TIØ 1: Financial Engineering in Energy Markets

Afzal Siddiqui
Department of Statistical Science
University College London
London WC1E 6BT, UK
afzal@stats.ucl.ac.uk

UCL

COURSE OUTLINE

- ★ Introduction (Chs 1–2)
- ★ Mathematical Background (Chs 3–4)
- ★ Investment and Operational Timing (Chs 5–6)
- ★ Entry, Exit, Lay-Up, and Scrapping (Ch 7)
- ★ Recent Theoretical Work I: Capacity Sizing
- ★ Recent Theoretical Work II: Risk Aversion and Multiple Risk Factors
- ★ Applications to the Energy Sector I: Capacity Sizing, Timing, and Operational Flexibility
- ★ Applications to the Energy Sector II: Modularity and Technology Choice

LECTURE OUTLINE

- ★ Stochastic processes
- ★ Wiener process and GBM
- ★ Itô's lemma
- ★ Dynamic programming
- ★ Contingent claims

STOCHASTIC PROCESSES: Discrete Time and Discrete State

- A variable that evolves over time in at least a partially random manner is a stochastic process
- More formally, a stochastic process is a law for the evolution of variable x_t over time t that allows us to calculate for various $t_1 < t_2 < t_3 < \ldots$ the joint probability $\mathcal{P}\left\{a_1 < x_1 \leq b_1, a_2 < x_2 \leq b_2, a_3 < x_3 \leq b_3, \ldots\right\}$
  - Stationary processes have statistical properties that are constant over long periods of time, e.g., temperature
  - Non-stationary processes may be things like stock prices
- Discrete-time processes change values only at discrete points in time, e.g., random walk
  - Starting with x_0, x_t takes independent jumps of size 1 (either up or down) at discrete points $t = 1, 2, 3, \ldots$ each with probability $\frac{1}{2}$
  - Thus, x_t has a binomial distribution: $\mathcal{P}[x_t = t - 2n] = \binom{t}{n}\left(\frac{1}{2}\right)^n\left(\frac{1}{2}\right)^{t-n} = \binom{t}{n}2^{-t}$ is the probability that there are n downward jumps and $t - n$ upward jumps by time t

STOCHASTIC PROCESSES: Discrete Time and Continuous State

- Instead of being a Bernoulli RV, the size of each jump may be a continuous RV, e.g., normal with mean zero and SD σ

- Another example is a first-order AR process, i.e., AR(1): $x_t = \delta + \rho x_{t-1} + \epsilon_t$, where $-1 < \rho < 1$ and ϵ_t is a standard normal RV
  - Stationary process with long-run expected value $\frac{\delta}{1-\rho}$
  - A mean-reverting process

- Both the random walk and AR(1) are Markov processes, i.e., the probability distribution for x_{t+1} depends only on x_t and is independent of anything that happened before time t

STOCHASTIC PROCESSES: Continuous Time

- A Wiener process (or Brownian motion) has the following properties:
  - Markov process
  - Independent increments
  - Changes over any finite time interval are normally distributed with variance that increases linearly in time
- Nice property that past patterns have no forecasting value
- For prices, it makes more sense to assume that changes in their logarithms are normally distributed, i.e., prices are lognormally distributed
- More formally for a Wiener process $\{z(t), t \geq 0\}$:
  - $\Delta z = \epsilon_t \sqrt{\Delta t}$, where $\epsilon_t \sim \mathcal{N}(0, 1)$
  - ϵ_t are serially uncorrelated, i.e., $\mathcal{E}[\epsilon_t \epsilon_s] = 0$ for $t \neq s$

STOCHASTIC PROCESSES: Continuous Time

- Implications of the two conditions are examined by breaking up the time interval T into n units of length Δt each
  - Change in z over T is $z(s+T) - z(s) = \sum_{i=1}^{n} \epsilon_i \sqrt{\Delta t}$, where the ϵ_i are independent
  - Via the CLT, $z(s+T) - z(s)$ is $\mathcal{N}(0, n\Delta t = T)$
  - Variance of the changes increases linearly in time
- Letting Δt become infinitesimally small implies $dz = \epsilon_t \sqrt{dt}$, where $\epsilon_t \sim \mathcal{N}(0,1)$
- This implies that $\mathcal{E}[dz] = 0$ and $\mathcal{V}(dz) = \mathcal{E}[(dz)^2] = dt$
- Coefficient of correlation between two Wiener processes, $z_1(t)$ and $z_2(t)$: $\mathcal{E}[dz_1 dz_2] = \rho_{12} dt$

STOCHASTIC PROCESSES: Brownian Motion with Drift

- Generalise the Wiener process: $dx = \alpha dt + \sigma dz$, where dz is the increment of the Wiener process, α is the drift parameter, and σ is the variance parameter
  - Over time interval Δt, Δx is normal with mean $\mathcal{E}[\Delta x] = \alpha \Delta t$ and variance $\mathcal{V}(\Delta x) = \sigma^2 \Delta t$
  - Given x_0, it is possible to generate sample paths
  - For example, if $\alpha = 0.2$ and $\sigma = 1.0$, then the discretisation with $\Delta t = \frac{1}{12}$ is $x_t = x_{t-1} + 0.01667 + 0.2887\epsilon_t$ (Figure 3.1)

- Optimal forecast is $\hat{x}_{t+T} = x_t + 0.01667T$ and 66% CI is $x_t + 0.01667T \pm 0.2887\sqrt{T}$ (Figure 3.2)

- Mean of $x_t - x_0$ is αt and its SD is $\sigma\sqrt{t}$, so the trend dominates in the long run

