markov process to construct a process that displays stochastic arithmetic growth. We show how to extract a linear time trend and a martingale from that growing process. In chapter 7, we use findings of this chapter to model a process that displays geometric growth.

4.1 Definition

A k-dimensional stochastic process $\{W_{t+1} : t = 0, 1, ...\}$ is a set of unanticipated economic shocks. Let $\{X_t : t = 0, 1, ...\} = \{X_t\}$ be a discrete-time stationary Markov process that is generated by initial distribution Q for X_0 and transition equation

$$X_{t+1} = \phi(X_t, W_{t+1}),$$

where ϕ is a Borel measurable function. Let $\{\mathfrak{F}_t : t = 0, 1, ...\}$ be the filtration generated by histories of W and X; $\mathfrak{F}_t$ serves as the information set (sigma algebra) generated by $X_0, W_1, \ldots, W_t$. To insist that the process $\{W_{t+1}\}$ represents unanticipated shocks, we restrict it to satisfy

$$E\left(W_{t+1} | \mathfrak{F}_t\right) = 0.$$

We condition on a statistical model in the sense of section 2.6 and assume that the X_t process is ergodic.[1] The Markov structure of $\{X_t\}$ makes the distribution of (X_{t+1}, W_{t+1}) conditioned on $\mathfrak{F}_t$ depend only on X_t. Like $\{X_t\}$, the pair $\{(X_t, W_t)\}$ is a first-order Markov process. Because the shock W_{t+1} is unpredictable and X_t is the only relevant state vector for (X_{t+1}, W_{t+1}), the composite system $\{(X_t, W_t)\}$ has a triangular structure in the sense that it can be expressed as

$$\begin{aligned} X_{t+1} &= \phi(X_t, W_{t+1}) \\ W_{t+1} &= W_{t+1} \end{aligned}$$

Definition 4.1.1. *A process $\{Y_t\}$ is said to be an **additive functional** if it can be represented as*

$$Y_{t+1} - Y_t = \kappa(X_t, W_{t+1}) \tag{4.1}$$

for a (Borel measurable) function $\kappa : \mathbb{R}^n \times \mathbb{R}^k \to \mathbb{R}$, or equivalently

$$Y_t = Y_0 + \sum_{j=1}^{t} \kappa(X_{j-1}, W_j),$$

where we initialize Y_0 at some arbitrary (Borel measurable) function of X_0.

[1] In the spirit of chapter 2, if we want to acknowledge model uncertainty, we can apply the construction offered here to each of element of a set of statistical models and then form weighted averages over that set of models.

We make Y_0 a function of X_0 because this allows us to construct Y_t as a function of the underlying Markov process between dates zero and t. Using a more general initial condition would have straightforward consequences for results to be stated in this chapter.[2]

Definition 4.1.2. *An additive functional* $\{Y_t : t = 0, 1, ...\}$ *is said to be an* ***additive martingale*** *if* $E\left[\kappa(X_t, W_{t+1})|X_t\right] = 0$.

Example 4.1.3. *Suppose*

$$\begin{aligned} Y_{t+1} - Y_t &= \mu(X_t) + \sigma(X_t)W_{t+1} \\ X_{t+1} &= AX_t + BW_{t+1} \end{aligned}$$

where $\{W_{t+1}\}$ *is an iid sequence of multivariate normally distributed random vectors,* A *is a stable matrix, and* B *has full column rank. Here* $\mu(X_t)$ *is the conditional mean of* $Y_{t+1} - Y_t$ *and* $|\sigma(X_t)|^2$ *is its conditional variance. This is called a stochastic volatility model because* $|\sigma(X_t)|^2$ *is a stochastic process.*

When the conditional mean $\mu(X_t) = 0$ in example (4.1.3), the process $\{Y_t\}$ becomes a martingale. Note that $E\left[\kappa(X_t, W_{t+1})|X_t\right] = 0$ implies the usual martingale restriction

$$E\left[Y_{t+1}|\mathfrak{F}_t\right] = Y_t, \quad \text{for} \quad t = 0, 1, ...$$

4.2 Extracting Martingales

An additive functional can be decomposed into a sum of components, one of which is an additive martingale that summarizes all long-run stochastic variation in the additive functional. In this section, we show how to extract the martingale component of an additive functional. We begin with an algorithm that applies to a special type of additive functional for which $\kappa(x, w^*) = f(x)$ and $\int f(x)Q(dx) = 0$. This algorithm will be the key tool in extracting martingales from additive functionals with general $\kappa(x, w^*)$ functions.

[2]There is great flexibility in initializing the process. We refer to an outcome of this more general construction as an *additive process* instead of an *additive functional*.

Algorithm 4.2.1. *Suppose that* $f \in \mathcal{N}$ *and*

$$Y_{t+1} - Y_t = f(X_t).$$

Solve $g(x) = f(x) + \mathbb{T}g(x)$ *for* $g(x)$ *to get*

$$g(x) = (\mathbb{I} - \mathbb{T})^{-1} f(x) = \sum_{j=0}^{\infty} \mathbb{T}^j f(x) = \sum_{j=0}^{\infty} E\left[f(X_{t+j}) | X_t = x \right], \qquad (4.2)$$

a legitimate calculation provided that the infinite sum on the right side of (4.2) *is finite. A sufficient condition for the sum on the right side of* (4.2) *to be finite is that* $\mathbb{T}^m$ *is a strong contraction for some integer* $m \geq 1$. *The function* g *is the best forecast of the long-term limit of the additive functional* $\{Y_t : t = 0, 1, ...\}$ *as a function of the current Markov state. Where* $(^*)$ *denotes a next period value, let*

$$\kappa_1(x, w^*) \doteq g[\phi(x, w^*)] - g(x) + f(x)$$

and note that $(\mathbb{I} - \mathbb{T})g(x) = f(x)$ *implies that*

$$\kappa_1(x, w^*) = g(x^*) - \mathbb{T}g(x).$$

Thus, $\kappa_1(X_t, W_{t+1})$ *is the error in forecasting* $g(X_{t+1})$ *given* X_t, *so*

$$E\left[\kappa_1(X_t, W_{t+1}) | X_t \right] = 0.$$

Therefore,

$$\begin{aligned} Y_t &= \sum_{j=1}^{t} f(X_{j-1}) + Y_0 \\ &= \sum_{j=1}^{t} \kappa_1(X_{j-1}, W_j) - g(X_t) + g(X_0) + Y_0. \end{aligned}$$

The component $\sum_{j=1}^{t} \kappa_1(X_{j-1}, W_j)$ *is a martingale, while* $-g(X_t)$ *is stationary and* $g(X_0) + Y_0$ *is constant.*

Algorithm 4.2.1 is an instance of a more general construction of Gordin (1969) that we have obtained by specializing to a Markov process for X_t.[3] We use algorithm 4.2.1 as a component of the following algorithm for extracting a martingale from an additive functional.

