A Continuous Time Multi-Dimensional Asset Pricing Model with Duffie-Epstein Preferences and Kreps-Porteus Utility

Peter Kelly

University of Notre Dame

Department of Mathematics

Advisors: Alex A. Himonas and Tom F. Cosimano

April 28, 2009

Abstract

Our research creates a well-defined problem to find an analytic solution for the price-dividend ratio. We use a two-dimensional consumption process and assume Duffie-Epstein preferences and Kreps-Porteus utility. We start with a one-dimensional model, the solution to which we will use in our more complicated model. We also derive the famous Black-Scholes formula for call options.

Acknowledgements

First, I would like to thank Alex Himonas and Tom Cosimano for their invaluable help throughout this process. They have been great advisors and I have thoroughly enjoyed working with them. I would also like to thank Jon Poelhuis and Kate Manley. I worked with Jon and Kate over the summer and am very grateful for their friendship and help during this process. Specifically, I would like to thank Kate Manley for teaching me how to create pictures and for creating a few of the pictures in this document.

1. Introduction

The price to dividend ratio is a number that describes the ratio of the price of a stock to its yearly dividends. For example, if the price of a current stock is eighty dollars and the dividends paid out last year equals two dollars then the price to dividend ratio is forty. A typical average price to dividend ratio for the market and, in fact, the one that we use in our research is twenty three. The price to dividend ratio is a very important tool because it can be used to price derivatives. A derivative is a contract based on factors such as assets or indices. They are very important because they can be used to reduce risk. Risk reduction is a critical component of any investment. The classic example that shows how a derivative can be helpful concerns an airline company worried about the rising price of oil. A call option on oil will allow this airline company to purchase oil at a strike price K at a future time T no matter what the price of oil happens to be, thereby eliminating risk associated with the volatility of oil prices. It is clear that these instruments are vital to hedging risk and permit a determination of their fair prices.

My research seeks to find a closed form solution for the price to dividend ratio. A closed form solution for the price to dividend ratio is one that solves for the price to dividend ratio with a number of well known functions. In my research, the price to dividend ratio will have a power series solution.

This process starts with the utility function. The utility function is an ordinal mapping that measures a person's happiness. Nearly every utility function has three basic properties. The first assumes that the utility function increases with consumption. That is, people are happier with higher levels of consumption. The second standard assumption is that the utility function is increasing at a decreasing rate. That is, a millionaire will not receive the same satisfaction from receiving a ten dollar bill as a poor person would. Finally, we assume that the utility function is time separable. That is, we assume that the utility function is not changing across time. Put in other words, a certain level of consumption will provide the same utility in each time period. In our research, we will also make a representative investor assumption. This means that we assume that there is one investor, who behaves like the average investor in the economy. Obviously, this makes the time separable assumption far more plausible. An accurate utility function is essential because it tells us the level of happiness that an investor will receive from every level of consumption. This will tell us how much an investor is opposed to taking gambles, or how averse he/she is to fluctuations in consumption.

One important indicator of these utility functions is the coefficient of relative risk aversion. The coefficient of relative risk aversion for a given utility function is a measure that tells us how opposed the representative investor is to taking an actuarially fair gamble. It is standard to assume that the utility function exhibits a constant coefficient of relative risk aversion. Another important description of a utility function is the intertemporal elasticity of substitution. This essentially tells us how willing an investor is to substitute consumption across time. A problem with most utility functions, including the power utility function, is that the intertemporal elasticity of substitution is tied to the coefficient of relative risk aversion (i.e.

it is equal to the inverse). There is no reason to expect this since both are measures of two inherently different things (one is a measure of aversion to different levels of consumption and one is a measure of how opposed one is to switching consumption from one period to another). Additionally, empirically we see that these two measures should not be related. Therefore, because of the data, and because of intuition, we use a utility function, the Kreps-Porteus utility function (12.41), from Kreps and Porteus' *Temporal Resolution of Uncertainty and Dynamic Choice Theory* Econometrica paper, with Duffie-Epstein preferences, from Duffie and Epstein's *Stochastic Differential Utility* Econometrica paper, that separates these two measures.

In 1999, Campbell and Cochrane introduced an extremely accurate and successful utility function in their *By Force of Habit, a Consumption-Based Explanation of Aggregate Stock Market Behavior* JPE paper. Campbell and Cochrane describe happiness as a function of how much the representative investor had consumed relative to his/her past consumption. This utility function is commonly described as the "keeping up with the Jones" utility function. However, this utility function is problematic in that it is backward looking. We will use a utility function that is forward looking. We will still incorporate the past into our function, but, the past will be used to form our expectation for the future so that utility will be a function, called the aggregator function, of consumption today and expected future consumption. In this sense, utility will be a recursive function. This means that utility today will be a function of consumption today and expected utility tomorrow. Using the tools of mathematics, we show how this recursive utility function leads us to a differential equation. And, again, using the tools of mathematics and computer programming, we solve this differential equation.

Next, we will try to show how the price-dividend ratio comes into play. We start with the investor maximization problem. This problem essentially states that an investor is trying to maximize his expected utility. We represent this maximum, say M, as an infinite sum where utility is measured each period and each period to the future is given a certain discount factor, because the future does not matter quite as much as the present. Through the use of advanced mathematics, we work through this in the deterministic case. It can be shown that this maximization problem can be rewritten as a functional equation where M at a time t is equal to the utility from consumption today plus M at time t plus one. As one might expect and we prove that this is acceptable, we solve for the optimal condition by looking at the first order condition. The first order condition is what is normally referred to as the Euler equation. We include multiple proofs of this equation including an "intuitive proof". Through mathematics, we show how this Euler equation will lead us to a differential equation for the price-dividend ratio.

In order to use the Euler condition, we first must choose our consumption and dividend processes. In 2004, Ravi Bansal (Duke) and Amir Yaron (Wharton) modeled consumption and dividend growth in their *Risks for the Long Run: A Potential Resolution of Asset Pricing Puzzle* Journal of Finance paper. The Bansal and Yaron model of consumption and dividend growth explained key financial phenomena including the equity premium puzzle. The equity premium puzzle explores the paradox of why potential shareholders do not invest in stocks despite the undisputed fact that stocks have had much higher annual growth than other investments. Is the deterrent risk to capital? Does the allure of guaranteed preservation of capital in other financial instruments such as treasuries offset the unattractive lower rate of annual growth? Given the risk aversion coefficient, those are obviously insufficient responses. To address such insufficient responses, the key component of the Bansal/Yaron model is a multi-dimensional consumption process which presents a long-run process with an expected return. However, within the model, to humanize it, the expected growth rate of consumption was also subjected to a trend and random motion yielding the empirical result.

In my research, I applied this empirical result of the Bansal and Yaron model to the aforementioned model to determine the price-dividend ratio as a function of that consumption process. The ratio within that crafted model, the price-dividend ratio simply compares the price of a stock to the dividends that are paid on the stock. My crafted model sought to determine the relationship of the multi dimensional consumption process to the movement of the price dividend ratio. Because the Bansal/Yaron model's consumption process suggests startling results, including a potential explanation of the equity premium puzzle, I used their model to determine if it is aligned with, or compatible to, the observed behavior of the price to dividend ratio. My research, in order to access this price to dividend function, solves three differential equations which analyze how things change in relation to each other. Finally, it proposes a well-defined problem, that is, one open to solution, specifically a fourth differential equation, second-order linear partial differential equation, to solve for this final price to dividend function. This is interesting research because the results to this well-defined problem may help explain some of the most intractable financial questions in existence. For example, once a determination of the price-dividend ratio has been made, utilizing the model to find option prices would follow closely on its heels and the model would then, theoretically, be able to access and explain the observed behavior of options.

2. Stochastic Processes and Setting up the Equation for the One-Dimensional Case

A stochastic process will represent a drift term, the process will drift in this direction over time, and a Brownian motion. $d\omega$ will represent Brownian motion and dt will represent the drift term. (Please see 11.1 for a discussion of Brownian motion and the Brownian motion multiplication rules, which we will use throughout our research.) We begin with the two stochastic processes

$$dc = (x + \bar{x})dt + \sigma d\omega_1, \tag{2.1}$$

and

$$dx = (\rho - 1)xdt + \varphi_e \sigma d\omega_2. \tag{2.2}$$

c represents consumption and x represents the stochastic process for expected consumption, the drift term. $\bar{x}, \sigma, \rho$, and φ_e are all constants. It is clear that dx goes to zero since it has a negative drift if x is positive and a positive drift if x is negative. Therefore, we shall just consider $dx = 0$ in our initial condition, and $x =$ some constant. Without loss of generality (We can just change the constant $\bar{x}$) we will consider this constant to be zero. Therefore, we are left with the stochastic process:

$$dc = \bar{x}dt + \sigma d\omega_1. \tag{2.3}$$

and since $C = e^c$ this implies that (by Ito's Lemma (11.56)):

$$\frac{dC}{C} = (\bar{x} + \frac{1}{2}\sigma^2)dt + \sigma d\omega_1. \tag{2.4}$$

Essentially, and intuitively, what we are doing is taking a snapshot of this consumption process when $x_{i,t}, dx = 0$. We are essentially studying changes in C, and hence c, at this point. This is the one-dimensional process that we refer to in our introduction.

3. The Stochastic Process for the Pricing Kernel

We begin with the following relationship (Please see 12.3 for a derivation):

$$E_t[dH(C)] + \beta\frac{\mu_\rho(C(t)) - g(H(C))}{g'(H(C))}dt = 0. \tag{3.1}$$

where we define:

$$\mu_\rho(C(t)) = \frac{C(t)^\rho}{\rho}, \tag{3.2}$$

and

$$g(H(C)) = \frac{((1-\gamma)H)^{\frac{\rho}{1-\gamma}}}{\rho}. \tag{3.3}$$

This implies that(remember to use chain rule):

$$g'(H(C)) = ((1-\gamma)H)^{\frac{\rho}{1-\gamma}-1}. \tag{3.4}$$

Applying Ito's Lemma (11.56) to $H(C)$, we see that:

$$E_t[\frac{\partial H}{\partial C}(dC) + \frac{1}{2}\frac{\partial^2 H}{\partial C^2}(dC)^2] + \beta\frac{\mu_p(C(t))dt - g(H(C))}{g'(H(C))}dt = 0. \tag{3.5}$$

Substituting, dividing by dt, and using the fact that $E[dw] = 0$, we see that this equation is equivalent to:

$$H'(\bar{x} + \frac{1}{2}\sigma^2)C + \frac{1}{2}H''(C\sigma)^2 + \frac{\beta}{\rho}(1-\gamma)\left[\frac{C^\rho}{((1-\gamma)H)^{\rho/(1-\gamma)}} - 1\right]H. \tag{3.6}$$

We now propose the following change of variable:

$$H(C) = \frac{g(C)^{(1-\gamma)/\rho}}{1-\gamma}. \tag{3.7}$$

Calculating the derivatives $H'(C)$ and $H''(C)$:

$$H'(C) = \frac{1-\gamma}{\rho}\frac{g'(C)}{g(C)}H(C) \tag{3.8}$$

$$H''(C) = \frac{1-\gamma}{\rho}\left[\frac{g''(C)}{g(C)} + \frac{1-\gamma-\rho}{\rho}\frac{g'(C)^2}{g(C)^2}\right]H(C). \tag{3.9}$$

Plugging this back into our original differential equation and dividing by H:

$$\frac{\sigma^2}{2}C^2\frac{g''(C)}{g(C)} + \frac{\sigma^2}{2}\frac{1-\gamma-\rho}{\rho}C^2\frac{g'(C)^2}{g(C)^2} + (\bar{x} + \frac{1}{2}\sigma^2)C\frac{g'(C)}{g(C)} + \beta\frac{C^\rho}{g(C)} = \beta. \tag{3.10}$$

Now, we re-introduce the variable c, which we defined earlier according to the following relation: $C = e^c$. Use the chain rule to get

$$\frac{dg}{dc} = \frac{dg}{dC}\frac{dC}{dc} = e^c\frac{dg}{dC} \tag{3.11}$$

and

$$\frac{d^2g}{dc^2} = e^c\frac{dg}{dC} + e^c\frac{d^2g}{dC^2}\frac{dC}{dc} = e^c\frac{dg}{dC} + e^{2c}\frac{d^2g}{dC^2}. \tag{3.12}$$

Solve (3.11) and (3.12) for dg/dC and d^2g/dC^2, respectively, to yield

$$\frac{dg}{dC} = e^{-c}\frac{dg}{dc} \quad \text{and} \quad \frac{d^2g}{dC^2} = e^{-2c}\left[\frac{d^2g}{dc^2} - \frac{dg}{dc}\right].$$

The differential equation (3.10) is equivalent to

$$\frac{\sigma^2}{2}\frac{\left[g''(c) - g'(c)\right]}{g(c)} + \frac{\sigma^2}{2}\frac{1-\gamma-\rho}{\rho}\frac{g'(c)^2}{g(c)^2} + (\bar{x} + \frac{1}{2}\sigma^2)\frac{g'(c)}{g(c)} + \beta\frac{e^{\rho c}}{g(c)} = \beta. \tag{3.13}$$

which can be rewritten as the following differential equation

$$g(c)g''(c) = B_1 g'(c)^2 + B_2(g'(c))(g(c)) + B_3 g(c)^2 + B_4(c)g(c). \tag{3.14}$$

where

$$B_1 = -\frac{1-\gamma-\rho}{\rho},\ B_2 = \frac{-2(\bar{x})}{\sigma^2},\ B_3 = \frac{2\beta}{\sigma^2}\ ,$$
$$B_4(c) = -\beta\frac{2e^{\rho c}}{\sigma^2}$$

We now need the initial conditions for this differential equation. Using (5.39), we see that

$$\frac{g'(c)}{g(c)} = (\frac{E_t[R^e(c)] - R^b(c)}{-(\frac{p'(c)}{p(c)} + \varphi_d\rho_{13})\sigma^2} + 1 - \rho)\frac{\rho}{1-\gamma-\rho}. \tag{3.15}$$