STOCHASTIC PROCESSES: Figures 3.1 and 3.2

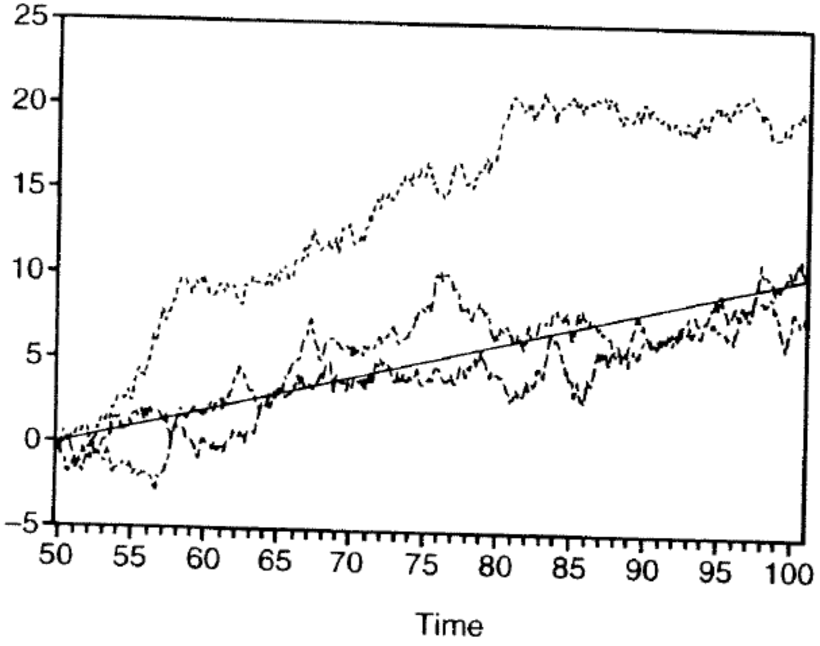

Figure 3.1. Sample Paths of Brownian Motion with Drift

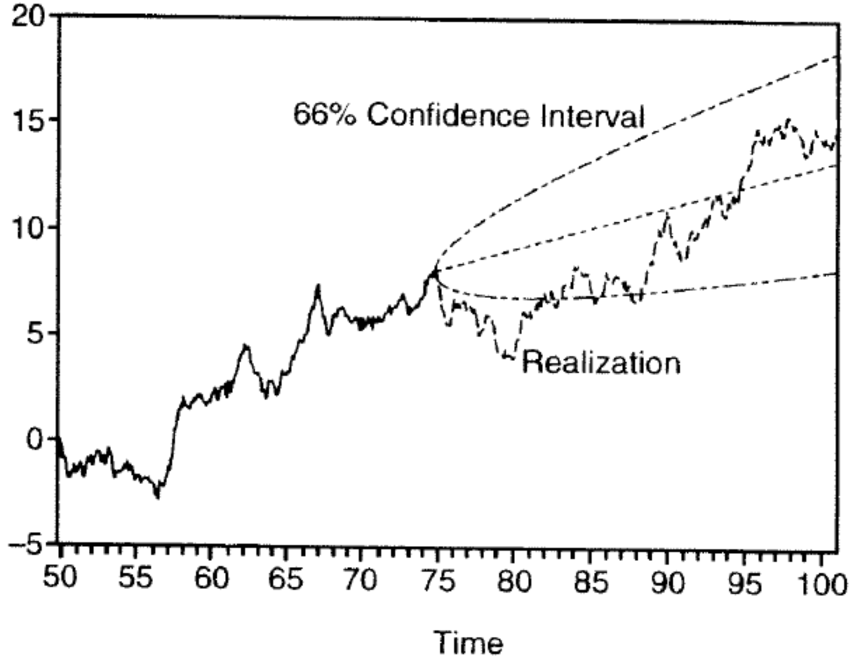

Figure 3.2. Optimal Forecast of Brownian Motion with Drift

STOCHASTIC PROCESSES: Brownian Motion and Random Walks

- Suppose that a discrete-time random walk for which the position is described by variable x makes jumps of $\pm\Delta h$ every Δt time units given the initial position x_0
  - The probability of an upward (downward) jump is p ($q = 1 - p$)
  - Thus, x follows a Markov process with independent increments, i.e., probability distribution of its future position depends only on its current position (Figure 3.3)
- Mean: $\mathcal{E}[\Delta x] = (p - q)\Delta h$; second moment: $\mathcal{E}[(\Delta x)^2] = p(\Delta h)^2 + q(\Delta h)^2 = (\Delta h)^2$; variance: $\mathcal{V}(\Delta x) = (\Delta h)^2[1 - (p - q)^2] = [1 - (2p - 1)^2](\Delta h)^2 = 4pq(\Delta h)^2$
- Thus, if t has $n = \frac{t}{\Delta t}$ steps, then $x_t - x_0$ is a binomial RV with mean $n\mathcal{E}[\Delta x] = \frac{t(p-q)\Delta h}{\Delta t}$ and variance $n\mathcal{V}(\Delta x) = \frac{4pqt(\Delta h)^2}{\Delta t}$