[3] Also see Hall and Heyde (1980).

Algorithm 4.2.2. *Let $\{X_t\}$ be a stationary, ergodic Markov process. Let $\{Y_t\}$ be an additive functional. Perform the following steps.*

(i) Compute the conditional expectation $E[\kappa(X_t, W_{t+1})|X_t = x] = \bar{f}(x)$ of $Y_{t+1} - Y_t$ and form the deviation of $\kappa(X_t, W_{t+1})$ from its conditional mean

$$\kappa_2(X_t, W_{t+1}) = \kappa(X_t, W_{t+1}) - \bar{f}(X_t).$$

Note that $E[\kappa_2(X_t, W_{t+1})|X_t = x] = 0$.

(ii) Let ν be the unconditional mean $\nu \doteq \int \bar{f}(x)Q(dx)$. Let $f(x)$ be the deviation of the conditional mean of $Y_{t+1} - Y_t$, $\bar{f}(x)$ from its unconditional mean ν: $f(x) = \bar{f}(x) - \nu$. By construction $f \in \mathcal{N}$. Take f as defined here to be the f that appears in algorithm 4.2.1 and form g and κ_1 in the algorithm 4.2.1 decomposition $f(x) = \kappa_1(x, w^*) - g(x^*) + g(x)$.

(iii) Note that

$$\begin{aligned} \kappa(x, w^*) &= \kappa_2(x, w^*) + \bar{f}(x) \\ &= \kappa_2(x, w^*) + f(x) + \nu \\ &= \kappa_2(x, w^*) + \kappa_1(x, w^*) - g[\phi(x, w^*)] + g(x) + \nu. \end{aligned}$$

(iv) It follows that

$$Y_t = t\nu + \left[\sum_{j=1}^{t} \kappa_a(X_{j-1}, W_j)\right] - g(X_t) + g(X_0) + Y_0 \tag{4.3}$$

where $\kappa_a(x, w^*) = \kappa_1(x, w^*) + \kappa_2(x, w^*)$ and $E[\kappa_a(X_j, W_{j+1})|X_j] = 0$.

Via algorithm 4.2.2, we have established[4]

Proposition 4.2.3. *Suppose that $\{Y_t\}$ is an additive functional, that $\mathbb{T}^m$ is a strong contraction on $\mathcal{N}$ for some m, and that $E[\kappa(X_t, W_{t+1})^2] < \infty$. Then*

$$Y_t = \underbrace{t\nu}_{\textbf{trend}} + \underbrace{\sum_{j=1}^{t} \kappa_a(X_{j-1}, W_j)}_{\textbf{martingale}} - \underbrace{g(X_t)}_{\textbf{stationary}} + \underbrace{Y_0 + g(X_0)}_{\textbf{invariant}}.$$

[4] Proposition 4.2.3 can be viewed as a special case of proposition 5.1.2, which assumes stationary rather than Markov increments.

Each term in the decomposition is itself an additive functional. The first term is a linear time trend, the second an additive martingale, the third a stationary process with mean zero, and the fourth a time-invariant constant. If $Y_0 = -g(X_0)$, then the fourth term is zero.

We use the decomposition in Proposition 4.2.3 as a way to identify a linear time trend and "permanent shock" associated with an additive functional. The permanent shock is the increment to the martingale. There are multiple ways to construct transitory components, some of which yield transitory shocks that are correlated with permanent shocks.

Application to a VAR

In this subsection, we apply the four-step construction in algorithm 4.2.2 to an example in which the Markov state $\{X_t\}$ follows a first-order VAR

$$X_{t+1} = AX_t + BW_{t+1}, \tag{4.4}$$

where A is a stable matrix and $\{W_{t+1}\}$ is a sequence of independent and identically normally distributed random variables with mean zero and covariance matrix I. Thus, the one-step ahead conditional covariance matrix of the innovations BW_{t+1} to X_{t+1} equals BB'. Let

$$Y_{t+1} - Y_t = \kappa(X_t, W_{t+1}) = D \cdot X_t + \nu + F \cdot W_{t+1}, \tag{4.5}$$

where D and F are vectors with the same dimensions as X_t and W_{t+1}, respectively. For this example, the four steps of algorithm 4.2.2 are:

(i) Form the conditional growth rate

$$\bar{f}(x) = D \cdot x + \nu$$

and the deviation

$$\kappa_2(X_t, W_{t+1}) = F \cdot W_{t+1}.$$

(ii) Compute $f(x)$ by subtracting from $\bar{f}(x)$ its unconditional mean:

$$f(x) = D \cdot x + \nu - \nu = D \cdot x.$$

Here we are using the fact that the unconditional mean of X is zero because A is a stable matrix.

(iii) Form

$$\begin{aligned}\kappa_1(x, w^*) &= f(x) + g(x^*) - g(x) \\ &= D \cdot x + D'(I-A)^{-1}(Ax + Bw^*) - D'(I-A)^{-1}x \\ &= [B'(I-A')^{-1}D] \cdot w^*,\end{aligned}$$

where $g(x) \doteq (\mathbb{I} - \mathbb{T})^{-1} f(x) = D'(I-A)^{-1}x$.

(iv) It follows that $\kappa_a = \kappa_1 + \kappa_2$ is

$$\kappa_a(X_t, W_{t+1}) = [F + B'(I-A')^{-1}D] \cdot W_{t+1}. \tag{4.6}$$

We shall use formula (4.6) when we construct examples below.