Using (5.33), we see that er

$$\frac{g'(c)}{g(c)} = (\frac{E_t[R^e(c)] - R^b(c)}{-(-\rho_{13}\varphi_d \pm \sqrt{\rho_{13}^2\varphi_d^2 - \varphi_d^2 + \frac{\Sigma(\bar{c})^2}{\sigma^2}} + \varphi_d\rho_{13})\sigma^2} + 1 - \rho)\frac{\rho}{1-\gamma-\rho}. \tag{3.16}$$

We standardize $g(\bar{c})$ to be equal to one and then plug in the parameters to get our second initial condition.

$$g_0 = g(\bar{c}) = 1 \tag{3.17}$$

and

$$g_1 = g'(\bar{c}) \tag{3.18}$$

Lemma 1. *The initial value problem* (**??**), (3.17) *and* (3.18) *has an analytic solution $g(c)$ whose power series expansion*

$$g(c) = \sum g_k(c-\bar{c})^k \tag{3.19}$$

has coefficients given by (3.24) *and which converges for*

$$|c - \bar{c}| < 1.2\sigma \tag{3.20}$$

Proof. We now assume that this initial value problem has a power series solution:

$$g(c) = \sum_{n=0}^{\infty} g_n(c-\bar{c})^n \quad \text{where} \tag{3.21}$$

$$g_0 = 1 \text{ and } g_1 = (\frac{E_t[R^e(\bar{c})] - R^b(\bar{c})}{-(\frac{p'(\bar{c})}{p(\bar{c})} + \varphi_d\rho_{13})\sigma^2} + 1 - \rho)\frac{\rho}{1-\gamma-\rho}. \tag{3.22}$$

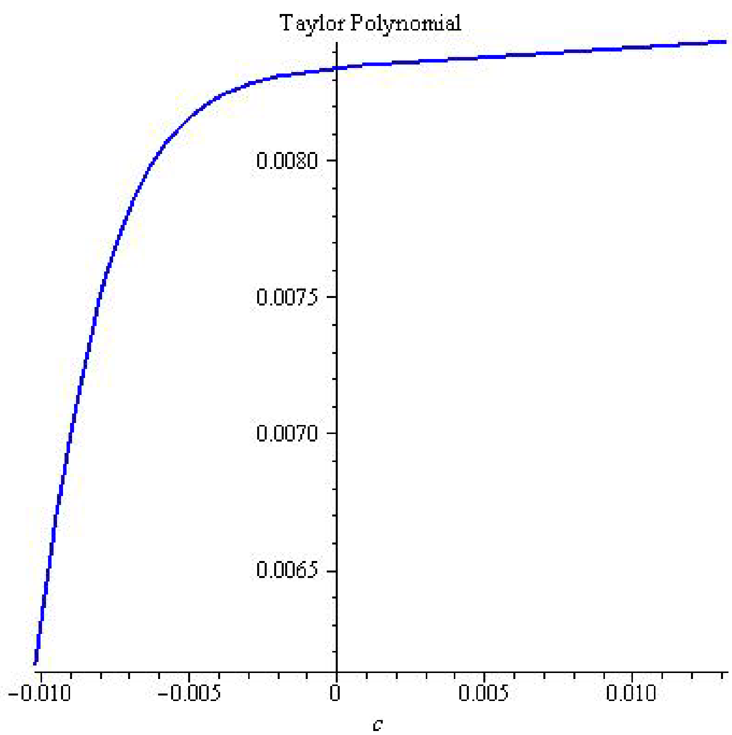
Find the power series representations for the coefficients in the differential equation (3.13):

$$\begin{aligned}
g'(c) &= \sum_{n=0}^{\infty}(n+1)g_{n+1}(c-\bar{c})^n\,,\\
g''(c) &= \sum_{n=0}^{\infty}(n+1)(n+2)g_{n+2}(c-\bar{c})^n\,,\\
g''(c)g(c) &= \sum_{n=0}^{\infty}\sum_{k=0}^{n}(k+1)(k+2)g_{k+2}g_{n-k}(c-\bar{c})^n\,,\\
g(c)^2 &= \sum_{n=0}^{\infty}\sum_{k=0}^{n} g_k g_{n-k}(c-\bar{c})^n\,,\\
g'(c)g(c) &= \sum_{n=0}^{\infty}\sum_{k=0}^{n}(k+1)g_{k+1}g_{n-k}(c-\bar{c})^n\,,\\
g'(c)^2 &= \sum_{n=0}^{\infty}\sum_{k=0}^{n}(k+1)(n-k+1)g_{k+1}g_{n-k+1}(c-\bar{c})^n\,,\\
\frac{-2\beta}{\sigma^2}e^{\rho c}g(c) &= \frac{-2\beta}{\sigma^2}e^{\rho\bar{c}}\sum_{n=0}^{\infty}\sum_{k=0}^{n}\frac{(\rho)^k}{k!}g_{n-k}(c-\bar{c})^n\,.
\end{aligned} \tag{3.23}$$

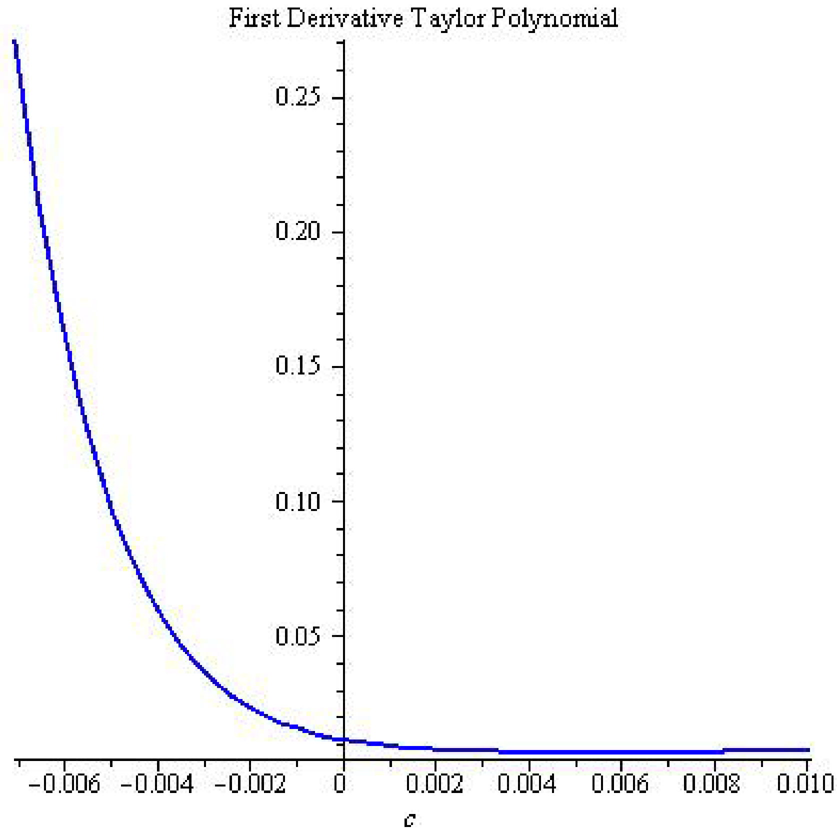
Substitute all the equations in (3.21) and (3.23) into the differential equation (3.13). Equate the coefficients of the terms of degree n and solve the equation for $(n+1)(n+2)g_0g_{n+2}$. The recurrence relation for the g_n's is given by

$$
\begin{aligned}
&(n+1)(n+2)g_0g_{n+2} \\
&= \sum_{k=0}^{n}\left\{\frac{2\beta}{\sigma^2}\left[g_k - e^{\rho\bar{c}}\frac{(\rho)^k}{k!}\right]g_{n-k} + (k+1)\left[-\frac{2\bar{x}}{\sigma^2}g_{n-k} - \frac{1-\gamma-\rho}{\rho}(n-k+1)g_{n-k+1}\right]g_{k+1}\right\} \\
&+ \sum_{k=0}^{n-1}(k+1)\Big[-(k+2)g_{k+2}g_{n-k}\Big] \qquad (3.24)
\end{aligned}
$$

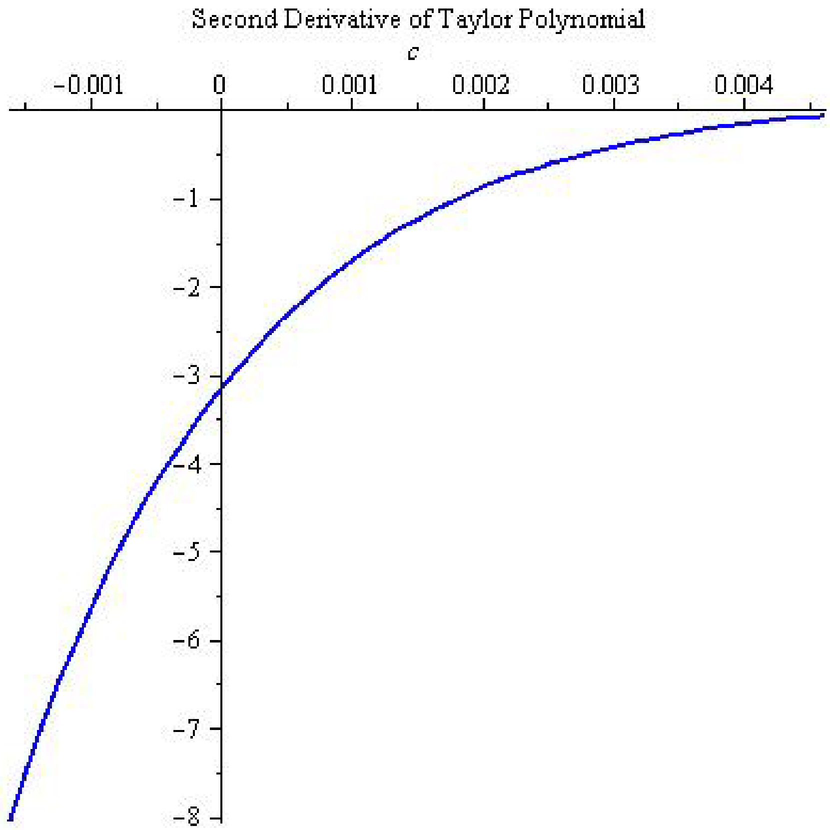
for $n = 0, 1, 2, \ldots$. Having solved for these coefficients, I present the resulting graph depicting a functional form of the utility function, which should have the same properties as the utility function, in one-dimension:



The following graph is of the derivative of the above function. Standard economic theory suggests that as consumption increases utility increases. Obviously, this indicates, for a differentiable utility function, that the derivative will be positive. This is clearly the result we notice in the graph below.



This graph is of the second derivative. Standard economic theory suggests that the second derivative of a differentiable utility function should be negative. (The second bite is not worth as much as the first.) Our utility function clearly has the negative second-derivative property one would expect from the utility function.



3.1. **Convergence and error analysis.** Please note that this subsection is essentially a modification of Yu Chen's "Asset Pricing Model with Duffie-Epstein Preferences" notes.

Lemma 2. *Let $A \geq 0$ be a real number. If b and d are nonnegative integers with $b + d > 0$, then*

$$\lim_{n\to\infty} \sum_{k=0}^{n-b} \frac{A^k}{k!} \cdot \frac{1}{n-k+d} = 0\,. \tag{3.25}$$

Proof. Since $\lim_{k\to\infty} A^k/(k-1)! = 0$, there is an integer $a > 0$ such that $A^k/(k-1)! \leq 1$ for all $k \geq a$.

$$\begin{aligned}
0 \leq \sum_{k=0}^{n-b} \frac{A^k}{k!} \cdot \frac{1}{n-k+d} &= \sum_{k=0}^{a-1} \frac{A^k}{k!} \cdot \frac{1}{n-k+d} + \sum_{k=a}^{n-b} \frac{A^k}{k!} \cdot \frac{1}{n-k+d} \\
&\leq \frac{1}{n-a+1+d} \sum_{k=0}^{a-1} \frac{A^k}{k!} + \sum_{k=a}^{n-b} \frac{A^k}{(k-1)!} \cdot \frac{1}{k(n-k+d)} \\
&\leq \frac{e^A}{n-a+1+d} + \sum_{k=a}^{n-b} \frac{1}{k(n-k+d)}
\end{aligned}$$

We know that

$$\lim_{n\to\infty} \sum_{k=a}^{n-b} \frac{1}{k(n-k+d)} = 0\,.$$

Hence the required equation follows from the Squeeze Theorem. $\square$

Rewrite the recurrence relation (3.24) in the following equivalent form:

$$
\begin{aligned}
(n+1)(n+2)g_0 g_{n+2} = {} & C_0 \sum_{k=0}^{n} g_k g_{n-k} - (D_0 g_0) \sum_{k=0}^{n} [(\rho)^k / k!] g_{n-k} + B_0 \sum_{k=0}^{n} (k+1) g_{k+1} g_{n-k} \\
& - A_0 \sum_{k=0}^{n} (k+1)(n-k+1) g_{k+1} g_{n-k+1} \\
& - \sum_{k=0}^{n-1} (k+1)(k+2) g_{k+2} g_{n-k} \,. \qquad (3.26)
\end{aligned}
$$

Here,

$$
C_0 = 2\frac{\beta}{\sigma^2} \qquad D_0 = e^{\rho \bar{c}} / g_0 \,, \qquad B_0 = \frac{-2\bar{x}}{\sigma^2} \,, \qquad A_0 = \frac{1-\gamma-\rho}{\rho} \,. \qquad (3.27)
$$