STOCHASTIC PROCESSES: Figure 3.3

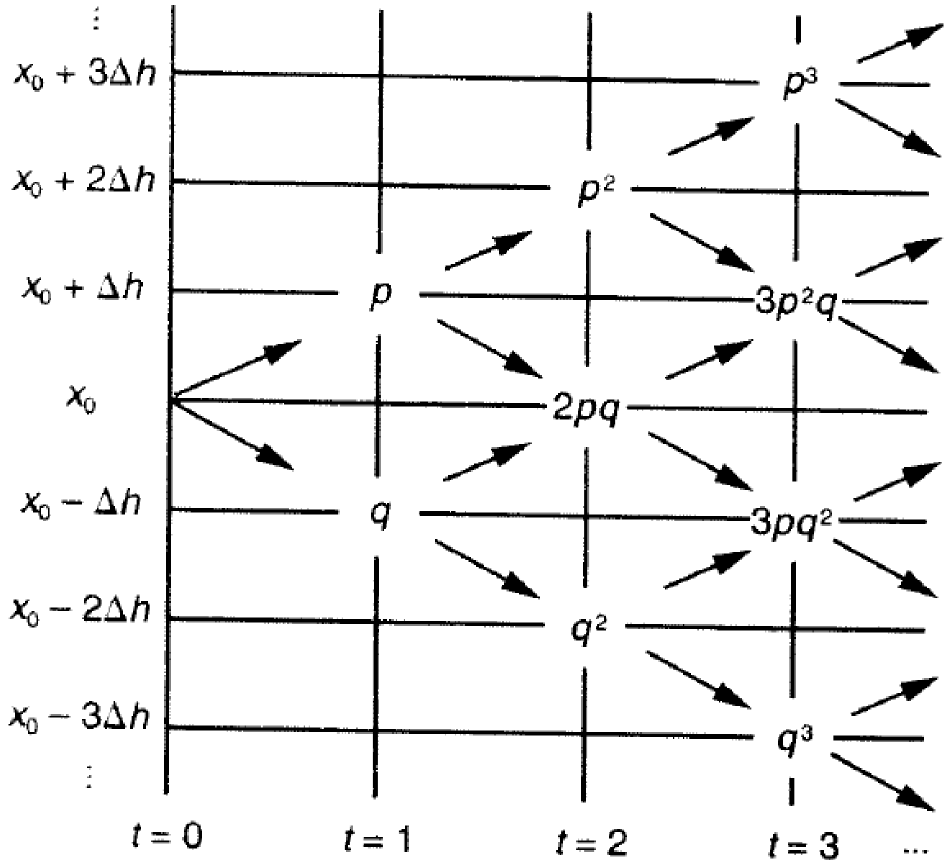

Figure 3.3. Random Walk Representation of Brownian Motion

STOCHASTIC PROCESSES: Brownian Motion and Random Walks

- Choose Δh, Δt, p, and q so that the random walk converges to a Brownian motion as $\Delta t \to 0$
  - $\Delta h = \sigma\sqrt{\Delta t}$
  - $p = \frac{1}{2}\left[1 + \frac{\alpha}{\sigma}\sqrt{\Delta t}\right]$, $q = \frac{1}{2}\left[1 - \frac{\alpha}{\sigma}\sqrt{\Delta t}\right]$
  - Thus, $p - q = \frac{\alpha}{\sigma}\sqrt{\Delta t} = \frac{\alpha}{\sigma^2}\Delta h$
- Substitute these into the formulas for the mean and variance $x_t - x_0$:
  - Mean: $\mathcal{E}[x_t - x_0] = \frac{t\alpha(\Delta h)^2}{\sigma^2\Delta t} = \frac{t\alpha\sigma^2\Delta t}{\sigma^2\Delta t} = \alpha t$; variance: $\mathcal{V}(x_t - x_0) = \frac{4pqt(\Delta h)^2}{\Delta t} = \frac{4t\sigma^2\Delta t\left[1 - \frac{\alpha^2}{\sigma^2}\Delta t\right]}{4\Delta t} = t\sigma^2\left[1 - \frac{\alpha^2}{\sigma^2}\Delta t\right]$, which goes to $t\sigma^2$ as $\Delta t \to 0$
- Hence, these are the mean and variance of a Brownian motion; furthermore, the binomial distribution approaches the normal one for large n

GENERALISED BROWNIAN MOTION

- An Itô process is $dx = a(x,t)dt + b(x,t)dz$, where dz is the increment of a Wiener process, and both $a(x,t)$ and $b(x,t)$ are known but may be functions of both x and t
  - Mean: $\mathcal{E}[dx] = a(x,t)dt$; second moment: $\mathcal{E}[(dx)^2] = \mathcal{E}[a^2(x,t)(dt)^2 + b^2(x,t)(dz)^2 + 2a(x,t)b(x,t)dtdz] = b^2(x,t)dt$; variance: $\mathcal{V}(dx) = \mathcal{E}[(dx)^2] - (\mathcal{E}[dx])^2 = b^2(x,t)dt$
- A geometric Brownian motion (GBM) has $a(x,t) = \alpha x$ and $b(x,t) = \sigma x$, which implies $dx = \alpha x dt + \sigma x dz$
  - Percentage changes in x are normally distributed, or absolute changes in x are lognormally distributed
  - If $\{y(t), t \geq 0\}$ is a BM with parameters $\left(\alpha - \frac{1}{2}\sigma^2\right) t$ and $\sigma^2 t$, then $\{x(t) \equiv x_0 e^{y(t)}, t \geq 0\}$ is a GBM
  - $m_y(s) = \mathcal{E}[e^{sy(t)}] = e^{s\alpha t - \frac{s\sigma^2 t}{2} + \frac{s^2\sigma^2 t}{2}}$, which implies $\mathcal{E}[y(t)] = \left(\alpha - \frac{1}{2}\sigma^2\right) t$ and $\mathcal{V}(y(t)) = \sigma^2 t$
  - Thus, $\mathcal{E}_{x_0}[x(t)] = \mathcal{E}_{x_0}[x_0 e^{y(t)}] = x_0 m_y(1) = x_0 e^{\alpha t}$ and $\mathcal{V}_{x_0}(x(t)) = \mathcal{E}_{x_0}[(x(t))^2] - (\mathcal{E}_{x_0}[x(t)])^2 = x_0^2 \mathcal{E}_{x_0}[e^{2y(t)}] - x_0^2 e^{2\alpha t} = x_0^2 e^{2\alpha t}[e^{\sigma^2 t} - 1]$

GEOMETRIC BROWNIAN MOTION TRAJECTORIES

- Expected PV of a GBM assuming discount rate $r > \alpha$ is $\mathcal{E}_{x_0}\left[\int_0^\infty x(t)e^{-rt}dt\right] = \int_0^\infty \mathcal{E}_{x_0}[x(t)]e^{-rt}dt = \int_0^\infty x_0 e^{\alpha t}e^{-rt}dt = \frac{x_0}{r-\alpha}$

- Generate sample paths for $\alpha = 0.09$ and $\sigma = 0.2$ per annum using $x_{1950} = 100$ and one-month intervals, i.e., $x_t - x_{t-1} = 0.0075x_{t-1} + 0.0577x_{t-1}\epsilon_t$, where $\epsilon_t \sim \mathcal{N}(0,1)$ (Figure 3.4)
  - Trend line is obtained by setting $\epsilon_t = 0$
  - Optimal forecast given x_{1974} is $\hat{x}_{1974+T} = (1.0075)^T x_{1974}$, while the CI is $(1.0075)^T(1.0577)^{\pm\sqrt{T}}x_{1974}$ (Figure 3.5)