Example 4.2.4. *Beveridge and Nelson (1981) decomposed a univariate time series Y_t into permanent and transitory components.*[5] *In terms of our notation, Beveridge and Nelson let a univariate $\{W_{t+1}\}$ process drive a serially correlated univariate process that we can map into a first-order vector process X_t. For Beveridge and Nelson (1981), $[F + B'(I-A')^{-1}D] \cdot W_{t+1}$ is the permanent shock in a proposition 4.2.3 decomposition of a univariate time series into permanent and transitory components. Because $\{W_{t+1}\}$ is a univariate process, permanent and transitory shocks are necessarily perfectly correlated.*

4.3 Examples of Additive Functionals

Example 4.3.1. *(Long-term risk)*

Let C denote consumption. The logarithm of consumption evolves as

$$\log C_{t+1} - \log C_t = \nu + X_t + F \cdot W_{t+1}$$

where

$$X_{t+1} = AX_t + BW_{t+1},$$

$|A| < 1$ *is a scalar, the process $\{X_t\}$ is univariate, and the i.i.d. $\mathcal{N}(0, I)$ shock vector W_{t+1} is 2×1. The 2×1 vectors F and B are such that one component of W_{t+1} disturbs consumption growth directly, while the other*

[5] We can regard them as seeking to generalize the model studied by Muth (1960).

component disturbs the Markov state X_t. The $j-$step ahead conditional mean of $\log C_{t+1} - \log C_t$ is $\mu + A^j X_t$, so the Markov state X_t contributes a predictable component to consumption growth. When A is close to 1, there is said to be substantial "long-run risk" in consumption. The impulse response function of $\log C_{t+j+1} - \log C_{t+j}$ to the shock vector W_{t+1} is an infinite sequence

$$F, B', B'A, B'A^2, \cdots . \tag{4.7}$$

It is mathematically convenient to represent this sequence by constructing a function with elements of the sequence as coefficients of a power series. This function is called a z transform and provides us with a useful bookkeeping device:

$$F + \sum_{j=1}^{\infty} B'A^{j-1} z^j = F + zB'(I - Az)^{-1},$$

where z is a complex valued scalar satisfying $|z| \leq 1$. The impulse response of $\log C_{t+1}$ cumulates impulse responses of $\log C_{t+1} - \log C_t$, so its z transform is

$$\left(\frac{1}{1-z} \right) \left[F + zB'(I - Az)^{-1} \right],$$

which is well defined as a power series for $|z| < 1$. It is also well defined as a function of z for $|z| \leq 1$, except when $z = 1$.[6] *Division by $1 - z$ in effect accumulates the impulse responses of the first difference of log consumption.*

The increment to the martingale component in a proposition 4.2.3 decomposition of $\log C_t$ scaled to have a unit standard deviation is evidently $F^* \cdot W_{t+1}$, where

$$F^* = \frac{1}{|F + B'(I - A)^{-1}|} [F + B'(I - A)^{-1}].$$

We call $F^* \cdot W_{t+1}$ a permanent shock *and can calculate the impulse response of $\log C_t$ to it. The permanent shock is a linear combination of components of a bivariate impulse response function, with a z-transform that is the same linear combination of z transforms of the two components:*

$$\left(\frac{1}{1-z} \right) F^* \cdot [F + zB'(I - zA)^{-1}].$$

[6] Formally, the transform has a pole at $z = 1$.

4.4 Cointegration

Remark 4.4.1. *A linear combination of two additive functionals is an additive functional.*[7] *Specifically, let* X_t *be governed by the vector autoregression* (4.4) *and let* $\tilde{\kappa}_1(x, w^*)$ *and* $\tilde{\kappa}_2(x, w^*)$ *be two functions that can play the role of* $\kappa(x, w^*)$ *in constructing additive functionals. For real valued scalars* $\mathbf{r}_1$ *and* $\mathbf{r}_2$*, form*

$$Y_t = \mathbf{r}_1 Y_t^{[1]} + \mathbf{r}_2 Y_t^{[2]}$$

where $Y_t^{[1]}$ *is built with* $\tilde{\kappa}_1$ *and* $Y_t^{[2]}$ *is built with* $\tilde{\kappa}_2$*. Thus, we can build*

$$\begin{aligned} Y_t = \mathbf{r}_1 Y_t^{[1]} + \mathbf{r}_2 Y_t^{[2]} &= \sum_{j=1}^{t} \left[\mathbf{r}_1 \tilde{\kappa}_1(X_{j-1}, W_j) + \mathbf{r}_2 \tilde{\kappa}_2(X_{j-1}, W_j) \right] \\ &+ \mathbf{r}_1 Y_0^{[1]} + \mathbf{r}_2 Y_0^{[2]}. \end{aligned}$$

The Proposition 4.2.3 martingale component of $\{Y_t : t = 0, 1, ...\}$ *is the corresponding linear combination of the martingale components of* $\{Y_t^{[1]} : t = 0, 1, ...\}$ *and* $\{Y_t^{[2]} : t = 0, 1, ...\}$*. The Proposition 4.2.3 trend component of* $\{Y_t : t = 0, 1, ...\}$ *is the corresponding linear combination of the trend components of* $\{Y_t^{[1]} : t = 0, 1, ...\}$ *and* $\{Y_t^{[2]} : t = 0, 1, ...\}$*.*

Engle and Granger (1987) focused on a special set of linear combinations of two additive functionals whose linear trend and martingale components are both zero. Engle and Granger call two processes *cointegrated* if there exists a linear combination of them that is stationary,[8] which is true when there exist real valued scalars $\mathbf{r}_1$ and $\mathbf{r}_2$ such that

$$\begin{aligned} \mathbf{r}_1 \nu_1 + \mathbf{r}_2 \nu_2 &= 0 \\ \mathbf{r}_1 \kappa_{a1} + \mathbf{r}_2 \kappa_{a2} &= 0, \end{aligned}$$

where the ν's and κ_a's correspond to the first two components of the representation in Proposition 4.2.3. These two zero restrictions imply that the time trend and the martingale component for the linear combination Y_t are

[7] An analogous statement applies to *additive processes*.

[8] Their definition can readily be extended to require only that the linear combination be asymptotically stationary. That would allow transients in the cointegrating residual ignited by initial conditions X_0 far in the tails of the stationary distribution of X.

both zero.[9] When $\mathbf{r}_1 = 1$ and $\mathbf{r}_2 = -1$, the component additive functionals $Y_t^{[1]}$ and $Y_t^{[2]}$ share a common growth component.