Dividing the equation (3.26) by g_0^2 yields

$$
\begin{aligned}
& (n+1)(n+2)(g_{n+2}/g_0) \\
& = C_0 \sum_{k=0}^{n} (g_k/g_0)(g_{n-k}/g_0) - D_0 \sum_{k=0}^{n} [(\rho)^k/k!](g_{n-k}/g_0) + B_0 \sum_{k=0}^{n} (k+1)(g_{k+1}/g_0)(g_{n-k}/g_0) \\
& - A_0 \sum_{k=0}^{n} (k+1)(n-k+1)(g_{k+1}/g_0)(g_{n-k+1}/g_0) \\
& - \sum_{k=0}^{n-1} (k+1)(k+2)(g_{k+2}/g_0)(g_{n-k}/g_0) \,. \qquad (3.28)
\end{aligned}
$$

When $n \geq 2$, the equation (3.28) is equivalent to

$$
\begin{aligned}
& (n+1)(n+2)(g_{n+2}/g_0) \\
& = C_0 \sum_{k=1}^{n-1} (g_k/g_0)(g_{n-k}/g_0) - D_0 \sum_{k=0}^{n-1} [(\rho)^k/k!](g_{n-k}/g_0) + B_0 \sum_{k=0}^{n-1} (k+1)(g_{k+1}/g_0)(g_{n-k}/g_0) \\
& - A_0 \sum_{k=0}^{n} (k+1)(n-k+1)(g_{k+1}/g_0)(g_{n-k+1}/g_0) \\
& - \sum_{k=0}^{n-1} (k+1)(k+2)(g_{k+2}/g_0)(g_{n-k}/g_0) \\
& + C_0(g_n/g_0) - D_0(\rho)^n/n! + B_0(n+1)(g_{n+1}/g_0) \,. \qquad (3.29)
\end{aligned}
$$

Define $\tilde{g}_n = n^2 L(g_n/g_0)$ for $n \geq 1$, where L is a positive number to be determined later. When $n \geq 2$, we can rewrite the equation (3.29) in terms of $\tilde{g}_n$.

$$\begin{aligned}
\tilde{g}_{n+2} = \frac{n+2}{n+1}\Bigg\{ & \frac{C_0}{L}\sum_{k=1}^{n-1}\frac{1}{k^2(n-k)^2}\tilde{g}_k\tilde{g}_{n-k} - D_0\sum_{k=0}^{n-1}\frac{(\rho)^k}{k!}\cdot\frac{1}{(n-k)^2}\tilde{g}_{n-k} \\
& + \frac{B_0}{L}\sum_{k=0}^{n-1}\frac{1}{(k+1)(n-k)^2}\tilde{g}_{k+1}\tilde{g}_{n-k} - \frac{A_0}{L}\sum_{k=0}^{n}\frac{1}{(k+1)(n-k+1)}\tilde{g}_{k+1}\tilde{g}_{n-k+1} \\
& - \frac{1}{L}\sum_{k=0}^{n-1}\frac{k+1}{(k+2)(n-k)^2}\tilde{g}_{k+2}\tilde{g}_{n-k} \\
& + \frac{C_0}{n^2}\tilde{g}_n - \frac{LD_0(\rho)^n}{n!} + \frac{B_0}{n+1}\tilde{g}_{n+1}\Bigg\}
\end{aligned} \tag{3.30}$$

Note that $\sum_{k=1}^{\infty}(1/k^2) = \pi^2/6$. By our lemma and our knowledge of power series, we know that we can find a real number $L > 1$ and an integer $N \geq 2$ such that for all $n \geq N$, we have

$$\begin{aligned}
& \frac{n+2}{n+1}\Bigg[\frac{|C_0|}{L}\sum_{k=1}^{n-1}\frac{1}{k^2(n-k)^2} + |D_0|\sum_{k=0}^{n-1}\frac{|\rho|^k}{k!}\cdot\frac{1}{(n-k)^2} + \frac{|B_0|}{L}\sum_{k=0}^{n-1}\frac{1}{(k+1)(n-k)^2} \\
& + \frac{|A_0|}{L}\sum_{k=0}^{n}\frac{1}{(k+1)(n-k+1)} + \frac{1}{L}\sum_{k=0}^{n-1}\frac{k+1}{(k+2)(n-k)^2} \\
& + \frac{|C_0|}{n^2} + \frac{L|D_0||\rho|^n}{n!} + \frac{|B_0|}{n+1}\Bigg] < 1\,.
\end{aligned} \tag{3.31}$$

Apply the following algorithm to construct a sequence $\{G_n\}$ of nonnegative real numbers.

(1) Use the recurrence relation (3.24) and the initial values g_0, g_1 to calculate g_n, where $2 \leq n \leq N+1$.
(2) Calculate $G_n = n^2 L|g_n/g_0|$ for $1 \leq n \leq N+1$.
(3) Calculate G_{n+2}, where $n \geq N$, by using the recurrence relation:

$$\begin{aligned}
G_{n+2} = \frac{n+2}{n+1}\Bigg[& \frac{|C_0|}{L}\sum_{k=1}^{n-1}\frac{1}{k^2(n-k)^2}G_kG_{n-k} + |D_0|\sum_{k=0}^{n-1}\frac{|\rho|^k}{k!}\cdot\frac{1}{(n-k)^2}G_{n-k} \\
& + \frac{|B_0|}{L}\sum_{k=0}^{n-1}\frac{1}{(k+1)(n-k)^2}G_{k+1}G_{n-k} + \frac{|A_0|}{L}\sum_{k=0}^{n}\frac{1}{(k+1)(n-k+1)}G_{k+1}G_{n-k+1} \\
& + \frac{1}{L}\sum_{k=0}^{n-1}\frac{k+1}{(k+2)(n-k)^2}G_{k+2}G_{n-k} \\
& + \frac{|C_0|}{n^2}G_n + \frac{L|D_0||\rho|^n}{n!} + \frac{|B_0|}{n+1}G_{n+1}\Bigg\}\,.
\end{aligned} \tag{3.32}$$

Let $M_g \geq 1$ be such that $G_n \leq M_g^n$ for $1 \leq n \leq N+1$. By mathematical induction, we can show that

$$n^2L|g_n/g_0| \leq G_n \leq M_g^n \quad \text{or} \quad |g_n| \leq \frac{|g_0|}{L}\cdot\frac{M_g^n}{n^2} \quad \text{for } n \geq 1\,. \tag{3.33}$$

Lemma 3. *Choose a real number $L \geq 1$ and an integer $N \geq 2$ such that the inequality (3.31) holds for $n \geq N$ and set*

$$M_g = \max\{1, \sqrt[n]{n^2 L |g_n/g_0|} \colon 1 \leq n \leq N+1\} \quad \textit{and} \quad r_g = 1/M_g\,. \tag{3.34}$$

The power series solution $g(c) = \sum_{n=0}^{\infty} g_n (c-\bar{c})^n$ of the initial value problem (3.14)-(3.18) converges in the open interval $\bar{c} - r_g < c < \bar{c} + r_g$, where the g_n's are determined by the recurrence relation given in (3.24).

Our error analysis, done in Maple, leads us to the conclusion that the utility function converges in a radius of convergence equal to 1.2 times the standard deviation.

4. The Discount Factor

We start with the solution H(we arrive at this by solving for G, and then solving back for H through our change of variable) from this ODE. This H satisfies

$$H(C) = U_t = E_t \left[\int_t^{\infty} \bar{f}(C_s, U_s)\, ds \right]. \tag{4.1}$$

We will derive an ordinary differential equation for the equilibrium price-dividend ratio in the DEKP model, given the power series of the lifetime utility function $H(C)$ around $C = \bar{C}$. The pricing kernel for the investor is given by

$$\Lambda(C,U,t) = e^{-\delta t} \frac{\partial \bar{f}}{\partial C}(C,U) = \frac{\partial}{\partial C}\left[\frac{\beta}{\rho} \cdot \frac{C^{\rho} - ((1-\gamma)U)^{\rho/(1-\gamma)}}{((1-\gamma)U)^{\rho/(1-\gamma)-1}} \right] = \frac{\beta C^{\rho-1}}{((1-\gamma)U)^{\rho/(1-\gamma)-1}}\,. \tag{4.2}$$

The first-order partial derivatives of $\Lambda(C,U,t)$ (Note: when we take derivatives with respect to C, we are considering C not contained in U) are:

$$\frac{\partial \Lambda}{\partial C} = (\rho - 1)\frac{\Lambda}{C} \quad \text{and} \quad \frac{\partial \Lambda}{\partial U} = \frac{1-\gamma-\rho}{1-\gamma}\frac{\Lambda}{U}\,. \quad \text{and} \quad \frac{\partial \Lambda}{\partial t} = (-\delta)\Lambda \tag{4.3}$$

The second-order partial derivatives of $\Lambda(C,U)$ are

$$\begin{aligned} \frac{\partial^2 \Lambda}{\partial C^2} &= (\rho-1)(\rho-2)\frac{\Lambda}{C^2}\,, \\ \frac{\partial^2 \Lambda}{\partial U^2} &= -\frac{\rho(1-\gamma-\rho)}{(1-\gamma)^2}\frac{\Lambda}{U^2}\,, \\ \frac{\partial^2 \Lambda}{\partial C \partial U} &= \frac{(\rho-1)(1-\gamma-\rho)}{1-\gamma}\frac{\Lambda}{CU}\,. \end{aligned} \tag{4.4}$$

Invoking Ito's Lemma (11.56), we note that:

$$dU = H'(C)dC + \frac{1}{2}H"(C)(dC)^2 \tag{4.5}$$

which implies that:

$$(dU)^2 = (H'(C))^2 (dC)^2 \tag{4.6}$$

and

$$(dU)(dC) = H'(C)(dC)^2 \tag{4.7}$$

By Ito's Lemma, the pricing kernel $\Lambda(t, C, U)$ follows the stochastic process(Remember that $U = H(C)$):

$$d\Lambda = \frac{\partial\Lambda}{\partial t}dt + \frac{\partial\Lambda}{\partial C}(dC) + \frac{\partial\Lambda}{\partial U}(dU) + \frac{1}{2}\frac{\partial^2\Lambda}{\partial C^2}(dC)^2 + \frac{1}{2}\frac{\partial^2\Lambda}{\partial U^2}(dU)^2 + \frac{\partial^2\Lambda}{\partial U\partial C}(dU)(dC) \tag{4.8}$$

This implies that:

$$\begin{aligned}
d\Lambda &= \frac{\partial\Lambda}{\partial t}dt + \frac{\partial\Lambda}{\partial C}(dC) + \frac{\partial\Lambda}{\partial U}(H'(C)dC + \frac{1}{2}H"(C)(dC)^2) \\
&+ \frac{1}{2}\frac{\partial^2\Lambda}{\partial C^2}(dC)^2 + \frac{1}{2}\frac{\partial^2\Lambda}{\partial U^2}(H'(C))^2(dC)^2 + \frac{\partial^2\Lambda}{\partial U\partial C}(dU)(dC) \qquad (4.9)\\
&= (-\delta)\Lambda dt + (\rho-1)\frac{\Lambda}{C}(dC) + \frac{1-\gamma-\rho}{1-\gamma}\frac{\Lambda}{U}(H'(C)dC + \frac{1}{2}H"(C)(dC)^2) \\
&+ \frac{1}{2}(\rho-1)(\rho-2)\frac{\Lambda}{C^2}(dC)^2 + \frac{-1}{2}\frac{\rho(1-\gamma-\rho)}{(1-\gamma)^2}\frac{\Lambda}{U^2}(H'(C))^2(dC)^2 \\
&+ \frac{(\rho-1)(1-\gamma-\rho)}{1-\gamma}\frac{\Lambda}{CU}H'(C)(dC)^2 \qquad (4.10)\\
&= (-\delta)\Lambda dt + ((\rho-1)\frac{\Lambda}{C} + \frac{1-\gamma-\rho}{1-\gamma}\frac{\Lambda}{U}H'(C))dC + (\frac{1}{2}\frac{1-\gamma-\rho}{1-\gamma}\frac{\Lambda}{U}H"(C) + \\
&\frac{1}{2}(\rho-1)(\rho-2)\frac{\Lambda}{C^2} + \frac{-1}{2}\frac{\rho(1-\gamma-\rho)}{(1-\gamma)^2}\frac{\Lambda}{U^2}(H'(C))^2 + \frac{(\rho-1)(1-\gamma-\rho)}{1-\gamma}\frac{\Lambda}{CU}H'(C))(dC)^2
\end{aligned} \tag{4.11}$$

Divide Eq.(4.9) by Λ and use the relation $U = H(C)$ to get

$$\begin{aligned}
\frac{d\Lambda}{\Lambda} &= (-\delta)dt + ((\rho-1)\frac{1}{C} + \frac{1-\gamma-\rho}{1-\gamma}\frac{1}{H(C)}H'(C))dC + (\frac{1-\gamma-\rho}{1-\gamma}\frac{1}{H(C)}\frac{1}{2}H"(C) + \\
&\frac{1}{2}(\rho-1)(\rho-2)\frac{1}{C^2} + \frac{-1}{2}\frac{\rho(1-\gamma-\rho)}{(1-\gamma)^2}\frac{1}{H(C)^2}(H'(C))^2 + \frac{(\rho-1)(1-\gamma-\rho)}{1-\gamma}\frac{1}{CH(C)}H'(C))(dC)^2
\end{aligned} \tag{4.12}$$

From our Brownian motion multiplication rules (11.57), we know that:

$$(dC)^2 = C^2\sigma^2 dt \tag{4.13}$$

Use Ito's rule for multiplication, $(dt)^2 = 0$, $dtd\omega_1 = 0$, $d\omega_1 d\omega_2 = adt$, where a represents the correlation between the first and second Brownian motion, $dtd\omega_2 = 0$ and $(d\omega_1)^2 = dt$, to get

$$\begin{aligned}
\frac{d\Lambda}{\Lambda} &= (-\delta + ((\rho-1)\frac{1}{C} + \frac{1-\gamma-\rho}{1-\gamma}\frac{1}{H(C)}H'(C))(\bar{x} + \frac{\sigma^2}{2})C + \Bigg(\frac{1-\gamma-\rho}{1-\gamma}\frac{1}{H(C)}\frac{1}{2}H"(C) + \\
&\frac{1}{2}(\rho-1)(\rho-2)\frac{1}{C^2} + \frac{-1}{2}\frac{\rho(1-\gamma-\rho)}{(1-\gamma)^2}\frac{1}{H(C)^2}(H'(C))^2 \\
&+ \frac{(\rho-1)(1-\gamma-\rho)}{1-\gamma}\frac{1}{CH(C)}(H'(C))\Bigg)(C^2\sigma^2)dt \\
&+ ((\rho-1)\frac{1}{C} + \frac{1-\gamma-\rho}{1-\gamma}\frac{1}{H(C)}H'(C))C\sigma d\omega_1
\end{aligned} \tag{4.14}$$