GEOMETRIC BROWNIAN MOTION TRAJECTORIES: Figures 3.4 and 3.5

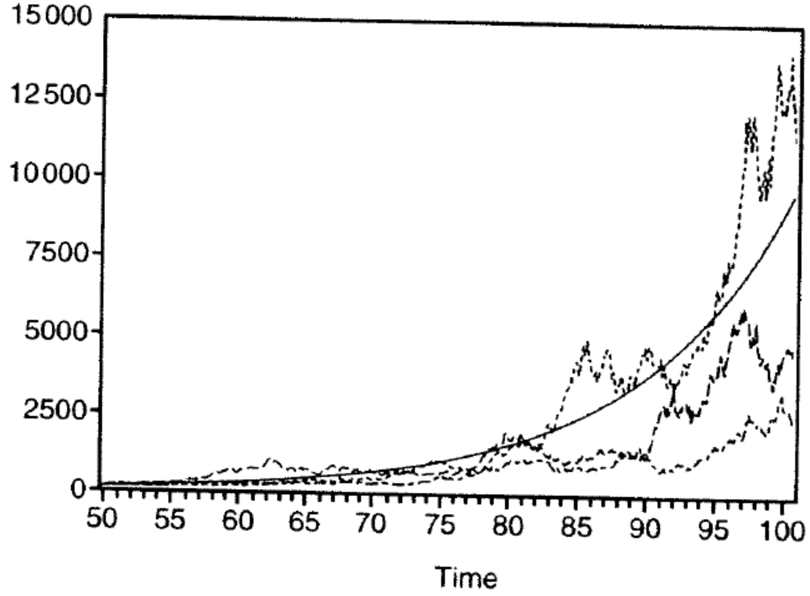

Figure 3.4. Sample Paths of Geometric Brownian Motion

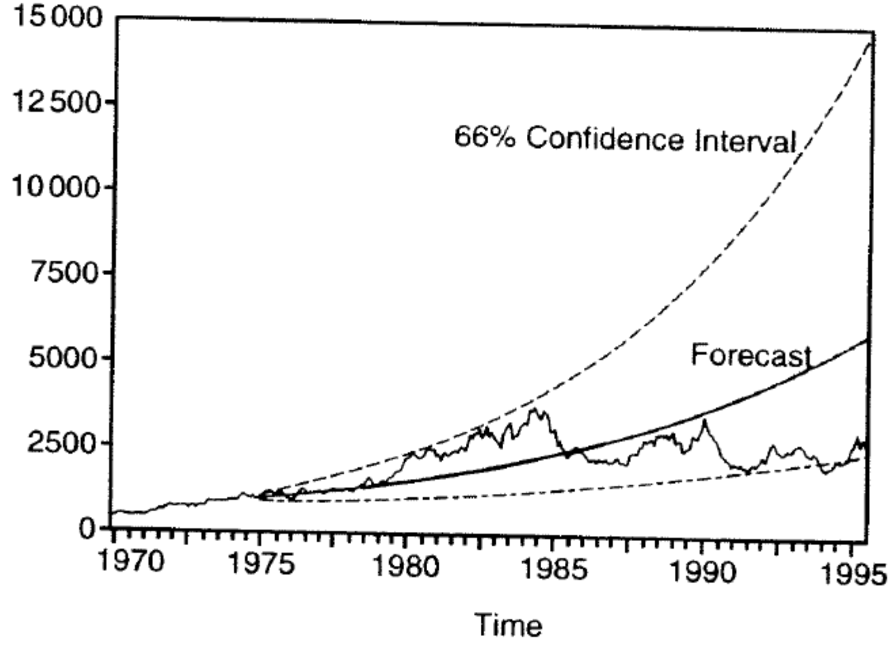

Figure 3.5. Optimal Forecast of Geometric Brownian Motion

MEAN-REVERTING PROCESSES

- Certain commodity prices tend to stay near their long-run marginal production costs, e.g., oil or copper

- Simplest mean-reverting (MR) process is the Ornstein-Uhlenbeck process: $dx = \eta(\overline{x} - x)dt + \sigma dz$
  - Satisfies the Markov property but does not have independent increments
  - Given $x(t) = x_0$, we have $\mathcal{E}_{x_0}[x(t)] = \overline{x} + (x_0 - \overline{x})e^{-\eta t}$ and $V_{x_0}[x(t) - \overline{x}] = \frac{\sigma^2}{2\eta}(1 - e^{-2\eta t})$
  - Note that as $t \to \infty$, the mean converges to $\overline{x}$ and the variance converges to $\frac{\sigma^2}{2\eta}$
  - As $\eta \to \infty$, the variance goes to zero
  - As $\eta \to 0$, $\{x(t), t \geq 0\}$ becomes a BM with variance $\sigma^2 t$
  - Figure 3.6 shows sample paths for $\overline{x} = 1$, $x_0 = 1$, $\sigma = 0.05$, and various values of η
  - Figure 3.7 shows the optimal forecast and CI

MEAN-REVERTING PROCESSES: Figures 3.6 and 3.7

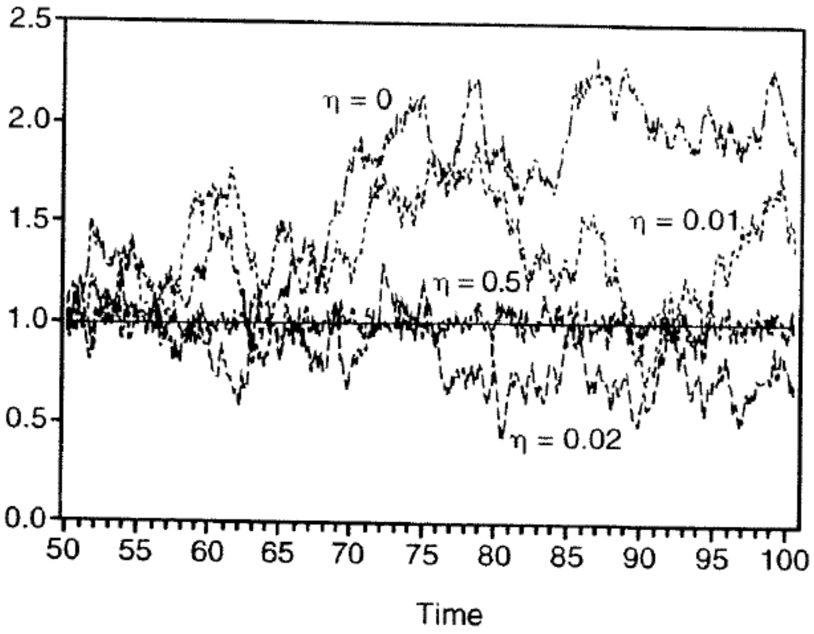

Figure 3.6. *Sample Paths of Mean-Reverting Process:* $dx = \eta(\bar{x} - x)dt + \sigma dz$

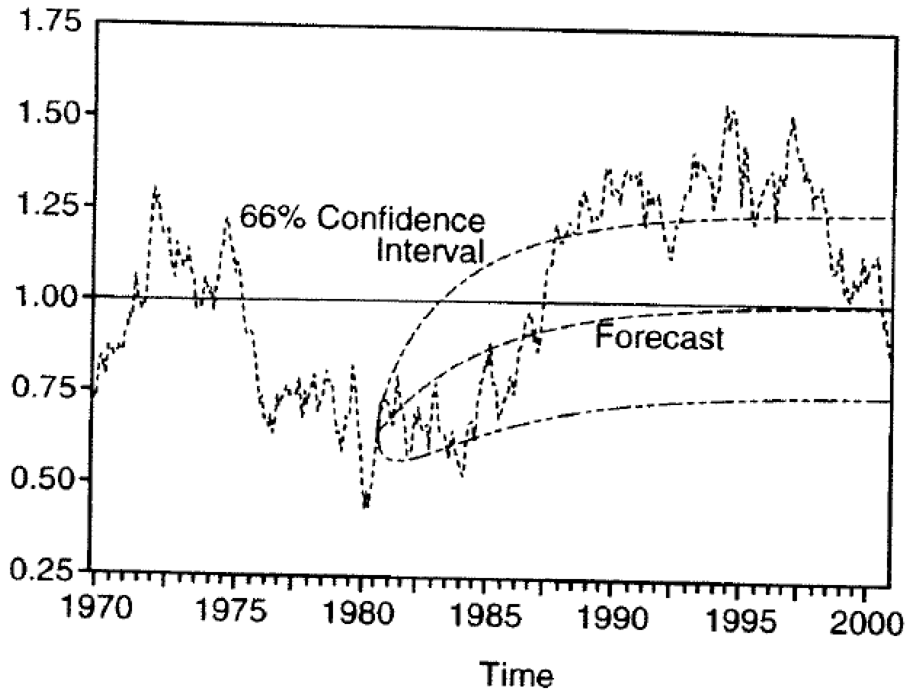

Figure 3.7. *Optimal Forecast of Mean-Reverting Process*

MEAN-REVERTING PROCESSES

- Equation for first-order autoregressive process is $x_t - x_{t-1} = \overline{x}(1-e^{-\eta}) + (e^{-\eta}-1)x_{t-1} + \epsilon_t$, where $\epsilon_t \sim \mathcal{N}(0, \sigma_\epsilon)$ and $\sigma_\epsilon^2 = \frac{\sigma^2}{2\eta}(1-e^{-2\eta})$
  - Estimate parameters by running the regression $x_t - x_{t-1} = a + bx_{t-1} + \epsilon_t$
  - Thus, $\overline{x} = -\frac{\hat{a}}{\hat{b}}$, $\hat{\eta} = -\ln(1+\hat{b})$, and $\hat{\sigma}^2 = \frac{\hat{\sigma}_\epsilon^2 \ln(1+\hat{b})^2}{(1+\hat{b})^2 - 1}$
- Can also have a geometric MR process: $dx = \eta x(\overline{x} - x)dt + \sigma x dz$
- In order to check for mean reversion, perform unit root tests on many years of data
  - Figures 3.8 and 3.9 indicate that commodity prices are mean reverting but with a low rate of mean reversion