Example 4.4.2. *(hyperinflation)*

Sargent (1977) constructed a model of hyperinflation in which Cagan's adaptive expectations model for inflation is an implication of rational expectations. Let $q_t = p_t - p_{t-1}$ *and* $k_t = m_t - m_{t-1}$ *where* p_t *is the log of the price level and* m_t *is the log of money supply. Let* $\bar{q}_t$ *be the public's time* t *forecast of time* $t+1$ *inflation. Cagan's adaptive expectations rule for forecasting inflation is*

$$\bar{q}_t = \lambda \bar{q}_{t-1} + (1-\lambda) q_t$$

for $0 < \lambda < 1$*. To make this adaptive expectations rule optimal, we assume that*

$$q_{t+1} = \bar{q}_t + \sigma_f \cdot W_{t+1},$$

which implies that

$$q_{t+1} = \lambda q_t + (1-\lambda) q_t + \sigma_f \cdot W_{t+1} - \lambda \sigma_f \cdot W_t$$

or

$$q_{t+1} - q_t = \sigma_f \cdot W_{t+1} - \lambda \sigma_f \cdot W_t.$$

The demand for real balances is

$$\begin{aligned} q_t - k_t &= \alpha \bar{q}_t + \sigma_d \cdot W_t \\ &= \alpha q_t - \lambda \sigma_f \cdot W_t + \sigma_d \cdot W_t, \end{aligned}$$

where $\alpha > 0$ *and* $-\alpha$ *is the semi elasticity of the demand for real balances with respect to expected inflation. So*

$$k_t = (1-\alpha) q_t + \lambda \sigma_f \cdot W_t - \sigma_d \cdot W_t.$$

From this relation, it follows that the martingale components of $\{k_t\}$ *and* $\{q_t\}$ *are proportional and that* $\begin{bmatrix} -1 & (\alpha - 1) \end{bmatrix}$ *is a cointegrating vector for* $\{(k_t, q_t)\}$*.*

[9] The cointegration vector $(\mathbf{r}_1, \mathbf{r}_2)$ is determined only up to scale.

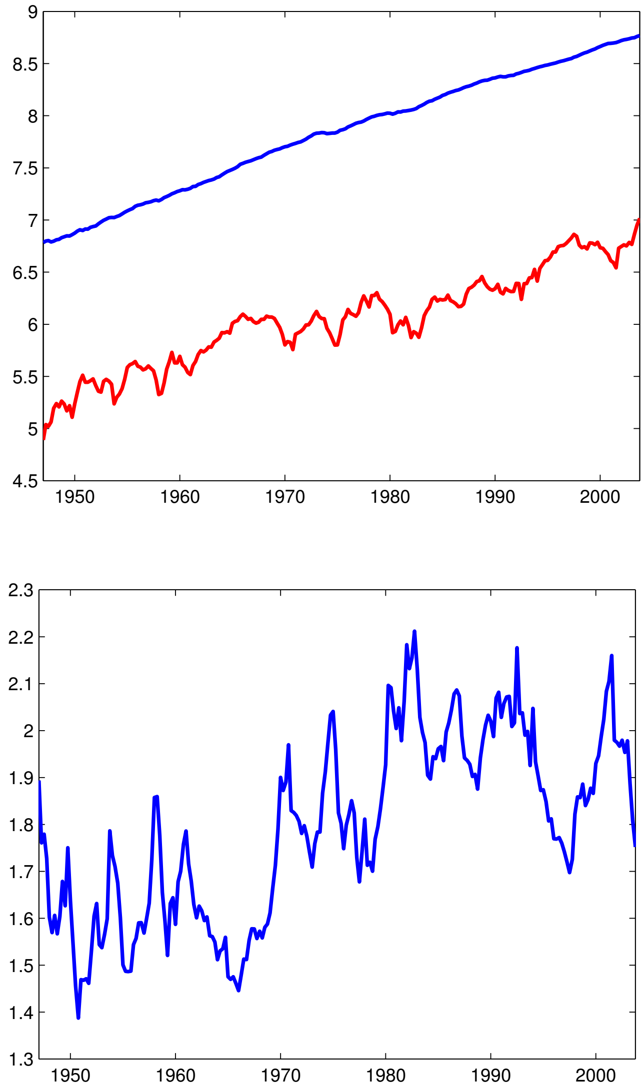

Figure 41: The top panel plots the logarithm of consumption (smooth blue series) and logarithm of corporate earnings (choppy red series). The bottom panel plots the difference in the logarithms of consumption and corporate earnings.

Example 4.4.3. *(Long-term consumption risk, II)*

Hansen et al. (2008) had the idea of using covariation with other time series to help infer long-term stochastic components of consumption. Figure 41 plots logarithms of nondurable consumption C_t and corporate earnings N_t. The absence of an obvious trend or martingale in the second panel, which plots the difference between the logarithms of nondurable consumption and corporate earnings, suggests the presence of common trend and martingale components in the two series themselves, an observation that led Hansen et al. to impose co-integration between the logarithms of consumption and corporate earnings and thereby restrict them to grow together. A way to impose the sought after co-integration is to let X_t be governed by the VAR

$$X_{t+1} = AX_t + BW_{t+1},$$

where A is a stable matrix and $\{W_{t+1}\}$ is an i.i.d. sequence of $\mathcal{N}(0, I)$ random vectors; then to choose X_t to have the growth rate of consumption (expressed in logarithms) as its first entry and the logarithm of corporate earnings minus the logarithm of consumption in the second position, then to fill in the remaining components of X_t with lags of these and any other variables that help forecast the logarithms of corporate earnings and consumption. This specification leaves us with two additive functionals with increments:

$$\begin{aligned} \log Y_{t+1}^{[1]} - \log Y_t^{[1]} &= \nu_1 + X_{t+1}^{[1]} \\ \log Y_{t+1}^{[2]} - \log Y_t^{[2]} &= \nu_2 + X_{t+1}^{[2]} - X_t^{[2]} + X_{t+1}^{[1]}, \end{aligned}$$

where $Y_{t+1}^{[1]} = C_{t+1}$ and $Y_{t+1}^{[2]} = N_{t+1}$. The two additive functionals $\{\log Y_{t+1}^{[1]}\}$ and $\{\log Y_{t+1}^{[2]}\}$ share the same martingale and trend components but have different transitory components.