We now make the change of variable $C = e^c$. We can use the chain rule to get

$$\frac{dH}{dc} = \frac{dH}{dC}\frac{dC}{dc} = e^c \frac{dH}{dC} \tag{4.15}$$

and

$$\frac{d^2H}{dc^2} = e^c \frac{dH}{dC} + e^c \frac{d^2H}{dC^2}\frac{dC}{dc} = e^c \frac{dH}{dC} + e^{2c}\frac{d^2H}{dC^2}\,. \tag{4.16}$$

Solve (4.15) and (4.16) for dG/dC and d^2G/dC^2, respectively, to yield

$$\frac{dH}{dC} = e^{-c}\frac{dH}{dc} \quad \text{and} \quad \frac{d^2H}{dC^2} = e^{-2c}\left[\frac{d^2H}{dc^2} - \frac{dH}{dc}\right]\,.$$

Plugging these values back in, we arrive at the new expression:

$$\begin{aligned}\frac{d\Lambda}{\Lambda} &= (-\delta + ((\rho-1) + \frac{1-\gamma-\rho}{1-\gamma}\frac{1}{H(c)}H'(c))(\bar{x} + \frac{\sigma^2}{2}) + \Bigg(\frac{1-\gamma-\rho}{1-\gamma}\frac{1}{H(c)}\frac{1}{2}(H"(c) - H'(c)) + \\ &\quad \frac{1}{2}(\rho-1)(\rho-2) + \frac{-1}{2}\frac{\rho(1-\gamma-\rho)}{(1-\gamma)^2}\frac{1}{H(c)^2}(H'(c))^2 \\ &\quad + \frac{(\rho-1)(1-\gamma-\rho)}{1-\gamma}\frac{1}{H(c)}(H'(c))\Bigg)(\sigma^2)\Bigg)dt \\ &\quad + ((\rho-1) + \frac{1-\gamma-\rho}{1-\gamma}\frac{1}{H(c)}H'(c))\sigma d\omega_1 \end{aligned} \tag{4.17}$$

5. Equilibrium Price-Dividend Function in the DEKP Model

Cochrane (2005) shows that the equilibrium price of stocks satisfies the Euler equation:

$$\Lambda(t)D(t)dt + E_t\left[d(\Lambda(t)P(t))\right] = 0 \tag{5.1}$$

where $P(t)$ is the price of a stock at time t and $D(t)$ is the dividend paid by this stock at time t.

Definition 1. Define the price-dividend ratio to be $p(t) = P(t)/D(t)$.

The Euler condition (8.1) is equivalent to

$$\Lambda(t)D(t)\,dt + E_t\left[d(\Lambda(t)p(t)D(t))\right] = 0\,. \tag{5.2}$$

By Ito's lemma, we have

$$\frac{d(\Lambda pD)}{\Lambda pD} = \frac{d\Lambda}{\Lambda} + \frac{dp}{p} + \frac{dD}{D} + \frac{d\Lambda dp}{\Lambda p} + \frac{dDdp}{Dp} + \frac{d\Lambda dD}{\Lambda D}\,. \tag{5.3}$$

The Euler condition (8.1) is equivalent further to

$$\frac{1}{p}\,dt + E_t\left[\frac{d\Lambda}{\Lambda} + \frac{dp}{p} + \frac{dD}{D} + \frac{d\Lambda dp}{\Lambda p} + \frac{dDdp}{Dp} + \frac{d\Lambda dD}{\Lambda D}\right] = 0\,. \tag{5.4}$$

We are given a process for dividends:

$$\frac{dD}{D} = (\phi x + \bar{x})dt + \varphi_d \sigma d\omega_3. \tag{5.5}$$

Since we are just dealing with initial conditions ($x = 0$) we rewrite this process as:

$$\frac{dD}{D} = (\bar{x})dt + \varphi_d \sigma d\omega_3. \tag{5.6}$$

We seek a price-dividend function of the form $p = p(c)$ that represents the equilibrium behavior of the stock price when the investor has DEKP preferences.

By Ito's rule, we have

$$dp = p'(c)dc + \frac{1}{2}p"(c)(dc)^2 \tag{5.7}$$

$$dc^2 = \sigma^2 dt \tag{5.8}$$

This implies that:

$$dp = p'(c)((\bar{x})dt + \sigma d\omega_1) + \frac{1}{2}p"(c)\sigma^2 dt \tag{5.9}$$

Which implies that:

$$\frac{dp}{p} = (\frac{p'(c)}{p(c)}(\bar{x}) + \frac{p"(c)}{p(c)}\sigma^2)dt + \frac{p'(c)\sigma}{p(c)}d\omega_1 \tag{5.10}$$

Now, we calculate $E_t[d\Lambda/\Lambda]$, $E_t[dp/p]$, $E_t[dD/D]$, $E_t[d\Lambda dp/\Lambda p]$, $E_t[dDdp/Dp]$, and $E_t[d\Lambda dD/\Lambda D]$.

$$\begin{aligned}
E_t\left[\frac{d\Lambda}{\Lambda}\right] &= \Bigg(- \delta + ((\rho - 1) + \frac{1-\gamma-\rho}{1-\gamma}\frac{1}{H(c)}H'(c))(\bar{x} + \frac{\sigma^2}{2}) + \Bigg(\frac{1-\gamma-\rho}{1-\gamma}\frac{1}{H(c)}\frac{1}{2}(H"(c) - H'(c)) + \\
&\quad \frac{1}{2}(\rho-1)(\rho-2) + \frac{-1}{2}\frac{\rho(1-\gamma-\rho)}{(1-\gamma)^2}\frac{1}{H(c)^2}(H'(c))^2 \\
&\quad + \frac{(\rho-1)(1-\gamma-\rho)}{1-\gamma}\frac{1}{H(c)}H'(c)\Bigg)(\sigma^2)\Bigg)dt \\
E_t\left[\frac{dp}{p}\right] &= \frac{((p'(c)(\bar{x})) + \frac{\sigma^2}{2}p"(c))dt}{p(c)} \\
E_t\left[\frac{dD}{D}\right] &= \bar{x}\,dt \\
E_t\left[\frac{d\Lambda dp}{\Lambda p}\right] &= ((((\rho-1) + \frac{1-\gamma-\rho}{1-\gamma}\frac{1}{H(c)}H'(c))\frac{p'(c)\sigma^2}{p(c)}))dt \\
E_t\left[\frac{dDdp}{Dp}\right] &= (\frac{\varphi_d\sigma^2 p'(c)}{p(c)}\rho_{13})dt \\
E_t\left[\frac{d\Lambda dD}{\Lambda D}\right] &= (((\rho-1) + (\frac{1-\gamma-\rho}{1-\gamma}\frac{1}{H(c)}H'(c)))\varphi_d\sigma^2\rho_{13})dt
\end{aligned}$$

where ρ_{13} represents the correlation between the first Brownian motion and the third Brownian motion.

By (8.2), we obtain the partial differential equation

$$\frac{1}{p}dt + \Bigg[(-\delta + ((\rho-1) + \frac{1-\gamma-\rho}{1-\gamma}\frac{1}{H(c)}H'(c))(\bar{x} + \frac{\sigma^2}{2})$$
$$+ \Bigg(\frac{1-\gamma-\rho}{1-\gamma}\frac{1}{H(c)}\frac{1}{2}(H''(c) - H'(c)) + \frac{1}{2}(\rho-1)(\rho-2) + \frac{-1}{2}\frac{\rho(1-\gamma-\rho)}{(1-\gamma)^2}\frac{1}{H(c)^2}(H'(c))^2$$
$$+ \frac{(\rho-1)(1-\gamma-\rho)}{1-\gamma}\frac{1}{H(c)}H'(c)\Bigg)(\sigma^2)) +$$
$$(p'(c)(\bar{x}) + \frac{\sigma^2}{2}p''(c))\frac{1}{p(c)} + \bar{x} + (((\rho-1) +$$
$$\frac{1-\gamma-\rho}{1-\gamma}\frac{1}{H(c)}H'(c))\frac{\sigma^2 p'(c)}{p(c)}) +$$
$$+ (\frac{\varphi_d\sigma^2 p'(c)}{p(c)}\rho_{13}) + ((\rho-1) + \frac{1-\gamma-\rho}{1-\gamma}\frac{1}{H(c)}H'(c))\varphi_d\sigma^2\rho_{13}\Bigg]dt = 0 \tag{5.11}$$

Multiplying by $\frac{p}{dt}$, and grouping common terms we see that:
This can be rewritten as(please remember that $C = e^c$):

$$A_1(c) + A_2(c)p(c) + A_3p'(c) + A_4p''(c) = 0 \tag{5.12}$$

where

$$A_1(c) = 1 + ((\rho-1) + (\frac{1-\gamma-\rho}{1-\gamma}\frac{1}{H(c)}H'(c)))\varphi_d\sigma^2\rho_{13} + \bar{x} \tag{5.13}$$

$$A_2(c) = (-\delta + ((\rho-1) + \frac{1-\gamma-\rho}{1-\gamma}\frac{1}{H(c)}H'(c))(\bar{x} + \frac{\sigma^2}{2}) + (\frac{1-\gamma-\rho}{1-\gamma}\frac{1}{H(c)}\frac{1}{2}(H''(c) - H'(c)) +$$
$$\frac{1}{2}(\rho-1)(\rho-2) + \frac{-1}{2}\frac{\rho(1-\gamma-\rho)}{(1-\gamma)^2}\frac{1}{H(c)^2}(H'(c))^2 + \frac{(\rho-1)(1-\gamma-\rho)}{1-\gamma}\frac{1}{H(c)}H'(c))(\sigma^2)) \tag{5.14}$$

$$A_3 = \varphi_d\sigma^2\rho_{13} + ((((\rho-1) + \frac{1-\gamma-\rho}{1-\gamma}\frac{1}{H(c)}H'(c))\sigma^2) + \bar{x} \text{ and } A_4 = \frac{\sigma^2}{2} \tag{5.15}$$

Using (3.7), and its subsequent derivative equivalents, we can transform this differential equation to be of the form:

$$ggp'' = A_1gg + A_2gg' + A_3ggp + A_4gg'p + A_5gg''p + A_6g'g'p + A_7ggp + A_8gg'p' \tag{5.16}$$

where

$$A_1 = -2(1 + ((\rho-1)\varphi_d\sigma\sigma)\rho_{13} + \bar{x})/(\sigma\sigma) \tag{5.17}$$

$$A_2 = -2(\frac{(1-\gamma-\rho)\varphi_d\sigma}{\rho\sigma})\rho_{13}/(\sigma\sigma) \tag{5.18}$$

$$A_3 = -2\frac{(-\delta + (\rho-1)(\bar{x} + (1/2)(\sigma^2)))}{\sigma^2} + (1/2)\frac{((\rho-1)(\rho-2)\sigma^2))}{(\sigma^2)} \tag{5.19}$$

$$A_4 = -2((1-\gamma-\rho)\sigma^2)(\bar{x} + (1/2)(\sigma^2) - 3/2 + \rho)/((\sigma^2)\rho) \tag{5.20}$$

$$A_5 = -\frac{2}{\sigma^2}((\sigma^2)\frac{(1-\gamma-\rho)}{2\rho}) \tag{5.21}$$

$$A_6 = -2(((1-\gamma-\rho)^2)/(2\rho^2) - (1-\gamma-\rho)/(2\rho))/(\sigma^2) \tag{5.22}$$

$$A_7 = \frac{-2((\rho-1)\sigma + \bar{x} + (\varphi_d\sigma)\rho_{13})}{(\sigma^2)} \tag{5.23}$$

$$A_8 = \frac{-2(1-\gamma-\rho)}{\rho} \tag{5.24}$$

The initial conditions for this ODE are

$$p_0 = p(\bar{c}), p_1 = p'(\bar{c}) \tag{5.25}$$

where p represents the price-dividend ratio from the data. To derive the second initial condition, we first recall that the instantaneous return on equity is given by:

$$R^e(c)dt = \frac{dP}{P} + \frac{D}{p}dt \tag{5.26}$$

The price-dividend ratio p implies that $P = pD$, which, using Ito's Lemma, tells us that:

$$R^e(c)dt = \frac{dp}{p} + \frac{dD}{D} + \frac{dp}{p}\frac{dD}{D} + \frac{1}{p}dt \tag{5.27}$$

This implies that:

$$R^e(c)dt = (\frac{p'(c)}{p(c)}(\bar{x}) + \frac{1}{2}p"(c)\sigma^2)dt + \frac{p'(c)\sigma}{p(c)}d\omega_1 + (\bar{x})dt + \varphi_d\sigma d\omega_3 + (\frac{\varphi_d\sigma^2 p'(c)}{p(c)}\rho_{13})dt + \frac{1}{p}dt \tag{5.28}$$

Combining terms, this implies that:

$$R^e(c)dt = (\frac{p'(c)}{p(c)}(\bar{x}) + \frac{1}{2}p"(c)\sigma^2 + \bar{x} + (\frac{\varphi_d\sigma^2 p'(c)}{p(c)}\rho_{13}) + \frac{1}{p})dt + \frac{p'(c)\sigma}{p(c)}d\omega_1 + \varphi_d\sigma d\omega_3 \tag{5.29}$$