MEAN-REVERTING PROCESSES: Figures 3.8 and 3.9

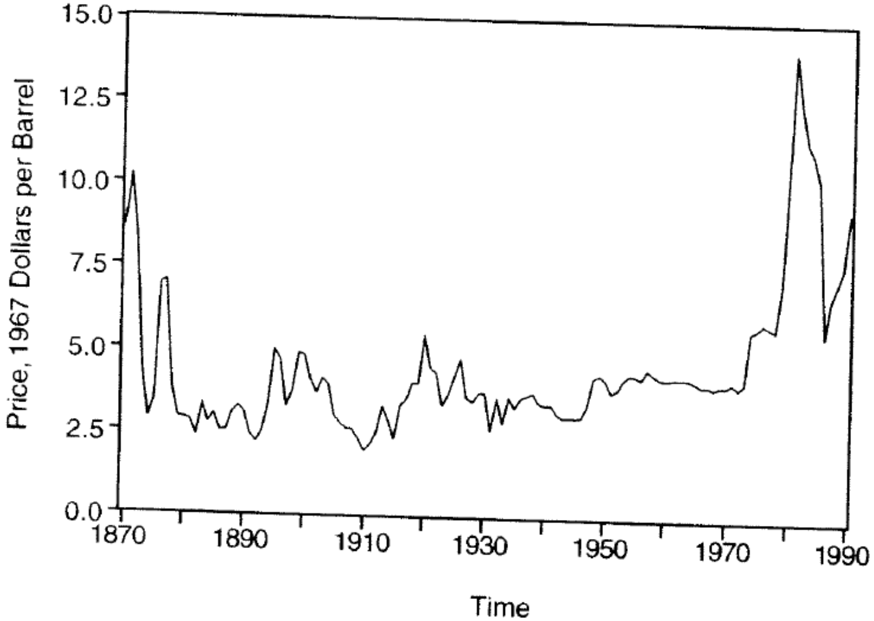

Figure 3.8. *Price of Crude Oil in 1967 Dollars per Barrel*

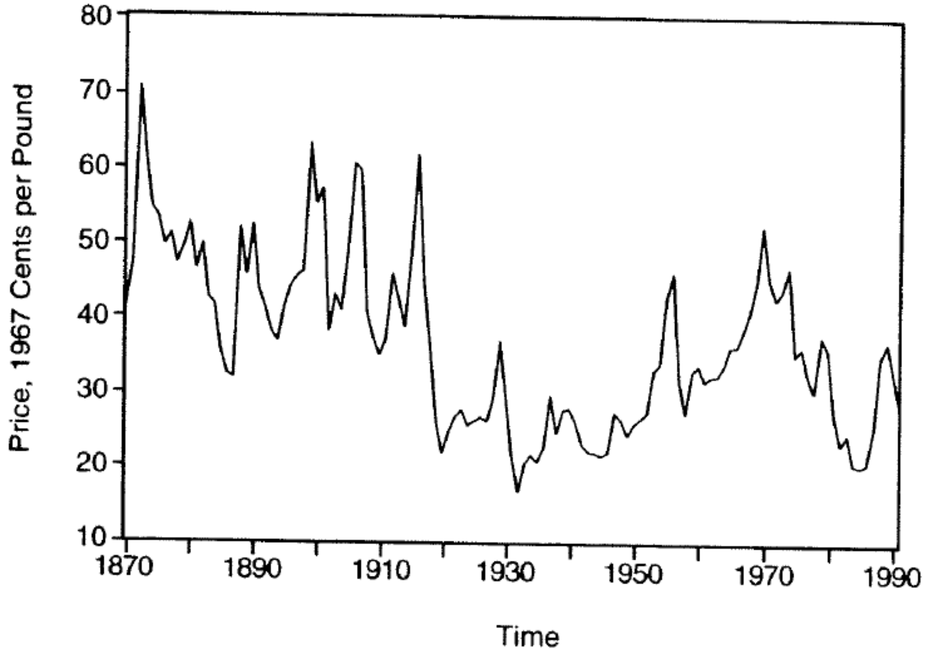

Figure 3.9. *Price of Copper in 1967 Cents per Pound*

ITÔ'S LEMMA

- Itô's lemma allows us to integrate and differentiate functions of Itô processes
  - Recall Taylor series expansion for $F(x,t)$: $dF = \frac{\partial F}{\partial x}dx + \frac{\partial F}{\partial t}dt + \frac{1}{2}\frac{\partial^2 F}{\partial x^2}(dx)^2 + \frac{1}{6}\frac{\partial^3 F}{\partial x^3}(dx)^3 + \cdots$
  - Usually, higher-order terms vanish, but here $(dx)^2 = b^2(x,t)dt$ (once terms in $(dt)^{\frac{3}{2}}$ and $(dt)^2$ are ignored), which is linear in dt
  - Thus, $dF = \frac{\partial F}{\partial x}dx + \frac{\partial F}{\partial t}dt + \frac{1}{2}\frac{\partial^2 F}{\partial x^2}(dx)^2 \Rightarrow dF = \left[\frac{\partial F}{\partial t} + a(x,t)\frac{\partial F}{\partial x} + \frac{1}{2}b^2(x,t)\frac{\partial^2 F}{\partial x^2}\right]dt + b(x,t)\frac{\partial F}{\partial x}dz$
  - Intuitively, even if $a(x,t) = 0$ and $\frac{\partial F}{\partial t} = 0$, then $\mathcal{E}[dx] = 0$, but $\mathcal{E}[dF] \neq 0$ because of Jensen's inequality

- Generalise to m Itô processes with $dx_i = a_i(x_1,\ldots,x_m,t)dt + b_i(x_1,\ldots,x_m,t)dx_i$ and $\mathcal{E}[dz_i dz_j] = \rho_{ij}dt$: $dF = \frac{\partial F}{\partial t}dt + \sum_i \frac{\partial F}{\partial x_i}dx_i + \frac{1}{2}\sum_i\sum_j \frac{\partial^2 F}{\partial x_i \partial x_j}dx_i dx_j$