Notice that

$$\log Y_{t+1}^{[1]} - \log Y_t^{[1]} = \nu_1 + D \cdot X_t + F \cdot W_{t+1}$$

where

$$D = A'U_1, \quad F = B'U_1$$

and U_1 is a vector of zeros except for a one in the first position. The impulse response vector of $\log C_{t+1} - \log C_t$ to W_{t+1} is

$$F, B'D, B'A'D, \cdots,$$

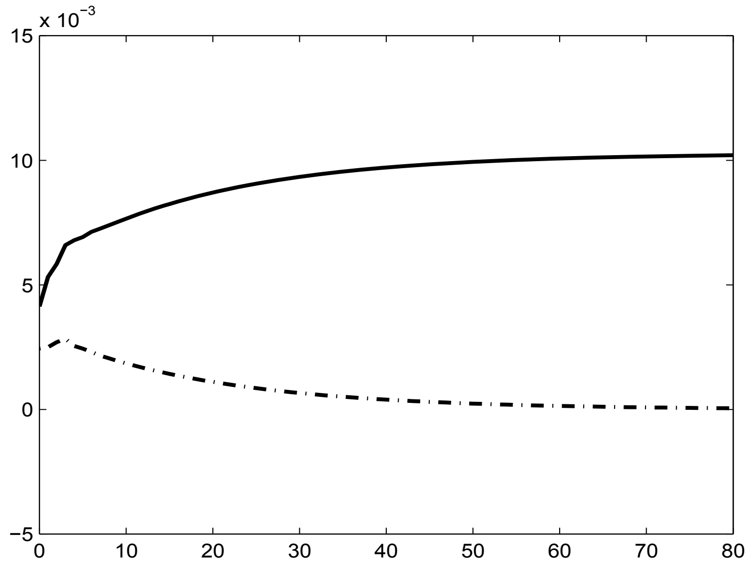

Figure 42: This figure plots the response of the logarithm of consumption to a permanent shock (solid) and and to a temporary shock (dashed line). The permanent shock is identified as the increment to the common martingale component of the logarithm of consumption and the logarithm of corporate earnings. The figure comes from Hansen et al. (2008).

which has z-transform:

$$F + \zeta B'(I - \zeta A')^{-1} D = B'(I - \zeta A')^{-1} U_1.$$

The martingale increment scaled to have unit standard deviation is $F^* \cdot W_{t+1}$*, where*

$$F^* = \frac{1}{|B'(I - A')^{-1} U_1|} B'(I - A')^{-1} U_1$$

and the z-transform of the impulse response function of $\log C_{t+1}$ *to the martingale increment is*

$$\left(\frac{1}{1-\zeta}\right) F^* \cdot [B'(I - \zeta A')^{-1} U_1].$$

Figure 42 shows how the logarithm of aggregate consumption responds to a shock identified as the common martingale component of the logarithm *of consumption and the logarithm of corporate earnings. The immediate response is less than half the long-term response. The long-term response equals the magnitude of the common martingale component of the logarithms of consumption and corporate earnings.*[10] *This figure illustrates a long-term consumption risk of the type featured by Bansal and Yaron (2004). The long-term response is about double that of the short-term response. We will discuss the accuracy of this estimate in chapter 9.*

4.5 Evaluating long-term risk

A model from Hansen et al. (2008) describes a representative household that cares especially about long-term components of risk in consumption of the type discussed in examples 4.4.3 and 4.4.3.

Additive functional for utility

A representative household ranks consumption processes $\{C_t\}_{t=0}^{\infty}$ with a utility functional $\{V_t\}_{t=0}^{\infty}$ that satisfies the recursion:

$$\log V_t = [1 - \exp(-\delta)] \log C_t + \exp(-\delta) \log \mathbb{R}_t\left(V_{t+1}\right) \tag{4.8}$$

where

$$\mathbb{R}_t(V_{t+1}) = \left(E\left[(V_{t+1})^{1-\gamma} | \mathcal{F}_t\right]\right)^{\frac{1}{1-\gamma}}. \tag{4.9}$$

Here V_t is the date t continuation value for current and future consumption, $\delta > 0$ is a subjective discount rate, and $\gamma \geq 1$ is a risk aversion parameter appearing in the 'risk-sensitivity' operator $\mathbb{R}_t(V_{t+1})$ defined in equation (4.8).

Remark 4.5.1. *The limit of* $\mathbb{R}$ *as* γ *approaches* 1 *is just ordinary expected logarithmic utility:*

$$\lim_{\gamma \downarrow 1} \log \mathbb{R}_t(V_{t+1}) = \lim_{\gamma \downarrow 1} \frac{\log E\left[(V_{t+1})^{1-\gamma} | \mathcal{F}_t\right]}{1-\gamma} = E\left(\log V_{t+1} | \mathcal{F}_t\right).$$

Suppose that $\{\log C_t\}_{t=0}^{\infty}$ is an additive functional described by

$$\log C_{t+1} - \log C_t = \nu + D \cdot X_t + F \cdot W_{t+1}$$

[10] This is a consequence of Proposition 6.3.2.

where

$$X_{t+1} = AX_t + BW_{t+1},$$

A is a stable matrix and $\{W_{t+1}\}_{t=0}^{\infty}$ is an i.i.d. sequence of $\mathcal{N}(0, I)$ random vectors.