This immediately implies (since the expected value of any Brownian motion is zero):

$$E[R^e(c)]dt = (\frac{p'(c)}{p(c)}(\bar{x}) + \frac{1}{2}p"(c)\sigma^2 + \bar{x} + (\frac{\varphi_d\sigma^2 p'(c)}{p(c)}\rho_{13}) + \frac{1}{p})dt \tag{5.30}$$

As a side note, we see immediately that the standard deviation on the return to equity is

$$\Sigma(c) = \sqrt{\frac{p'(c)^2\sigma^2}{p(c)^2} + 2\rho_{13}\frac{p'(c)\sigma}{p(c)}\varphi_d\sigma + \varphi_d^2\sigma^2}, \tag{5.31}$$

which, squaring both sides, implies that

$$\Sigma(c)^2 = \frac{p'(c)^2\sigma^2}{p(c)^2} + 2\rho_{13}\frac{p'(c)\sigma}{p(c)}\varphi_d\sigma + \varphi_d^2\sigma^2. \tag{5.32}$$

Using the quadratic formula, we see that

$$\frac{p'(c)}{p(c)} = -\rho_{13}\varphi_d \pm \sqrt{\rho_{13}^2\varphi_d^2 - \varphi_d^2 + \frac{\Sigma(c)^2}{\sigma^2}}. \tag{5.33}$$

Thus, we arrive at our second initial condition

$$p_1 = p'(\bar{c}) = p_0(-\rho_{13}\varphi_d \pm \sqrt{\rho_{13}^2\varphi_d^2 - \varphi_d^2 + \frac{\Sigma(\bar{c})^2}{\sigma^2}}). \tag{5.34}$$

From (5.27) we see that the expected return on equity is given by the equation:

$$E_t[R^e(c)]dt = \frac{dt}{p} + E_t\left[\frac{dp}{p} + \frac{dD}{D} + \frac{dDdp}{Dp}\right]. \tag{5.35}$$

Recall that $-R^b(c)dt = E_t[d\Lambda/\Lambda]$. This implies that:

$$\left[E_t[R^e(c)] - R^b(c)\right]dt = \frac{dt}{p} + E_t\left[\frac{d\Lambda}{\Lambda} + \frac{dp}{p} + \frac{dD}{D} + \frac{dDdp}{Dp}\right] \tag{5.36}$$

From (8.2), we see that:

$$\begin{aligned}\left[E_t[R^e(c)] - R^b(c)\right]dt &= \frac{dt}{p} + E_t\left[\frac{d\Lambda}{\Lambda} + \frac{dp}{p} + \frac{dD}{D} + \frac{dDdp}{Dp}\right] \\ &= -E_t\left[\frac{d\Lambda dp}{\Lambda p} + \frac{d\Lambda dD}{\Lambda D}\right].\end{aligned} \tag{5.37}$$

So

$$\begin{aligned}E_t[R^e(c)] - R^b(c) = &-(((\rho-1) + \frac{1-\gamma-\rho}{1-\gamma}\frac{1}{H(c)}H'(c))\frac{p'(c)\sigma^2}{p(c)}) \\ &-((\rho-1) + (\frac{1-\gamma-\rho}{1-\gamma}\frac{1}{H(c)}H'(c)))\varphi_d\sigma^2\rho_{13},\end{aligned} \tag{5.38}$$

which, using (3.8), we see that

$$E_t[R^e(c)] - R^b(c) = -((\rho-1) + (\frac{1-\gamma}{\rho} - 1)\frac{g'(c)}{g(c)})(\frac{p'(c)}{p(c)} + \varphi_d\rho_{13})\sigma^2 \tag{5.39}$$

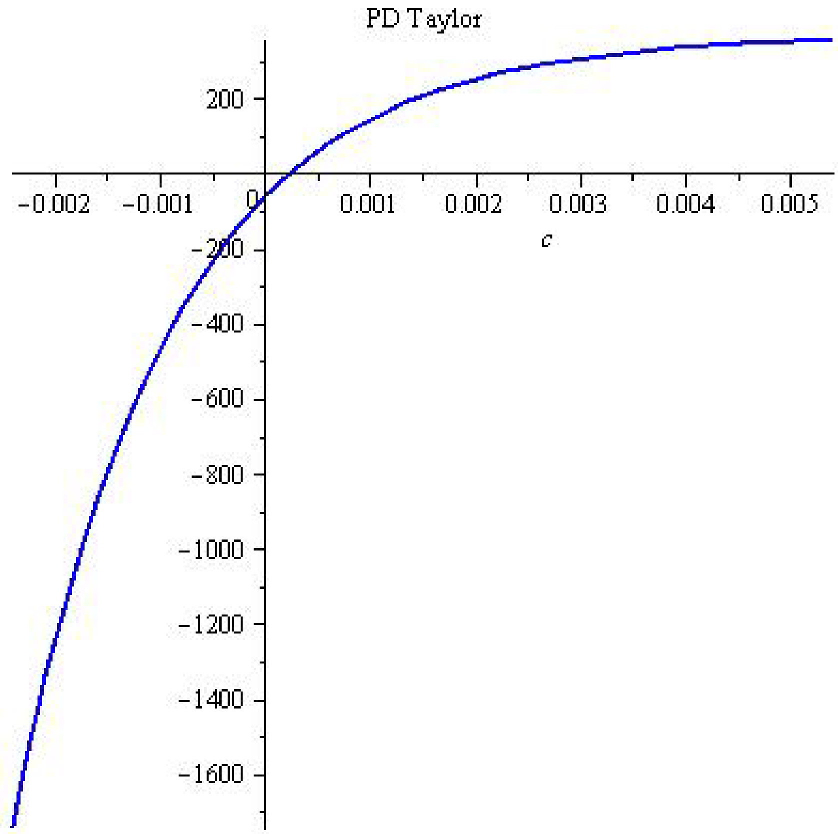
Lemma 4. *The initial value problem* (**??**), (5.25) *and* (**??**) *has an analytic solution* $p(c)$ *whose power series expansion*

$$p(c) = \sum p_k(c-\bar{c})^k \tag{5.40}$$

has coefficients given by (5.41)

$$\begin{aligned}(n+1)(n+2)g_0^2p_{n+2} = {}& A_1\sum_{k=0}^{N}\sum_{j=0}^{k} g_{k-j}g_j + A_2\sum_{k=0}^{N}\sum_{j=0}^{k}(j+1)g_{k-j}g_{j+1} + A_3\sum_{k=0}^{n}\sum_{j=0}^{k} p_{n-k}g_{k-j}g_j + \\ & A_4\sum_{k=0}^{n}\sum_{j=0}^{k} p_{n-k}(j+1)g_{k-j}g_{j+1} + A_5\sum_{k=0}^{n}\sum_{j=0}^{k}(j+2)(j+1)p_{n-k}g_{j+2}g_{k-j} + \\ & A_6\sum_{k=0}^{n}\sum_{j=0}^{k}(k-j+1)(j+1)p_{n-k}g_{j+1}g_{k-j+1} + A_7\sum_{k=0}^{n}\sum_{j=0}^{k}(n-k+1)p_{n-k+1}g_{k-j}g_j \\ & + \sum_{k=0}^{n}\sum_{j=0}^{k}(n-k+1)(j+1)p_{n-k+1}g_{k-j}g_{j+1} \\ & - \sum_{k=1}^{n}\sum_{j=0}^{k}(n-k+1)(n-k+2)p_{n-k+2}g_{k-j}g_j.\end{aligned} \tag{5.41}$$

The following graph depicts our results for the price-dividend ratio as a function of consumption. There is a postive relation between the price-dividend ratio and consumption as one might expect.:



6. Ordinary Differential Equation with the DEKP Preferences

We now return to the two-dimensional case. That is, we do not make any assumptions about the x process. Assume that the consumption $C(t)$ of the investor follows the stochastic process:

$$\frac{dC}{C} = dc = (x + \bar{x})dt + \sigma d\omega_1 \tag{6.1}$$

where

$$dx = (\rho - 1)xdt + \varphi_e \sigma d\omega_2. \tag{6.2}$$

We try to get an expression for C in terms of x and randomness. First, we integrate from 0 to t our expression for dx.

$$x(t) - x(0) = (\rho - 1)\int_0^t xdt + \varphi_e \sigma \int_0^t d\omega_2 \tag{6.3}$$

Rearranging terms we see that:

$$\int_0^t xdt = \frac{x(t) - x(0) - \varphi_e \sigma \int_0^t d\omega_2}{\rho - 1} \tag{6.4}$$

Now, integrating (6.1), we see that:

$$ln(C(t)) - ln(C(0)) = \bar{x}t + \int_0^t xdt + \sigma \int_0^t d\omega_1 \tag{6.5}$$

Plugging in (6.4), we arrive at a relationship between C and x

$$ln(C(t)) = ln(C(0)) + \bar{x}t + \frac{x(t) - x(0) - \varphi_e \sigma \int_0^t d\omega_2}{\rho - 1} + \sigma(\omega_1(t) - \omega_1(0)) \tag{6.6}$$

Here $\bar{x} = \ln \bar{C}$. which implies that:

$$C(t) = e^{ln(C(0)) + \bar{x}t + \frac{x(t) - x(0) - \varphi_e \sigma \int_0^t d\omega_2}{\rho - 1} + \sigma(\omega_1(t) - \omega_1(0))} \tag{6.7}$$

At this point, it is clear that C is a function of the state process x, and t, and unless we get rid of one of the Brownian motions, it is clear that we will not be able to write C solely as a function of x.

We now try to derive a partial differential equation for $H(C, x)$. We begin with equation (12.53), and our H function defined earlier:

$$E_t[dH(C, x)] + \beta \frac{\mu_p(C(t)) - g(H(C, x))}{g'(H(C, x))} dt = 0 \tag{6.8}$$

where we define:

$$\mu_p(C(t)) = \frac{C(t)^p}{p}, \tag{6.9}$$

and

$$g(H(C, x)) = \frac{((1 - \gamma)H)^{\frac{\rho}{1-\gamma}}}{\rho}. \tag{6.10}$$

This implies that(remember to use chain rule):

$$g'(H(C, x)) = ((1 - \gamma)H)^{\frac{\rho}{1-\gamma} - 1} \tag{6.11}$$

Applying Ito's Lemma to $H(C, x)$, we see that:

$$E_t[\frac{\partial H}{\partial C}(dC) + \frac{\partial H}{\partial x} dx + \frac{1}{2}\frac{\partial^2 H}{\partial C^2}(dC)^2 + \frac{1}{2}\frac{\partial^2 H}{\partial x^2} dx^2 + \frac{\partial^2 H}{\partial C \partial x}(dC)(dx)] + \beta \frac{\mu_p(C(t))dt - g(H(C, x))}{g'(H(C, x))} dt = 0 \tag{6.12}$$

From our Brownian motion rules we know that:

$$(dC)^2 = C^2 \sigma^2 dt \tag{6.13}$$

$$(dx)(dC) = a\varphi_e C \sigma^2 dt \tag{6.14}$$

where a represents the correlation between dw_1 and dw_2

$$dx^2 = \sigma^2 \varphi_e^2 dt \tag{6.15}$$

Using the results from our Brownian motion rules above, plugging in (6.9), (6.10), and (6.11), dividing by dt, and using the fact that E[dw]=0, we arrive at:

$$\frac{\sigma^2}{2}C^2\frac{\partial^2 H}{\partial C^2}+\frac{1}{2}\sigma^2\phi_e^2\frac{\partial^2 H}{\partial x^2}+$$
$$(x+\bar{x})C\frac{\partial H}{\partial C}+(\rho-1)x\frac{\partial H}{\partial x}+w\varphi_e\sigma^2 C\frac{\partial^2 H}{\partial C\partial x}+\frac{\beta}{\rho}(1-\gamma)\left[\frac{C^\rho}{((1-\gamma)H)^{\rho/(1-\gamma)}}-1\right]H=0\,. \tag{6.16}$$

This is the PDE, which ultimately describes the situation. To make our situation easier(i.e. this PDE is very difficult to solve), we assume the functional form(note: this is an educated guess because it makes sense to think of C as a sort of scaling factor) and hope that this reduces to a nonlinear ODE:

$$H(x,C)=\frac{C^{1-\gamma}}{1-\gamma}(h(x))^{\frac{1-\gamma}{\rho}} \tag{6.17}$$

The first-order partial derivatives of $H(x,C)$ are

$$\frac{\partial H}{\partial C}=(1-\gamma)\frac{H}{C}\quad\text{ and }\quad\frac{\partial H}{\partial x}=\frac{(1-\gamma)H}{\rho}\frac{h'}{h} \tag{6.18}$$

The second-order partial derivatives of $H(x,C)$ are

$$\frac{\partial^2 H}{\partial C^2}=(-\gamma)(1-\gamma)\frac{H}{C^2}\,,$$
$$\frac{\partial^2 H}{\partial x^2}=\frac{1-\gamma}{\rho}\left(H\left(\frac{hh''+\frac{1-\gamma-\rho}{\rho}(h')^2}{h^2}\right)\right)$$
$$\frac{\partial^2 H}{\partial C\partial x}=\frac{(1-\gamma)^2 H}{\rho C}\frac{h'}{h}\,. \tag{6.19}$$

Plugging these values in, we arrive at the differential equation:

$$\frac{\sigma^2}{2}(-\gamma)(1-\gamma)H+\frac{1}{2}\sigma^2\phi_e^2\frac{1-\gamma}{\rho}\left(H\left(\frac{hh''+\frac{1-\gamma-\rho}{\rho}(h')^2}{h^2}\right)\right)+$$
$$(x+\bar{x})(1-\gamma)H+(\rho-1)x\frac{(1-\gamma)H}{\rho}\frac{h'}{h}+w\varphi_e\sigma^2\frac{(1-\gamma)^2 H}{\rho}\frac{h'}{h}+\frac{\beta}{\rho}(1-\gamma)\left[\frac{1}{h}-1\right]H=0\,. \tag{6.20}$$

Now, dividing by $(1-\gamma)H$, and rearranging terms, we arrive at the following ordinary differential equation.