APPLICATION TO GBM

- If $dx = \alpha x dt + \sigma x dz$ and $F(x) = \ln(x)$, then $F(x)$ follows a BM with parameters $\alpha - \frac{1}{2}\sigma^2$ and σ
  - $\frac{\partial F}{\partial t} = 0$, $\frac{\partial F}{\partial x} = \frac{1}{x}$, $\frac{\partial^2 F}{\partial x^2} = -\frac{1}{x^2}$, which implies that $dF = \frac{dx}{x} - \frac{1}{2x^2}(dx)^2 = \alpha dt + \sigma dz - \frac{1}{2}\sigma^2 dt = (\alpha - \frac{1}{2}\sigma^2)dt + \sigma dz$
- Consider $F(x, y) = xy$ and $G = \ln F$ with $dx = \alpha_x x dt + \sigma_x x dz_x$, $dy = \alpha_y y dt + \sigma_y y dz_y$, and $\mathcal{E}[dz_x dz_y] = \rho dt$
  - $\frac{\partial^2 F}{\partial x^2} = \frac{\partial^2 F}{\partial y^2} = 0$ and $\frac{\partial^2 F}{\partial x \partial y} = 1$, which implies $dF = ydx + xdy + dxdy$
  - Substitute dx and dy: $dF = \alpha_x xydt + \sigma_x xydz_x + \alpha_y xydt + \sigma_y xydz_y + xy\sigma_x\sigma_y\rho dt \Rightarrow dF = (\alpha_x + \alpha_y + \rho\sigma_x\sigma_y)Fdt + (\sigma_x dz_x + \sigma_y dz_y)F$, i.e., F is also a GBM
  - Meanwhile, $dG = (\alpha_x + \alpha_y - \frac{1}{2}\sigma_x^2 - \frac{1}{2}\sigma_y^2)dt + \sigma_x dz_x + \sigma_y dz_y$
- Discounted PV: $F(x) = x^\theta$ and x follows a GBM
  - F follows a GBM, too: $dF = \theta x^{\theta-1}dx + \frac{1}{2}\theta(\theta - 1)x^{\theta-1}(dx)^2 = F[\theta\alpha + \frac{1}{2}\theta(\theta-1)\sigma^2]dt + \theta\sigma F dz \Rightarrow \mathcal{E}_{x_0}[F(x(t))] = F(x_0)e^{t(\theta\alpha + \frac{1}{2}\theta(\theta-1)\sigma^2)}$
  - Thus, $\mathcal{E}_{x_0}\left[\int_0^\infty F(x(t))e^{-rt}dt\right] = \frac{x_0^\theta}{r - \alpha\theta - \frac{1}{2}\theta(\theta-1)\sigma^2}$

STOCHASTIC DISCOUNT FACTOR

★ Proposition: The conditional expectation of the stochastic discount factor, $\mathcal{E}_p\left[e^{-\rho\tau}\right]$, is the power function, $\left(\frac{p}{P_I}\right)^{\beta_1}$, where $\tau \equiv \min\left\{t : P_t \geq P_I\right\}$

★ Proof: Let $g(p) \equiv \mathcal{E}_p\left[e^{-\rho\tau}\right]$

- $g(p) = o(dt)e^{-\rho dt} + (1 - o(dt))e^{-\rho dt}\mathcal{E}_p\left[g(p + dP)\right]$
- $\Rightarrow g(p) = o(dt)e^{-\rho dt} + (1 - o(dt))e^{-\rho dt}\mathcal{E}_p\left[g(p) + dPg^{'}(p) + \frac{1}{2}(dP)^2 g^{''}(p) + o(dt)\right]$
- $\Rightarrow g(p) = o(dt) + e^{-\rho dt}g(p) + e^{-\rho dt}\alpha p g^{'}(p)dt + e^{-\rho dt}\frac{1}{2}\sigma^2 p^2 g^{''}(p)dt$
- $\Rightarrow g(p) = o(dt) + (1 - \rho dt)g(p) + (1 - \rho dt)\alpha p g^{'}(p)dt + (1 - \rho dt)\frac{1}{2}\sigma^2 p^2 g^{''}(p)dt$
- $\Rightarrow -\rho g(p) + \alpha p g^{'}(p) + \frac{1}{2}\sigma^2 p^2 g^{''}(p) = \frac{o(dt)}{dt}$
- $\Rightarrow g(p) = a_1 p^{\beta_1} + a_2 p^{\beta_2}$
- $\lim_{p\to 0} g(p) = 0 \Rightarrow a_2 = 0$ and $g(P_I) = 1 \Rightarrow a_1 = \frac{1}{P_I^{\beta_1}}$

DYNAMIC PROGRAMMING: Many-Period Example

- Now, let the state variable x_t be continuous and the control variable u_t represent the possible choices made at time t
  - Let the immediate profit flow be $\pi_t(x_t, u_t)$ and $\Phi_t(x_{t+1}|x_t, u_t)$ be the CDF of the state variable next period given current information
  - Given the discount rate ρ and the Bellman Principle of Optimality, the expected NPV of the cash flows to go from period t is $F_t(x_t) = \max_{u_t}\left\{\pi_t(x_t, u_t) + \frac{1}{(1+\rho)}\mathcal{E}_t[F_{t+1}(x_{t+1})]\right\}$
  - Use the termination value at time T and work backwards to solve for successive values of u_t: $F_{T-1}(x_{T-1}) = \max_{u_{T-1}}\left\{\pi_{T-1}(x_{T-1}, u_{T-1}) + \frac{1}{(1+\rho)}\mathcal{E}_{T-1}[\Omega_T(x_T)]\right\}$
- With an infinite horizon, it is possible to solve the problem recursively due to independence from time and the downward scaling due to the discount factor: $F(x) = \max_u\left\{\pi(x, u) + \frac{1}{(1+\rho)}\mathcal{E}[F(x')|x, u]\right\}$

DYNAMIC PROGRAMMING: Optimal Stopping

★ Suppose that the choice is binary: either continue (to wait or to produce) or to terminate (waiting or production)

- Bellman equation is now $\max\left\{\Omega(x), \pi(x) + \frac{1}{(1+\rho)}\mathcal{E}[F(x')|x]\right\}$
- Focus on case where it is optimal to continue for $x > x^*$ and stop otherwise
- Continuation is more attractive for higher x if: (i) immediate profit from continuation becomes larger relative to the termination payoff, i.e., $\pi(x) + \frac{1}{(1+\rho)}\mathcal{E}[\Omega(x')|x] - \Omega(x)$ is increasing in x, and (ii) current advantage should not be likely to be reversed in the near future, i.e., require first-order stochastic dominance
- Both conditions are satisfied in the applications studied here: (i) always holds, and (ii) is true for random walks, Brownian motion, MR processes, and most other economic applications
- In general, may have stopping threshold that varies with time, $x^*(t)$