Proposition 4.5.2. *The value function process* $\{V_t\}_{t=0}^{\infty}$ *is described by*

$$\log V_t - \log C_t = U \cdot X_t + \mathbf{u} \tag{4.10}$$

where

$$U = \exp(-\delta)\left[I - \exp(-\delta)A'\right]^{-1} D, \tag{4.11}$$

and

$$\mathbf{u} = \frac{\exp(-\delta)}{1-\exp(-\delta)}\nu + \frac{(1-\gamma)}{2}\frac{\exp(-\delta)}{1-\exp(-\delta)}\left|D'\left[I - \exp(-\delta)A\right]^{-1} B + F\right|^2. \tag{4.12}$$

Proof. Transform the utility recursion (4.8) to

$$\log V_t - \log C_t = \exp(-\delta)\log \mathbb{R}_t\left[\left(\frac{V_{t+1}}{C_{t+1}}\right)\left(\frac{C_{t+1}}{C_t}\right)\right].$$

Guess that V_t has the form (8.9). Under this guess,

$$\left[\left(\frac{V_{t+1}}{C_{t+1}}\right)\left(\frac{C_{t+1}}{C_t}\right)\right]^{1-\gamma}$$

is a log-normal random variable with conditional mean

$$(1-\gamma)\left(A'U \cdot X_t + \mathbf{u} + D \cdot X_t + \nu\right)$$

and conditional variance

$$(1-\gamma)^2|U'B + F|^2.$$

Recall that for a log normal random variable, the logarithm of the mean is the mean of the associated normally distributed random variable plus one half the variance of this same random variable. It follows that U in the value function (8.9) satisfies $U = \exp(-\delta)A'U + \exp(-\delta)D$, which implies formula (4.11), which is independent of γ. Similarly,

$$\mathbf{u} = \exp(-\delta)\left[\mathbf{u} + \nu + \frac{1}{2}(1-\gamma)|U'B + F|^2\right],$$

which implies formula (4.12) for $\mathbf{u}$. □

In example 7.3.6 below, we will use a special case of this value functional to describe costs of random fluctuations in aggregate consumption.

Implied Stochastic Discount Factor Process

A stochastic discount factor process $\{S_t\}$ expresses how a consumer values exposures to risks. It provides a local way to assess how the decision maker responds to uncertainty. Such calculations have a variety of applications. First they provide the ingredients for the construction of asset pricing models. Second, they provide inputs into the computation of Pigouvian taxes for the purposes correcting for externalities in a socially optimal ways in the presence of uncertainty. Finally, they are sometimes valuable in assessing the impact of small (local) changes in policies.

The date zero value of a risky date t payout ξ_t is

$$\pi_0^t(\xi_t) = E\left[\left(\frac{S_t}{S_0}\right)\xi_t \middle| \mathcal{F}_0\right]. \tag{4.13}$$

We can compute the ratio $\frac{S_t}{S_0}$ that appears in formula (4.13) by evaluating the slope of an indifference curve that runs through both a baseline consumption process $\{C_t\}_{t=0}^{\infty}$ and a perturbed one

$$(C_0 - P_0(\mathbf{r}), C_1, C_2, \ldots, C_t + \mathbf{r}\xi_t, C_{t+1}, \ldots).$$

Here $P_0(\mathbf{r})$ expresses how much current period consumption must be reduced to keep a consumer on the same indifference curve when we replace C_t by $C_t + \mathbf{r}\xi_t$. We think of $\mathbf{r}$ as parameterizing an indifference curve. We set $\pi_0^t(\xi_t)$ defined in equation (4.13) equal to the slope of that indifference curve:

$$\pi_0^t(\xi_t) = \frac{d}{d\mathbf{r}} P_0(\mathbf{r})\big|_{\mathbf{r}=0}.$$

Applying this way of computing $\pi_0^t(\xi_t)$ in (4.13) to utility specification (4.8) results in

$$\frac{S_{t+1}}{S_t} = \exp(-\delta)\left(\frac{C_t}{C_{t+1}}\right)\left(\frac{(V_{t+1})^{1-\gamma}}{E\left[(V_{t+1})^{1-\gamma}|\mathcal{F}_t\right]}\right). \tag{4.14}$$

Fact 4.5.3. *The term* $\left(\frac{(V_{t+1})^{1-\gamma}}{E\left[(V_{t+1})^{1-\gamma}|\mathcal{F}_t\right]}\right)$ *is a nonnegative random variable with conditional expectation equal to unity. Therefore, it is a ratio of one-*

step transition probabilities that can be interpreted as a multiplicative increment to a likelihood ratio process, an object that will play a central role in chapter 8.

Taking logs on both sides to enable us to express equation (4.14) as

$$\begin{aligned}\log S_{t+1} - \log S_t = & -\delta - \log C_t + \log C_{t+1} \\ & + \left\{ (1-\gamma) \left[\log V_{t+1} - E\left(\log V_{t+1} | \mathcal{F}_t\right) \right] \right\} \\ & - \frac{(1-\gamma)^2}{2} \left| D' \left[I - \exp(-\delta) A \right]^{-1} B \exp(-\delta) + F' \right|^2 . \end{aligned} \tag{4.15}$$

From formulas (8.9), (4.11), and (4.12), the forward-looking term in braces on the right side of equation (4.15) is

$$\begin{aligned} & (1-\gamma) \left[(\log V_{t+1} - \log C_t) - E\left(\log V_{t+1} - \log C_t | \mathcal{F}_t\right) \right] \\ & = (1-\gamma) \left(D' \left[I - \exp(-\delta) A \right]^{-1} B \exp(-\delta) + F' \right) W_{t+1}. \end{aligned} \tag{4.16}$$

With this calculation, it is evident that the logarithm of the stochastic discount factor process is an additive functional.

Remark 4.5.4. *Notice that when* $\delta = 0$*,*

$$\left[D' \left[I - \exp(-\delta) A \right]^{-1} B \exp(-\delta) + F' \right]$$

appearing in (4.16) *equals the matrix* $[F + B'(I - A')^{-1} D]'$ *multiplying* W_{t+1} *in formula* (4.6) *for the martingale increment* $\kappa_a(X_t, W_{t+1})$ *of the additive functional that is* $\log C_t$*. Thus, formula* (4.16) *for the forward-looking term contributed by the continuation value is* $1-\gamma$ *times an approximation to the martingale increment of* $\{\log C_t\}$*, an approximation that becomes arbitrarily accurate when the subjective rate of discount* δ *becomes sufficiently small. Adding the contribution from* $\{\log C_t - \log C_{t+1}\}$ *on the right side of* (4.15)*, the martingale component of the logarithm of the stochastic discount factor then has an increment that approximates*

$$-\gamma \left[D' \left(I - A \right)^{-1} B + F' \right] W_{t+1},$$

when $\exp(-\delta)$ *is very close to* 1*. This is proportional to the martingale increment of* $\{\log C_t : t = 0, 1, 2, ...\}$*, the risk aversion parameter* γ *being* *the factor of proportionality. This martingale component of the logarithm of the stochastic discount factor process dominates the pricing of long-horizon risks. The minus sign in front of γ expresses that the representative consumer dislikes risk. The inner product of the vector $\left[D'(I-A)^{-1}B + F'\right]$ appearing in the martingale increment (with itself) turns out to be approximately equal to the variance of $\frac{1}{t}Y_t$ for large t (see formula (4.17) below).*