$$h(x)h''(x)=A(x)(h(x))^2+B(x)h(x)h'(x)+C(h'(x))^2+Dh(x) \tag{6.21}$$

where

$$A(x)=\frac{-2\rho}{\sigma^2\varphi_e^2}((x+\bar{x})-\frac{\beta}{\rho}-\frac{\sigma^2\gamma}{2})=c_{11}x+c_{12} \tag{6.22}$$

$$B(x)=\frac{-2\rho}{\sigma^2\varphi_e^2}(\frac{(\rho-1)x}{\rho}+\varphi_e\sigma^2\frac{(1-\gamma)w}{\rho})=c_{21}x+c_{22} \tag{6.23}$$

$$C=(\frac{-(1-\gamma-\rho)}{\rho})=c_3 \tag{6.24}$$

and

$$D = \frac{-2\rho}{\sigma^2\varphi_e^2}(\frac{\beta}{\rho}) = c_4 \tag{6.25}$$

We now apply a change of variable $\mu = ax$ (this is similar to the change of variable that Yu Chen did in Numerical Solutions of Portfolio Decision Problems, when he let $\mu = \epsilon * \sigma_{mu} * x$. This implies that $h(x) = h(\frac{\mu}{a})$, which implies that $h'(\mu) = \frac{1}{a}h'(x)$ and $h''(\mu) = \frac{1}{a^2}h''(x)$. From this we can see that:

$$h(\mu)h''(\mu) = A(\mu)(h(\mu))^2 + B(\mu)h(\mu)h'(\mu) + C(h'(\mu))^2 + Dh(\mu) \tag{6.26}$$

where

$$A(\mu) = \frac{-2\rho a^2}{\sigma^2\varphi_e^2}((\frac{\mu}{a} + \bar{x}) - \frac{\beta}{\rho} - \frac{\sigma^2\gamma}{2}) = c_{11}x + c_{12} \tag{6.27}$$

$$B(\mu) = \frac{-2\rho a}{\sigma^2\varphi_e^2}(\frac{(\rho-1)\frac{\mu}{a}}{\rho} + \varphi_e\sigma^2\frac{w(1-\gamma)}{\rho}) = c_{21}x + c_{22} \tag{6.28}$$

$$C = (\frac{-(1-\gamma-\rho)}{\rho}) = c_3 \tag{6.29}$$

and

$$D = \frac{-2\rho a^2}{\sigma^2\varphi_e^2}(\frac{\beta}{\rho}) = c_4. \tag{6.30}$$

We now derive initial conditions. We first note the relationship between $h(0)$ and $H(0)$, where we let $L(0) = (1-\gamma)H(0)$

$$h(0) = \left(\frac{L(0)}{C(0)^{1-\gamma}}\right)^{\frac{\rho}{1-\gamma}} \tag{6.31}$$

This is well defined, because(as we will see later), the utility function $H(x, C)$, is defined as negative. We also note that $H(0)$ is simply $J(0)$ from Marianne's thesis. $J(t)$, according to Marianne's thesis, is simply (we switch their H with a G to avoid confusing notation):

$$J(W(t), \mu(t)) = e^{-\nu t}G(\mu(t))^{-\frac{1-\gamma}{1-\psi}}\frac{W(t)^{1-\gamma}}{1-\gamma} \tag{6.32}$$

This implies that $J(0)$ is:

$$J(0) = G(0)^{-\frac{1-\gamma}{1-\psi}}\frac{W(0)^{1-\gamma}}{1-\gamma} \tag{6.33}$$

Equating $J(0)$ with our $H(0)$, we see that:

$$G(0)^{-\frac{1-\gamma}{1-\psi}}\frac{W(0)^{1-\gamma}}{1-\gamma} = \frac{C^{1-\gamma}}{1-\gamma}(h(x))^{\frac{1-\gamma}{\rho}} \tag{6.34}$$

Therefore, we still must find G(0). According to page 22 of Marianne's thesis, we see that:

$$G_0 = G(\mu(0)) = \beta^{\psi}\frac{W(0)}{C(0)}. \tag{6.35}$$

From (6.34), this immediately implies that:

$$(\beta^{\psi}\frac{W(0)}{C(0)})^{-\frac{1-\gamma}{1-\psi}}\frac{W(0)^{1-\gamma}}{1-\gamma} = \frac{C^{1-\gamma}}{1-\gamma}(h(x))^{\frac{1-\gamma}{\rho}} \tag{6.36}$$

We want to solve for $h(0)$. Therefore, moving terms to the other side:

$$h(0) = ((\beta^{\psi}\frac{W(0)}{C(0)})^{-\frac{1-\gamma}{1-\psi}}\frac{W(0)}{C(0)}^{1-\gamma})^{\frac{\rho}{1-\gamma}} \tag{6.37}$$

Using the properties of exponents, we see that:

$$h(0) = (\beta\frac{W(0)}{C(0)})^{-\frac{\psi\rho}{1-\psi}} \tag{6.38}$$

This is clearly well defined.

We now seek the second initial condition. From (7.2), we remember that:

$$\frac{\partial H}{\partial x} = \frac{(1-\gamma)H}{\rho}\frac{h'}{h} \tag{6.39}$$

This implies that:

$$\frac{\partial h}{\partial x} = h_1 = \frac{\rho}{(1-\gamma)H(0)}h(0)\frac{\partial H}{\partial x}(0) \tag{6.40}$$

To derive the second initial condition, we then look at the partial derivative of J with respect to μ, which is equivalent to the partial derivative of H with respect to x.

$$\frac{\partial J}{\partial \mu} = -e^{-\nu t}\frac{1-\gamma}{1-\psi}\frac{W(t)^{1-\gamma}}{1-\gamma}G(\mu(t))^{-\frac{1-\gamma}{1-\psi}-1}G' \tag{6.41}$$

$$= -e^{-\nu t}\frac{1-\gamma}{1-\psi}\frac{W(t)^{1-\gamma}}{1-\gamma}G(\mu(t))^{-\frac{1-\gamma}{1-\psi}}\frac{G'}{G} = -\frac{1-\gamma}{1-\psi}J\frac{G'}{G} \tag{6.42}$$

We still need to find G'. From Marianne's Thesis (p.23), we know that:

$$G'(0) = \frac{(1-\psi)}{\rho\sigma_{\mu}(1-\gamma)}\left(\frac{1}{\sigma_S} - \frac{\alpha(0)\gamma\sigma_S}{\mu(0)}\right)G((0)). \tag{6.43}$$

Plugging this all in, we see that:

$$h_1 = \frac{\rho}{(1-\gamma)H(0)}h(0)(-\frac{1-\gamma}{1-\psi}H(0)\frac{\frac{(1-\psi)}{\rho\sigma_{\mu}(1-\gamma)}\left(\frac{1}{\sigma_S} - \frac{\alpha(0)\gamma\sigma_S}{\mu(0)}\right)G((0))}{G(0)}) \tag{6.44}$$

Canceling terms, we see that:

$$h_1 = \frac{1}{(1-\gamma)}h(0)(\frac{-1}{\sigma_{\mu}}\left(\frac{1}{\sigma_S} - \frac{\alpha(0)\gamma\sigma_S}{\mu(0)}\right)) \tag{6.45}$$

All the coefficients(there is no forcing term) are affine functions, and therefore analytic with infinite radius of convergence. Applying the Cauchy-Kovalevsky Theorem, this means that $h(x)$ is also analytic with some radius of convergence (to determine this radius of convergence we use Yu Chen's program). We write $h(x)$ as such:

$$h(x) = \sum_{j=0}^{\infty} a_j x^j \tag{6.46}$$

We will now find the power series expansions of the $h(x)h''(x)$, $(h(x))^2$, $h(x)h'(x)$, and $(h'(x))^2$. We will then plug in these power series into our differential equation. We will then equate the coefficients associated with x^j. This will give us a recurrence relationship to determine each a_j in our power series expansion for $h(x)$

We know that:

$$h'(x) = \sum_{j=0}^{\infty}(j+1)a_{j+1}x^j \tag{6.47}$$

Using (6.46) and (6.47) we see that:

$$(h(x))^2 = \sum_{j=0}^{\infty}(\sum_{k=0}^{j} a_k a_{j-k})x^j \tag{6.48}$$

which implies that:

$$(c_{11}x + c_{12})(h(x))^2 = \sum_{j=0}^{\infty}(c_{11}\sum_{k=0}^{j-1} a_k a_{j-k-1} + c_{12}\sum_{k=0}^{j} a_k a_{j-k})x^j \tag{6.49}$$

Additionally,

$$(h'(x))^2 = \sum_{j=0}^{\infty}(\sum_{k=0}^{j}(k+1)(j-k+1)a_{k+1}a_{j-k+1})x^j \tag{6.50}$$

$$h(x)h''(x) = \sum_{j=0}^{\infty}(\sum_{k=0}^{j}(k+1)(k+2)a_{k+2}a_{j-k})x^j \tag{6.51}$$

$$h(x)h'(x) = \sum_{j=0}^{\infty}(\sum_{k=0}^{j}(k+1)a_{k+1}a_{j-k})x^j \tag{6.52}$$

which implies that:

$$(c_{21}x + c_{22})h(x)h'(x) = \sum_{j=0}^{\infty}(c_{21}\sum_{k=0}^{j-1}(k+1)a_{k+1}a_{j-k-1} + c_{22}\sum_{k=0}^{j}(k+1)a_{k+1}a_{j-k})x^j \tag{6.53}$$

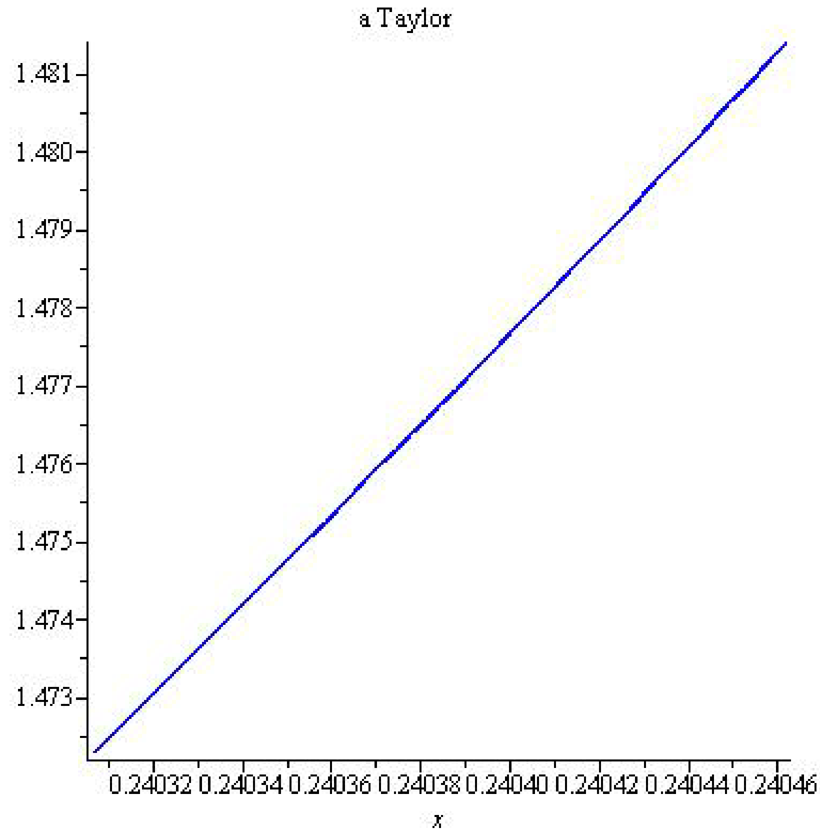
Plugging (6.53), (6.51) ,(6.50), (6.49), and (6.46) into our ODE, and then equating the coefficients for x^j, we see that:

$$\sum_{k=0}^{j}(k+1)(k+2)a_{k+2}a_{j-k} = (c_{11}\sum_{k=0}^{j-1} a_k a_{j-k-1} + c_{12}\sum_{k=0}^{j} a_k a_{j-k}) + (c_{21}\sum_{k=0}^{j-1}(k+1)a_{k+1}a_{j-k-1}+$$
$$c_{22}\sum_{k=0}^{j}(k+1)a_{k+1}a_{j-k}) + c_3(\sum_{k=0}^{j}(k+1)(j-k+1)a_{k+1}a_{j-k+1}) + c_4 a_j \tag{6.54}$$

This immediately implies our recurrence relation to determine each a_j:

$$(j+1)(j+2)a_{j+2}a_0 = (c_{11}\sum_{k=0}^{j-1} a_k a_{j-k-1} + c_{12}\sum_{k=0}^{j} a_k a_{j-k}) + (c_{21}\sum_{k=0}^{j-1}(k+1)a_{k+1}a_{j-k-1}$$
$$+ c_{22}\sum_{k=0}^{j}(k+1)a_{k+1}a_{k-j}) + c_3(\sum_{k=0}^{j}(k+1)(j-k+1)a_{k+1}a_{j-k+1}) + c_4 a_j - \sum_{k=0}^{j-1}(k+1)(k+2)a_{k+2}a_{j-k} \tag{6.55}$$

We have solved this. This function, which should have the same properties as our utility, exhibits a clear positive relation with consumption. This is certainly what we would expect:



We now try to develop the differential equation for the price-dividend ratio:

7. The Stochastic Process for the Pricing Kernel

We will derive an ordinary differential equation for the equilibrium price-dividend ratio in the DEKP model, given the power series of the lifetime utility function $H(C)$ around $C = \bar{C}$. The pricing kernel for the investor is given by

$$\Lambda(C, U, t) = e^{-\delta t} \frac{\partial \bar{f}}{\partial C}(C, U) = \frac{\partial}{\partial C}\left[\frac{\beta}{\rho} \cdot \frac{C^\rho - ((1-\gamma)U)^{\rho/(1-\gamma)}}{((1-\gamma)U)^{\rho/(1-\gamma)-1}}\right] = \frac{\beta C^{\rho-1}}{((1-\gamma)U)^{\rho/(1-\gamma)-1}}. \tag{7.1}$$