DYNAMIC PROGRAMMING: Continuous Time

- In continuous time, the length of the time period, Δt, goes to zero and all cash flows are expressed in terms of rates
  - Bellman equation is now $F(x,t) = \max_u \left\{ \pi(x,u,t)\Delta t + \frac{1}{(1+\rho\Delta t)} \mathcal{E}[F(x',t+\Delta t)|x,u] \right\}$
  - Multiply by $(1 + \rho\Delta t)$ and re-arrange: $\rho\Delta t F(x,t) = \max_u \{\pi(x,u,t)\Delta t(1+\rho\Delta t) + \mathcal{E}[F(x',t+\Delta t) - F(x,t)|x,u]\} = \max_u \{\pi(x,u,t)\Delta t(1+\rho\Delta t) + \mathcal{E}[\Delta F|x,u]\}$
  - Divide by Δt and let it go to zero to obtain $\rho F(x,t) = \max_u \left\{ \pi(x,u,t) + \frac{\mathcal{E}[dF|x,u]}{dt} \right\}$
  - Intuitively, the instantaneous rate of return on the asset must equal its expected net appreciation

DYNAMIC PROGRAMMING: Itô Processes

- Suppose that $dx = a(x,u,t)dt + b(x,u,t)dz$ and $x' = x + dx$
- Apply Itô's lemma to the value function, F:
  - $\mathcal{E}[F(x+\Delta, t+\Delta t)|x,u] = F(x,t) + [F_t(x,t) + a(x,u,t)F_x(x,t) + \frac{1}{2}b^2(x,u,t)F_{xx}(x,t)]\Delta t + o(\Delta t)$
  - Return equilibrium condition is now $\rho F(x,t) = \max_u \left\{\pi(x,u,t) + F_t(x,t) + a(x,u,t)F_x(x,t) + \frac{1}{2}b^2(x,u,t)F_{xx}(x,t)\right\}$
  - Next, find optimal u as a function of $F_t(x,t)$, $F_x(x,t)$, $F_{xx}(x,t)$, x, t, and underlying parameters
  - Subsitute it back into the return equilibrium condition to obtain a second-order PDE with F as the dependent variable and x and t as the independent ones
  - Solution procedure is typically to start at the terminal time T and work backwards
- When time horizon is infinite, t drops out of the equation:
  - $\rho F(x) = \max_u \left\{\pi(x,u) + a(x,u)F'(x) + \frac{1}{2}b^2(x,u)F''(x)\right\}$

DYNAMIC PROGRAMMING: Optimal Stopping and Smooth Pasting

★ Consider a binary decision problem: can either continue to obtain a profit flow (with continuation value) or stop to obtain a termination payoff where $dx = a(x,t)dt + b(x,t)dz$

- In this case, a threshold policy with $x^*(t)$ exists, and the Bellman equation is $\rho F(x,t)dt = \max\{\Omega(x,t)dt, \pi(x,t)dt + \mathcal{E}[dF|x]\}$
- The RHS is larger in the continuation region, so applying Itô's lemma gives $\frac{1}{2}b^2(x,t)F_{xx}(x,t) + a(x,t)F_x(x,t) + F_t(x,t) - \rho F(x,t) + \pi(x,t) = 0$
- The PDE can be solved for $F(x,t)$ for $x > x^*(t)$ subject to the boundary condition $F(x^*(t),t) = \Omega(x^*(t),t)\ \forall t$ (value-matching condition)
- A second condition is necessary to find the free boundary: $F_x(x^*(t),t) = \Omega_x(x^*(t),t)\ \forall t$ (smooth-pasting condition)
- The latter may be thought of as a first-order necessary condition, i.e., if the two curves met at a kink, then the optimal stopping would occur elsewhere

DYNAMIC PROGRAMMING EXAMPLE: Optimal Abandonment

- You own a machine that produces profit, x, that evolves according to a BM process, i.e., $dx = adt + bdz$, where $a < 0$ to reflect decay of the machine over time

- The lifetime of the machine is T years, discount rate is ρ, and we must find the optimal threshold profit level, $x^*(t)$, below which to abandon the machine (zero salvage value)
  - Corresponding PDE is $\frac{1}{2}b^2 F_{xx}(x,t) + aF_x(x,t) + F_t(x,t) - \rho F(x,t) + x = 0$
  - PDE is solved numerically for $T = 10$, $a = -0.1$, $b = 0.2$, and $\rho = 0.10$ using discrete time steps of $\Delta t = 0.01$
  - Solution in Figure 4.1 indicates that for lifetimes greater than ten years, the optimal abandonment threshold is about -0.17
  - As lifetime is reduced, it becomes easier to abandon the machine

DYNAMIC PROGRAMMING EXAMPLE: Figure 4.1

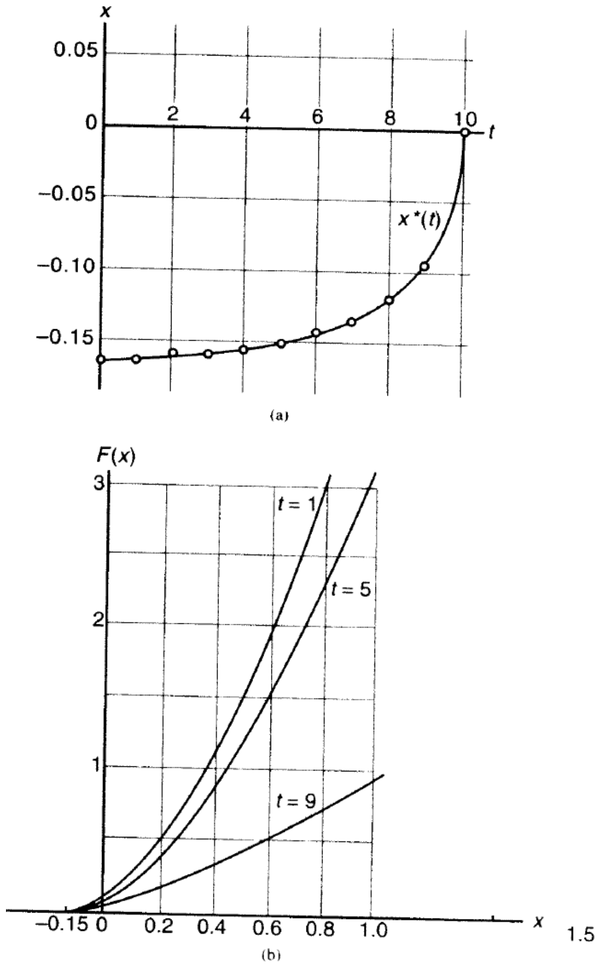

Figure 4.1. Depreciation and Abandonment

DYNAMIC PROGRAMMING EXAMPLE: Optimal Abandonment

★ Assume an effectively infinite lifetime to obtain an ODE instead of a PDE: $\frac{1}{2}b^2F''(x) + aF'(x) - \rho F(x) + x = 0$