4.6 Digression on robustness

To fit asset pricing models to data, it is common to assume a large value of the risk aversion parameter γ in (4.9). Some agree with Lucas (2003) that the required value of γ indicates implausibly high risk-aversion. In this context it is noteworth that we can reinterpret the parameter γ in the risk-sensitivity operator defined in equation (4.9) as reflecting a representative consumer's concerns about robustness to model misspecification instead of risk aversion. A value of γ sufficiently high to fit risk prices and consumption volatility may not be implausible when we adopt a point of view of Barillas et al. (2009) and regard it as measuring something other than risk aversion. In developing our reinterpretation, we can appeal to insights from the contributions of Jacobson (1973), Whittle (1981), and Hansen and Sargent (1995) to control theory and the economics paper of Hansen et al. (2006).

The reinterpretation regards γ as measuring a representative consumer's concern about robustness of his valuations with respect to misspecifications of the stochastic process governing consumption, accomplished by recognizing the risk-sensitivity operator as an indirect utility function emerging from a minimization problem that we shall discuss in section 8.9.

4.7 Central Limit Theory

Let $\{Y_t : t = 0, 1, ...\}$ be an additive martingale process whose increments $Y_{t+1} - Y_t$ are stationary, ergodic, martingale differences:

$$E\left(Y_{t+1} - Y_t | \mathfrak{F}_t\right) = 0.$$

Billingsley (1961) proved that this process obeys a central limit theorem asserting that

$$\frac{1}{\sqrt{t}} Y_t \Longrightarrow N(0, E[(Y_{t+1} - Y_t)^2])$$

where $\Longrightarrow$ means convergence in distribution.[11] In Billingsley's central limit theorem, the increments to Y are martingale differences rather than i.i.d. as they are in more "standard" central limit theorems.

Gordin (1969) extends Billingsley's result to allow for temporally dependent increments. We can regard Gordin's result as an application of Proposition 4.2.3.

Corollary 4.7.1. *(Gordin (1969)) Suppose that the assumptions of Proposition 4.2.3 apply and that* $\nu = 0$. *Then*

$$\frac{1}{\sqrt{t}} Y_t \Longrightarrow N(0, \sigma^2)$$

where $\sigma^2 = E\left([\kappa_a(X_j, W_{j+1})]^2\right)$.[12]

The variance formula

$$\sigma^2 = \lim_{t \to \infty} \frac{1}{t} \text{variance}(Y_t) = E\left([\kappa_a(X_j, W_{j+1})]^2\right)$$

shows how properly to take into account the temporal dependence of the increments $(Y_{t+1} - Y_t)$ when computing the "long-run" volatility of the level Y_t. Notice that all that matters is the martingale component, not the stationary $g(X_t)$ component.

To illustrate, we return to the first-order VAR example 4.2 with $\nu = 0$:

$$\begin{aligned} X_{t+1} &= AX_t + BW_{t+1} \\ Y_{t+1} - Y_t &= D \cdot X_t + F \cdot W_{t+1}. \end{aligned}$$

The variance of the martingale increment that appears in Corollary 4.7.1 is[13]

$$\sigma^2 = [F + B'(I - A')^{-1}D] \cdot [F + B'(I - A')^{-1}D]. \tag{4.17}$$

[11] Ergodicity can be dispensed with if we replace the variance by $E[(Y_1 - Y_0)^2 | \mathfrak{I}]$ in the variance used for the normal approximation.

[12] Hall and Heyde (1980) show how to extend this approach to functional counterparts to the Central Limit Theorem.

[13] This expression for σ^2 equals the spectral density of $\{Y_{t+1} - Y_t\}$ at zero frequency. Recall that this term also played a key role in remark 4.5.4.

This differs from both the conditional variance $|F|^2$ and the unconditional variance, $D'\Sigma D + |F|^2$ of $Y_{t+1} - Y_t$, where

$$\Sigma = \sum_{j=0}^{\infty} (A)^j BB'(A^j)' \tag{4.18}$$

is the covariance matrix of the stationary distribution of X_t.

Since linear combinations of additive functionals are additive functionals, Corollary 4.7.1 can be applied to any linear combination of a vector of additive functionals.

4.8 Growth-rate Regimes

In this section, we construct a Proposition 4.2.3 decomposition for a model with persistent switches in both the conditional mean and the conditional volatility of the growth rate $Y_{t+1} - Y_t$. Section 4.9 then describes the Proposition 4.2.3 decomposition for a model in which a growth rate has stochastic volatility that is a quadratic function of the state X_t.

Suppose that $\{X_t\}$ evolves according to an n-state Markov chain with transition matrix $\mathbb{P}$. Realized values of X_t are coordinate vectors in $\mathbb{R}^n$. Suppose that $\mathbb{P}$ has only one unit eigenvalue. Let $\mathbf{q}$ be the row eigenvector associated with that unit eigenvalue normalized so that $\mathbf{q} \cdot \mathbf{1}_n = 1$:

$$\mathbf{q}'\mathbb{P} = \mathbf{q}'.$$

Consider an additive functional satisfying

$$Y_{t+1} - Y_t = D \cdot X_t + X_t{}' F W_{1,t+1},$$

where $\{W_{1,t}\}$ is an i.i.d. sequence of multivariate standard normally distributed random vectors. Evidently, the stationary Markov $\{X_t\}$ process induces discrete changes in the conditional mean and conditional volatility of the growth rate process $\{Y_{t+1} - Y_t\}$. We can represent the Markov chain as

$$X_{t+1} = \mathbb{P} X_t + W_{2,t+1}$$

where $E(X_{t+1}|X_t) = \mathbb{P} X_t$ and $\{W_{2,t+1}\}$ is an $n \times 1$ vector process that satisfies $E(W_{2,t+1}|X_t) = 0$, which is therefore a martingale difference sequence adapted to $X_t, X_{t-1}, \ldots, X_0$. Thus, we construct a second component of the shock vector as