The first-order partial derivatives of $\Lambda(C, U, t)$ are

$$\frac{\partial \Lambda}{\partial C} = (\rho - 1)\frac{\Lambda}{C} \quad \text{and} \quad \frac{\partial \Lambda}{\partial U} = \frac{1-\gamma-\rho}{1-\gamma}\frac{\Lambda}{U}. \quad \text{and} \quad \frac{\partial \Lambda}{\partial t} = (-\delta)\Lambda \tag{7.2}$$

The second-order partial derivatives of $\Lambda(C, U)$ are

$$\begin{aligned}
\frac{\partial^2\Lambda}{\partial C^2} &= (\rho-1)(\rho-2)\frac{\Lambda}{C^2}\,,\\
\frac{\partial^2\Lambda}{\partial U^2} &= -\frac{\rho(1-\gamma-\rho)}{(1-\gamma)^2}\frac{\Lambda}{U^2}\,,\\
\frac{\partial^2\Lambda}{\partial C\partial U} &= \frac{(\rho-1)(1-\gamma-\rho)}{1-\gamma}\frac{\Lambda}{CU}\,.
\end{aligned} \tag{7.3}$$

After algebraic massaging, plugging in for our stochastic processes, and using our Brownian motion rules, we arrive at the following stochastic process for Λ:

$$\frac{d\Lambda}{\Lambda} = B_1 dt + B_2 d\omega_1 + B_3 d\omega_2 \tag{7.4}$$

where

$$\begin{aligned}
B_1 = &(-\delta + ((\rho-1)\frac{1}{C} + \frac{1-\gamma-\rho}{1-\gamma}\frac{1}{H}\frac{\partial H}{\partial C})(x+\bar{x})C + \frac{1-\gamma-\rho}{1-\gamma}\frac{1}{H}\frac{\partial H}{\partial x}(\rho-1)x\\
&+ \frac{1-\gamma-\rho}{1-\gamma}\frac{1}{H}\frac{1}{2}\frac{\partial^2 H}{\partial C^2} + \frac{1}{2}(\rho-1)(\rho-2)\frac{1}{C^2} + \frac{-1}{2}\frac{\rho(1-\gamma-\rho)}{(1-\gamma)^2}\frac{1}{H^2}(\frac{\partial H}{\partial C})^2\\
&+ \frac{(\rho-1)(1-\gamma-\rho)}{1-\gamma}\frac{1}{CH}\frac{\partial H}{\partial C}(C^2\sigma^2) + (\frac{1-\gamma-\rho}{1-\gamma}\frac{1}{H}\frac{\partial^2 H}{\partial x\partial C} +\\
&\frac{-1}{2}\frac{\rho(1-\gamma-\rho)}{(1-\gamma)^2}\frac{1}{H^2}2\frac{\partial H}{\partial C}\frac{\partial H}{\partial x} + \frac{1}{CH}\frac{\partial H}{\partial x})\varphi_e\sigma^2 Ca\\
&+ (\frac{1-\gamma-\rho}{1-\gamma}\frac{1}{H}\frac{1}{2}\frac{\partial^2 H}{\partial x^2} + \frac{-1}{2}\frac{\rho(1-\gamma-\rho)}{(1-\gamma)^2}\frac{1}{H^2}(\frac{\partial H}{\partial x})^2)\varphi_e^2\sigma^2)
\end{aligned} \tag{7.5}$$

and

$$B_2 = ((\rho-1)\frac{1}{C} + \frac{1-\gamma-\rho}{1-\gamma}\frac{1}{H}\frac{\partial H}{\partial C})C\sigma \text{ and } B_3 = (\frac{1-\gamma-\rho}{1-\gamma}\frac{1}{H}\frac{\partial H}{\partial x})\varphi_e\sigma \tag{7.6}$$

8. Equilibrium Price-Dividend Function in the DEKP Model

We begin with the Euler condition:

$$\Lambda(t)D(t)dt + E_t\left[d(\Lambda(t)P(t))\right] = 0 \tag{8.1}$$

where $P(t)$ is the price of a stock at time t and $D(t)$ is the dividend paid by this stock at time t.

Following algebraic manipulation, we arrive at an equivalent expression:

$$\frac{1}{p}\,dt + E_t\left[\frac{d\Lambda}{\Lambda} + \frac{dp}{p} + \frac{dD}{D} + \frac{d\Lambda dp}{\Lambda p} + \frac{dDdp}{Dp} + \frac{d\Lambda dD}{\Lambda D}\right] = 0\,. \tag{8.2}$$

where p is the price dividend ratio.

We now assume that p is a function of x and $c = ln(C)$ the two underlying state processes. From this we can calculate dp using Ito's Lemma. After employing Ito's Lemma, a little algebraic massaging, and plugging in for the stochastic processes we arrive at:

$$\frac{dp}{p} = D_1\frac{dt}{p} + D_2 d\omega_1 + D_3 d\omega_2 \tag{8.3}$$

where

$$D_1 =(\frac{\partial p}{\partial c}(x+\bar{x})+\frac{\partial p}{\partial x}(\rho-1)x+\frac{1}{2}\frac{\partial^2 p}{\partial x^2}\sigma^2\phi_e^2 \\ +\frac{1}{2}\frac{\partial^2 p}{\partial c^2}\sigma^2+\frac{\partial^2 p}{\partial c\partial x}a\sigma^2\phi_e) \tag{8.4}$$

$$D_2=\frac{\sigma}{p} \text{ and } D_3=\frac{\varphi_e\sigma}{p} \tag{8.5}$$

We have our equation for the dividend process (given):

$$\frac{dD}{D}=(\phi x+\bar{x})dt+\varphi_d\sigma d\omega_3 \tag{8.6}$$

We now have expressions for everything in the Euler equation. Plugging in, and after some algebraic massaging, we arrive at:

$$A_1+A_2p+A_3\frac{\partial p}{\partial c}+A_4\frac{\partial p}{\partial x}+A_5\frac{\partial^2 p}{\partial x^2}+A_6\frac{\partial^2 p}{\partial c^2}+A_7\frac{\partial^2 p}{\partial c\partial x}=0 \tag{8.7}$$

where

$$A_1 =1+(\frac{1-\gamma-\rho}{1-\gamma}\frac{1}{H}\frac{\partial H}{\partial x})\varphi_e^2\sigma^2+((\rho-1)\frac{1}{C}+\frac{1-\gamma-\rho}{1-\gamma}\frac{1}{H}\frac{\partial H}{\partial C})C\varphi_e\sigma^2 a \\ +\varphi_d\sigma^2 b+\varphi_e\varphi_d\sigma^2 z+(((\rho-1)\frac{1}{C}+\frac{1-\gamma-\rho}{1-\gamma}\frac{1}{H}\frac{\partial H}{\partial C})C\sigma^2) \\ +(\frac{1-\gamma-\rho}{1-\gamma}\frac{1}{H}\frac{\partial H}{\partial x})\varphi_e\sigma^2 a \tag{8.8}$$

$$A_2 =((\frac{1-\gamma-\rho}{1-\gamma}\frac{1}{H}\frac{\partial H}{\partial C})C\varphi_d\sigma^2 b+ \\ (\frac{1-\gamma-\rho}{1-\gamma}\frac{1}{H}\frac{\partial H}{\partial x})\varphi_e\varphi_d\sigma^2 z+\phi x+\bar{x} \\ (-\delta+((\rho-1)\frac{1}{C}+\frac{1-\gamma-\rho}{1-\gamma}\frac{1}{H}\frac{\partial H}{\partial C})(x+\bar{x})C+\frac{1-\gamma-\rho}{1-\gamma}\frac{1}{H}\frac{\partial H}{\partial x}(\rho-1)x \\ +\frac{1-\gamma-\rho}{1-\gamma}\frac{1}{H}\frac{1}{2}\frac{\partial^2 H}{\partial C^2}+\frac{1}{2}(\rho-1)(\rho-2)\frac{1}{C^2}+\frac{-1}{2}\frac{\rho(1-\gamma-\rho)}{(1-\gamma)^2}\frac{1}{H^2}(\frac{\partial H}{\partial C})^2 \\ +\frac{(\rho-1)(1-\gamma-\rho)}{1-\gamma}\frac{1}{CH}\frac{\partial H}{\partial C}(C^2\sigma^2)+(\frac{1-\gamma-\rho}{1-\gamma}\frac{1}{H}\frac{\partial^2 H}{\partial x\partial C}+ \\ \frac{-1}{2}\frac{\rho(1-\gamma-\rho)}{(1-\gamma)^2}\frac{1}{H^2}2\frac{\partial H}{\partial C}\frac{\partial H}{\partial x}+\frac{1}{CH}\frac{\partial H}{\partial x})\varphi_e\sigma^2 Ca \\ +(\frac{1-\gamma-\rho}{1-\gamma}\frac{1}{H}\frac{1}{2}\frac{\partial^2 H}{\partial x^2}+\frac{-1}{2}\frac{\rho(1-\gamma-\rho)}{(1-\gamma)^2}\frac{1}{H^2}(\frac{\partial H}{\partial x})^2)\varphi_e^2\sigma^2)) \tag{8.9}$$

$$A_3=(x+\bar{x})\ A_4=(\rho-1)x\ A_5=\frac{1}{2}\sigma^2\phi_e^2\ A_6=\frac{\sigma^2}{2}\ A_7=a\sigma^2\phi_e \tag{8.10}$$

p_0, the first initial condition, is equal to the power series defined by 4. We start with the formula for the return on equity.

$$R^e(x,c) = \frac{dp}{p} + \frac{dD}{D} + \frac{dp}{p}\frac{dD}{D} + \frac{1}{p}dt. \tag{8.11}$$

From my thesis, we recall that

$$\frac{dp}{p} = D_1\frac{dt}{p} + D_2 d\omega_1 + D_3 d\omega_2. \tag{8.12}$$

where

$$\begin{aligned} D_1 =&(\frac{\partial p}{\partial c}(x+\bar{x}) + \frac{\partial p}{\partial x}(\rho-1)x + \frac{1}{2}\frac{\partial^2 p}{\partial x^2}\sigma^2\phi_e^2 \\ &+ \frac{1}{2}\frac{\partial^2 p}{\partial c^2}\sigma^2 + \frac{\partial^2 p}{\partial c\partial x}a\sigma^2\phi_e) \end{aligned} \tag{8.13}$$

$$D_2 = \frac{\frac{\partial p}{\partial c}\sigma}{p} \text{ and } D_3 = \frac{\frac{\partial p}{\partial x}\varphi_e\sigma}{p}. \tag{8.14}$$

and that

$$\frac{dD}{D} = (\phi x + \bar{x})dt + \varphi_d\sigma d\omega_3. \tag{8.15}$$

Using our Brownian motion rules, this implies that (where a_{ij} represents the correlation between $d\omega_i$ and $d\omega_j$)

$$\frac{dp}{p}\frac{dD}{D} = (\frac{\rho_{13}\frac{\partial p}{\partial c}\varphi_d\sigma^2}{p} + \frac{\rho_{23}\frac{\partial p}{\partial x}\varphi_e\varphi_d\sigma^2}{p})dt. \tag{8.16}$$

Plugging (8.12), (8.15) and (8.16) into (8.11) we see that

$$R^e(x,c) = (\frac{\rho_{13}\varphi_d\sigma^2}{p} + \frac{\rho_{23}\varphi_e\varphi_d\sigma^2}{p} + \phi x + \bar{x} + \frac{1}{p} + \frac{D_1}{p})dt + \frac{\sigma\frac{\partial p}{\partial c}}{p}d\omega_1 + \frac{\varphi_e\sigma\frac{\partial p}{\partial x}}{p}d\omega_2 + \varphi_d\sigma d\omega_3. \tag{8.17}$$

This immediately implies that

$$E[R^e(x,c)] = (\frac{\rho_{13}\varphi_d\sigma^2}{p} + \frac{\rho_{23}\varphi_e\varphi_d\sigma^2}{p} + \phi x + \bar{x} + \frac{1}{p} + \frac{D_1}{p})dt. \tag{8.18}$$

We also know that the standard deviation of $\frac{\sigma}{p}\frac{\partial p}{\partial c}d\omega_1 + \frac{\varphi_e\sigma}{p}\frac{\partial p}{\partial x}d\omega_2 + \varphi_d\sigma d\omega_3$ is

$$\Sigma(x,c) = \sqrt{\frac{\sigma^2(\frac{\partial p}{\partial c})^2}{p^2} + 2\rho_{12}\frac{\frac{\partial p}{\partial c}\sigma}{p}\frac{\frac{\partial p}{\partial x}\varphi_e\sigma}{p} + 2\rho_{13}\frac{\frac{\partial p}{\partial c}\sigma^2\varphi_d}{p} + 2\rho_{23}\varphi_d\sigma\frac{\frac{\partial p}{\partial x}\varphi_e\sigma}{p} + \frac{\varphi_e^2\sigma^2\frac{\partial p}{\partial x}^2}{p^2} + \varphi_d^2\sigma^2}. \tag{8.19}$$

Plugging in (5.31), which we will label Σ_1, and p_1 for the price dividend ratio we found in the one-dimensional case, we see that

$$\Sigma(x,c) = \sqrt{\Sigma_1^2 + 2\rho_{12}\frac{\frac{\partial p_1}{\partial c}\sigma}{p_1}\frac{\frac{\partial p}{\partial x}\varphi_e\sigma}{p_1} + 2\rho_{23}\varphi_d\sigma\frac{\frac{\partial p}{\partial x}\varphi_e\sigma}{p_1} + \frac{\varphi_e^2\sigma^2\frac{\partial p}{\partial x}^2}{p_1^2}}. \tag{8.20}$$

This implies that

$$\frac{\varphi_e^2\sigma^2}{p_1^2}\frac{\partial p}{\partial x}^2 + (2\rho_{12}\frac{\frac{\partial p_1}{\partial c}\sigma}{p_1}\frac{\varphi_e\sigma}{p_1} + 2\rho_{23}\varphi_d\sigma\frac{\varphi_e\sigma}{p_1})\frac{\partial p}{\partial x} + \Sigma_1^2 - \Sigma^2(x,c) = 0. \tag{8.21}$$

Applying the quadratic formula, we know that

$$\frac{\partial p}{\partial x} = \frac{-(2\rho_{12}\frac{\frac{\partial p_1}{\partial c}\sigma}{p_1}\frac{\varphi_e\sigma}{p_1} + 2\rho_{23}\varphi_d\sigma\frac{\varphi_e\sigma}{p_1}) + \sqrt{(2\rho_{12}\frac{\frac{\partial p_1}{\partial c}\sigma}{p_1}\frac{\varphi_e\sigma}{p_1} + 2\rho_{23}\varphi_d\sigma\frac{\varphi_e\sigma}{p_1})^2 - 4(\frac{\varphi_e^2\sigma^2}{p_1^2})(\Sigma_1^2 - \Sigma^2(x,c))}}{2\frac{\varphi_e^2\sigma^2}{p_1^2}} \tag{8.22}$$

9. The Main Result

We have arrived at the final differential equation. The solution to this differential equation will describe the motion of the price-dividend ratio in the two-dimensional case. Before we state the main theorem it is necessary to give a little background information. Our first initial condition is our solution to the one-dimensional model. We derived our second initial condition in a fashion similar to the way that Cosimano and Chen derive their condition for the Wachter model. As always, p represents the price-dividend ratio, c, x reflect consumption and the variable that affects expected growth of consumption respectively, $\bar{c}, \bar{x}$ are constants that reflect the average consumption and the average drift term, or expected growth, of consumption respectively and a represents the coefficients of the power series of the price-dividend ratio.