- Homogeneous solution is $y(x) = c_1e^{r_1x} + c_2e^{r_2x}$
- Substituting derivatives into the homogeneous portion of the ODE yields $c_1e^{r_1x}(\frac{1}{2}b^2r_1^2 + ar_1 - \rho) + c_2e^{r_2x}(\frac{1}{2}b^2r_2^2 + ar_2 - \rho) = 0$
- The terms in the parentheses must be equal to zero, i.e., $r_1 = \frac{-a+\sqrt{a^2+2b\rho}}{b^2} = 5.584 > 0$ and $r_2 = \frac{-a-\sqrt{a^2+2b\rho}}{b^2} = -0.854 < 0$
- Particular solution: $Y(x) = Ax + B$, $Y'(x) = A$, and $Y''(x) = 0$
- Substituting these into the original ODE yields $aA - \rho(Ax + B) + x = 0 \Rightarrow A = \frac{1}{\rho}, B = \frac{a}{\rho^2}$
- Thus, $Y(x) = \frac{x}{\rho} + \frac{a}{\rho^2}$, and $F(x) = c_1e^{r_1x} + c_2e^{r_2x} + \frac{x}{\rho} + \frac{a}{\rho^2}$
- Boundary conditions: (i) $F(x^*) = 0$, (ii) $F'(x^*) = 0$, (iii) $\lim_{x\to\infty} F(x) = Y(x)$
- The third one implies that $c_1 = 0$, i.e., $F(x) = c_2e^{r_2x} + \frac{x}{\rho} + \frac{a}{\rho^2}$
- First two conditions imply $x^* = -\frac{a}{\rho} + \frac{1}{r_2} = -0.17$ and $c_2 = -\frac{e^{-r_2x^*}}{r_2\rho}$

CONTINGENT CLAIMS: Replicating Portfolio

- Dynamic programming uses an exogenous discount rate, ρ, which is assumed to the opportunity cost of capital
- Financial theory has a more sophisticated treatment of this topic in terms of relating this cost to the market portfolio
  - Assume profit flow, x, follows a GBM and the output of the firm can be traded in financial markets
  - Output held by investors if it provides a sufficiently high return: part of it from α and another from the convenience yield, $\delta = \mu - \alpha$
  - The risk-adjusted rate of return is obtained from CAPM: $\mu = r + \phi\sigma\rho_{xm}$, where ϕ is the market price of risk and ρ_{xm} is the correlation between returns

CONTINGENT CLAIMS: Replicating Portfolio

- Value of a firm, $F(x,t)$, with profit flow, $\pi(x,t)$, may be replicated by investing a dollar in the risk-free asset and holding n units of the output
  - Portfolio costs $\$(1+nx)$, and if held for dt time units, then it provides a safe return of rdt, a dividend of $n\delta x dt$, and a stochastic capital gain of $ndx = n\alpha x dt + n\sigma x dz$
  - The total return per dollar invested is $\frac{r+n(\alpha+\delta)x}{1+nx}dt + \frac{\sigma n x}{1+nx}dz$
  - Ownership of the firm over dt costs $F(x,t)$ and offers a profit flow $\pi(x,t)dt$ along with a stochastic capital gain $dF = [F_t(x,t) + \alpha x F_x(x,t) + \frac{1}{2}\sigma^2 x^2 F_{xx}(x,t)]dt + \sigma x F_x(x,t)dz$
  - Thus, total return per dollar is $\frac{\pi(x,t)+F_t(x,t)+\alpha x F_x(x,t)+\frac{1}{2}\sigma^2 x^2 F_{xx}(x,t)}{F(x,t)}dt + \frac{\sigma x F_x(x,t)}{F(x,t)}dz$

CONTINGENT CLAIMS: Replicating Portfolio

- Matching the risk terms gives $\frac{nx}{(1+nx)} = \frac{xF_x(x,t)}{F(x,t)} \Rightarrow n = \frac{F_x(x,t)}{(F(x,t)-xF_x(x,t))}$
  - Matching the return terms gives $\frac{\pi(x,t)+F_t(x,t)+\alpha xF_x(x,t)+\frac{1}{2}\sigma^2x^2F_{xx}(x,t)}{F(x,t)} = \frac{r+n(\alpha+\delta)x}{1+nx}$
  - Substituting for n implies that the RHS becomes $r\frac{(F(x,t)-xF_x(x,t))}{F(x,t)} + (\alpha+\delta)\frac{xF_x(x,t)}{F(x,t)}$
  - Re-arranging the return equation then yields $\frac{1}{2}\sigma^2x^2F_{xx}(x,t)+(r-\delta)xF_x(x,t) + F_t(x,t) - rF(x,t) + \pi(x,t) = 0$
  - Similar to the PDE obtained via dynamic programming
  - Can also use a risk-free portfolio by holding one unit of $F(x,t)$ and n units short of the underlying asset x

CONTINGENT CLAIMS: Spanning Assets

- If x is not directly traded, then we can use a spanning asset, i.e., one whose risk tracks the uncertainty in x
  - Suppose replicating asset follows $dX = A(x,t)Xdt + B(x,t)Xdz$, i.e., have the same dz even if the other coefficients are different
  - If there is a dividend flow rate, $D(x,t)$, then one dollar invested in X over time dt provides the return $[D(x,t)+A(x,t)]dt+B(x,t)dz$
  - An investor will require return $\mu_X(x,t) = r + \phi\rho_{xm}B(x,t)$, which must equal $D(x,t) + A(x,t)$
- Risk-free portfolio will cost $F - nX$ to buy and provide dividend flows of $[\pi - nDX]dt$
  - Capital gain on the portfolio is $dF - ndX = [F_t + aF_x + \frac{1}{2}b^2F_{xx} - nAX]dt + [bF_x - nBX]dz$, so risk-free portfolio requires $n = \frac{bF_x}{BX}$
  - Set expected net return on portfolio to the risk-free return on its cost: $r[F - nX]dt = [F_t + aF_x + \frac{1}{2}b^2F_{xx} - nAX]dt + \pi dt - nDXdt$
  - Thus: $\frac{1}{2}b^2F_{xx} + aF_x + F_t - rF + rnX - nDX - nAX + \pi = 0 \Rightarrow \frac{1}{2}b^2F_{xx} + aF_x + F_t rF + \frac{rbF_x}{B} - \frac{DbF_x}{B} - \frac{AbF_x}{B} + \pi = 0$
  - $\frac{1}{2}b^2F_{xx} + aF_x + F_t - rF + \frac{rbF_x}{B} - \frac{\mu_X bF_x}{B} + \pi = 0$

QUESTIONS