$$W_{2,t+1} = X_{t+1} - E\left(X_{t+1}|X_t\right).$$

To apply algorithm 4.2.2, first compute

$$\kappa_2(X_t, W_{t+1}) = X_t{}' F W_{1,t+1},$$

and

$$\nu = D \cdot \mathbf{q}.$$

Let $f(x) = \mathbf{f} \cdot x$ where

$$\mathbf{f} = D - \nu \mathbf{1}_n.$$

Then $g(x) = \mathbf{g} \cdot x$ where $\mathbf{g}$ solves

$$\begin{aligned} \left(\mathbb{I} - \mathbb{P}\right)\mathbf{g} &= \mathbf{f} \\ \mathbf{q}'\mathbf{g} &= 0, \end{aligned}$$

where we include the second equation because the matrix $(\mathbb{I} - \mathbb{P})$ is singular since

$$\mathbf{q}'\left(\mathbb{I} - \mathbb{P}\right) = 0.$$

Set

$$\kappa_1(x, w^*) = \mathbf{f} \cdot x + \mathbf{g} \cdot x^* - \mathbf{g} \cdot x.$$

Recall that $\kappa_1(x, w^*)$ is the part of $g \cdot x^*$ that cannot be predicted given x, so

$$\kappa_1(x, w^*) = \mathbf{g} \cdot (x^* - \mathbb{P}x) = \mathbf{g} \cdot w_2^*.$$

So

$$Y_t = t\nu + \left[\sum_{j=1}^{t} \kappa_a(X_{j-1}, W_j)\right] - \mathbf{g} \cdot X_t + \mathbf{g} \cdot X_0 + Y_0,$$

where $\kappa_a = \kappa_1 + \kappa_2$. The martingale increment has both continuous and discrete components:

$$\kappa_a(X_t, W_{t+1}) \;=\; \underbrace{X_t{}' F W_{1,t+1}}_{\textbf{continuous}} + \underbrace{\mathbf{g} \cdot W_{2,t+1}}_{\textbf{discrete}}.$$

4.9 Quadratic Model of Growth

In this section, we describe an additive process with *stochastic volatility*. In chapter 10, we will characterize the behavior of prices of the risks driving such a process.

Suppose that $\{X_t\}$ follows the first-order autoregression

$$X_{t+1} = AX_t + BW_{t+1},$$

where A has stable eigenvalues and $\{W_{t+1} : t = 0, 1, ...\}$ is a sequence of independent and identically normally distributed random variables with mean zero and covariance matrix I. Consider an additive functional $\{Y_t\}$ defined by

$$Y_{t+1} - Y_t = \epsilon + D \cdot X_t + \frac{1}{2} {X_t}' H X_t + F \cdot W_{t+1} + {X_t}' G W_{t+1},$$

where H is a symmetric matrix. The function ${X_t}'G$ expresses stochastic volatility of the shock vector ${X_t}'GW_{t+1}$ in the conditional distribution of $Y_{t+1} - Y_t$.

To apply algorithm 4.2.2, first compute

$$\kappa_2(X_t, W_{t+1}) = F \cdot W_{t+1} + {X_t}' G W_{t+1}.$$

Next compute

$$\nu = \epsilon + \frac{1}{2} E\left({X_t}' H X_t\right) = \epsilon + \frac{1}{2} \text{trace}(H\Sigma),$$

where Σ is the covariance matrix in a stochastic steady state given by formula (4.18), and

$$f(x) = D \cdot x + \frac{1}{2} x' H x - \frac{1}{2} \text{trace}(H\Sigma).$$

Recall that $g - \mathbb{T} g = f$ and guess that

$$g(x) = \widehat{D} \cdot x + \frac{1}{2} x' \widehat{H} x - \frac{1}{2} \text{trace}\left(\widehat{H}\Sigma\right).$$

This guess gives rise to the following two relations:

$$\begin{aligned} \widehat{D} - A'\widehat{D} &= D, \\ \widehat{H} - A'\widehat{H}A &= H. \end{aligned} \tag{4.19}$$

It can be verified that

$$\begin{aligned}\widehat{H} &= \sum_{j=0}^{\infty} \left(A^j\right)' H \left(A^j\right)\\ \widehat{D} &= (I - A')^{-1} D.\end{aligned}$$

Since $\Sigma = BB' + A\Sigma A$,

$$\begin{aligned}\text{trace}\left(\widehat{H}\Sigma\right) &= \text{trace}\left(\widehat{H}BB'\right) + \text{trace}\left(\widehat{H}A\Sigma A'\right)\\ &= \text{trace}\left(B'\widehat{H}B\right) + \text{trace}\left(A'\widehat{H}A\Sigma\right)\\ &= \text{trace}\left(B'\widehat{H}B\right) + \text{trace}\left[\left(\widehat{H} - H\right)\Sigma\right],\end{aligned}$$

where the last equality follows from (4.19). Thus,

$$\text{trace}\left(B'\widehat{H}B\right) = \text{trace}\left(H\Sigma\right). \tag{4.20}$$

The increment to the martingale component of the additive functional is

$$\begin{aligned}\kappa_a(X_t, W_{t+1}) =& F \cdot W_{t+1} + X_t{}'GW_{t+1} + \left(B'\widehat{D}\right) \cdot W_{t+1}\\ &+ \frac{1}{2} X_{t+1}{}'\widehat{H}X_{t+1} + \frac{1}{2} X_t{}'\left(H - \widehat{H}\right) X_t - \nu\\ =& \left(F + B'\widehat{D}\right) \cdot W_{t+1} + X_t{}'\left(G + A'\widehat{H}\right) W_{t+1}\\ &+ \frac{1}{2} W_{t+1}{}'B'\widehat{H}BW_{t+1} - \frac{1}{2}\ \text{trace}(H\Sigma)\\ =& \left(F + B'\widehat{D}\right) \cdot W_{t+1} + X_t{}'\left(G + A'\widehat{H}\right) W_{t+1}\\ &+ \frac{1}{2} W_{t+1}{}'B'\widehat{H}BW_{t+1} - \frac{1}{2}\ \text{trace}(B'\widehat{H}B),\end{aligned}$$

where the last equality follows from (4.20).