Theorem 1. *The initial value problem*

$$B_1\frac{\partial^2 p}{\partial x^2} = B_2 + B_3 p + B_4\frac{\partial p}{\partial c} + B_5\frac{\partial p}{\partial x} + B_6\frac{\partial^2 p}{\partial c^2} + B_7\frac{\partial^2 p}{\partial c\partial x} \tag{9.23}$$

$$p(\bar{x}, c) = p_0(c) \tag{9.24}$$

$$\frac{\partial p}{\partial x}(\bar{x}, c) = p_1(c) \tag{9.25}$$

where the coefficients B_j are given by

$$B_1 = \frac{1}{2}\sigma^2\phi_e^2 \tag{9.26}$$

$$\begin{aligned} B_2 = &-1 - (\frac{1-\gamma-\rho}{1-\gamma}\frac{1}{H}\frac{\partial H}{\partial x})\varphi_e^2\sigma^2 + ((\rho-1)\frac{1}{C} + \frac{1-\gamma-\rho}{1-\gamma}\frac{1}{H}\frac{\partial H}{\partial C})C\varphi_e\sigma^2 a \\ &+ \varphi_d\sigma^2 b + \varphi_e\varphi_d\sigma^2 z + (((\rho-1)\frac{1}{C} + \frac{1-\gamma-\rho}{1-\gamma}\frac{1}{H}\frac{\partial H}{\partial C})C\sigma^2) \\ &+ (\frac{1-\gamma-\rho}{1-\gamma}\frac{1}{H}\frac{\partial H}{\partial x})\varphi_e\sigma^2 a \end{aligned} \tag{9.27}$$

$$
\begin{aligned}
B_3 = &-\Big(\big(\frac{1-\gamma-\rho}{1-\gamma}\frac{1}{H}\frac{\partial H}{\partial C}\big)C\varphi_d\sigma^2 b+\\
&\big(\frac{1-\gamma-\rho}{1-\gamma}\frac{1}{H}\frac{\partial H}{\partial x}\big)\varphi_e\varphi_d\sigma^2 z+\phi x+\bar{x}\\
&\Big(-\delta+\big((\rho-1)\frac{1}{C}+\frac{1-\gamma-\rho}{1-\gamma}\frac{1}{H}\frac{\partial H}{\partial C}\big)(x+\bar{x})C+\frac{1-\gamma-\rho}{1-\gamma}\frac{1}{H}\frac{\partial H}{\partial x}(\rho-1)x\\
&+\frac{1-\gamma-\rho}{1-\gamma}\frac{1}{H}\frac{1}{2}\frac{\partial^2 H}{\partial C^2}+\frac{1}{2}(\rho-1)(\rho-2)\frac{1}{C^2}+\frac{-1}{2}\frac{\rho(1-\gamma-\rho)}{(1-\gamma)^2}\frac{1}{H^2}\big(\frac{\partial H}{\partial C}\big)^2\\
&+\frac{(\rho-1)(1-\gamma-\rho)}{1-\gamma}\frac{1}{CH}\frac{\partial H}{\partial C}(C^2\sigma^2)+\big(\frac{1-\gamma-\rho}{1-\gamma}\frac{1}{H}\frac{\partial^2 H}{\partial x\partial C}+\\
&\frac{-1}{2}\frac{\rho(1-\gamma-\rho)}{(1-\gamma)^2}\frac{1}{H^2}2\frac{\partial H}{\partial C}\frac{\partial H}{\partial x}+\frac{1}{CH}\frac{\partial H}{\partial x}\big)\varphi_e\sigma^2 Ca\\
&+\big(\frac{1-\gamma-\rho}{1-\gamma}\frac{1}{H}\frac{1}{2}\frac{\partial^2 H}{\partial x^2}+\frac{-1}{2}\frac{\rho(1-\gamma-\rho)}{(1-\gamma)^2}\frac{1}{H^2}\big(\frac{\partial H}{\partial x}\big)^2\big)\varphi_e^2\sigma^2\Big)\Big)
\end{aligned}
\tag{9.28}
$$

$$
B_4=-(x+\bar{x}),\ B_5=(1-\rho)x,\ B_6=\frac{-\sigma^2}{2},\ B_7=-a\sigma^2\phi_e \tag{9.29}
$$

and the initial conditions are given by,

$$
p_0=\sum p_k(c-\bar{c})^k \tag{9.30}
$$

and

$$
\frac{\partial p}{\partial x}=\frac{-(2\rho_{12}\frac{\frac{\partial p_1}{\partial c}\sigma}{p_1}\frac{\varphi_e\sigma}{p_1}+2\rho_{23}\varphi_d\sigma\frac{\varphi_e\sigma}{p_1})+\sqrt{(2\rho_{12}\frac{\frac{\partial p_1}{\partial c}\sigma}{p_1}\frac{\varphi_e\sigma}{p_1}+2\rho_{23}\varphi_d\sigma\frac{\varphi_e\sigma}{p_1})^2-4(\frac{\varphi_e^2\sigma^2}{p_1^2})(\Sigma_1^2-\Sigma^2(x,c))}}{2\frac{\varphi_e^2\sigma^2}{p_1^2}} \tag{9.31}
$$

has an analytic solution p, whose power series at $(\bar{x},\bar{c})$

$$
p(x,c)=\sum_{j,k=0}^{\infty}a_{j,k}(x-\bar{x})^j(c-\bar{c})^k \tag{9.32}
$$

has a radius of convergence.

Proof. That the solution is analytic in a domain containing region G follows from the Cauchy-Kovalevsky Theorem that we state in the appendix . We shall explain why the Cauchy-Kovalevsky theorem applies in this situation. First, we know that our initial conditions are analytic because we know that our power series p_0 is analytic. This comes from the fact that p_0 comes out of a second-order linear differential equation, whose coefficients are analytic. We have done error analysis on the coefficients in that differential equation and have proved that our g function, which appears in the coefficients of ps differential equation converges with a radius of convergence at least as big as 1.2 times the standard deviation. (By g we mean the power series solution $g(c)=\sum_{n=0}^{\infty}g_n(c-\bar{c})^n$ of the initial value problem (3.13)). We solved for the coefficients of the power series to the solution h by the recurrence relation (6.55) and it appears that there is a radius of convergence of at least .8 times the standard deviation. However, we have yet to do the error analysis. But, based on analytic coefficients and analytic initial conditions, it appears that the price-dividend function is analytic and has a radius of convergence. This follows from 3. The next step is to write down the recurrence relationship and produce explicit bounds for the coefficients. Therefore, our main result, is that we have created a well-defined problem to solve for the price-dividend ratio under the assumptions of a multi-dimensional consumption process, Duffie-Epstien preferences and Kreps-Porteus Utility.

10. Part II: The Black-Scholes Model

In the upcoming section we will describe the Black Scholes Model. This model was a revolutionary one; it gave a closed form solution for the pricing of call options. It is incredibly beautiful and any discussion of major contributions to mathematical finance certainly begins with this model.

The owner of a call option has the right to purchase a stock $s(t)$, for a strike price K, at a future time T. We will denote this call option $V = V(s,t)$ since we assume that it is a function of the stock price s and time t. We also assume that the stock price follows the following stochastic differential equation:

$$ds = asdt + \sigma sdw. \tag{10.1}$$

where dw is the increment of a Brownian motion (it has a mean of zero and a variance of dt), a and σ are positive constants and dt is an increment of time. We also assume that there is a constant riskless rate of return r, that there are no transaction costs, the market flows continuously and there are no dividends.

We let $B(t)$ represent the number of currency invested in the riskless asset, $O(t)$ represent the number of options the investor holds and $N(t)$ represent the number of stocks held. The value of the investors portfolio, at time t, is therefore:

$$\pi(t) = N(t)S(t) + O(t)V(t) + B(t). \tag{10.2}$$

which implies that:

$$d\pi(t) = N(t)dS(t) + O(t)dV(t) + dB(t). \tag{10.3}$$

From Ito's Lemma, we know that

$$dV = V_s ds + V_t dt + \frac{1}{2}V_{ss}(ds)^2. \tag{10.4}$$

From (10.1), and our Brownian Motion multiplication rules, we know that

$$(ds)^2 = \sigma^2 s^2 dt. \tag{10.5}$$

Plugging (10.5) and (10.1) into (10.4), we see that

$$dV = [V_t + asV_s + \frac{\sigma^2 s^2}{2}V_{ss}]dt + [\sigma s V_s]dw \tag{10.6}$$

By definition of a risk free asset we know that

$$dB = rBdt \tag{10.7}$$

Plugging in (10.7), (10.6), and (10.1) into (10.3) we see that

$$d\pi = [Nas + O(V_t + asV_s + \frac{\sigma^2 s^2}{2}V_{ss}) + rB]dt + [O\sigma s V_s + N\sigma s]dw \tag{10.8}$$

If we choose stocks and options such that

$$N = -OV_s, \tag{10.9}$$

then the portfolio is riskless and will earn the riskless rate of return r. That is

$$d\pi = \pi r dt. \tag{10.10}$$

Plugging (10.9) into (10.2), we see that

$$\pi = O[V - sV_s] + B. \tag{10.11}$$

Furthermore, plugging (10.9) into (10.8), we see that

$$d\pi = [O(V_t + \frac{\sigma^2 s^2}{2} V_{ss}) + rB]dt. \tag{10.12}$$

Finally, plugging (10.11) and (10.12) into (10.10), we see that

$$[O(V_t + \frac{\sigma^2 s^2}{2} V_{ss}) + rB]dt = (O[V - sV_s] + B)rdt \tag{10.13}$$

Canceling terms, and rearranging the equation, we arrive at the Black-Scholes Partial differential equation

$$\frac{\partial V}{\partial t} + rs\frac{\partial V}{\partial s} + \frac{\sigma^2 s^2}{2}\frac{\partial^2 V}{(\partial s)^2} - rV = 0. \tag{10.14}$$

where V is the price of a call option, s is the stock price, and r is the riskless rate of return(e.g. the return on a government bond). This PDE is also subject to the following boundary conditions:

$$V(s,T) = \lim_{t\uparrow T} V(s,t) = [s-K]^+. \tag{10.15}$$

where K represents the strike price, and T is the exercise time

$$V(0,t) = \lim_{s\downarrow 0} V(s,t) = 0 \ \forall \ t\epsilon[0,T] \tag{10.16}$$

Additionally,

$$\lim_{s\uparrow\infty}(V(s,t) - (s - e^{-r(T-t)})) = 0. \tag{10.17}$$

We now apply the change of variable:

$$V(s,t) = e^{-r\tau}G(x,y) \tag{10.18}$$

where x and y are functions of s and t, and $\tau = T - t$ Taking partial derivatives(by the definition of a total derivative), we see that:

$$V_s = e^{-r\tau}[G_x x_s + G_y y_s]. \tag{10.19}$$

$$V_t = re^{-r\tau}G + e^{-r\tau}[G_x x_t + G_y y_t]. \tag{10.20}$$

$$V_{ss} = e^{-r\tau}[G_{xx}x_s^2 + G_{xy}x_s y_s + G_x x_{ss} + G_{yx}x_s y_s + G_{yy}y_s^2 + G_y y_{ss}]. \tag{10.21}$$

Substituting (10.18) through (10.21), we arrive at an equivalent partial differential equation(by combining like terms and dividing through by $e^{-r\tau}$:

$$\frac{\sigma^2 s^2}{2}x_s^2 G_{xx} + \frac{\sigma^2 s^2}{2}x_s y_s G_{xy} + \frac{\sigma^2 s^2}{2}y_s^2 G_{yy} + [\frac{\sigma^2 s^2}{2}x_{ss} + rsx_s + x_t]G_x + [\frac{\sigma^2 s^2}{2}y_{ss} + rsy_s + y_t]G_y = 0. \tag{10.22}$$

In order to put this in the form of the heat equation (i.e. $G_{xx} = G_y$), we need the following equalities to hold:

$$\frac{\sigma^2 s^2}{2}x_s^2 + \frac{\sigma^2 s^2}{2}y_{ss} + rsy_s + y_t = 0, \tag{10.23}$$

$$\sigma^2 s^2 x_s y_s = 0, \tag{10.24}$$

$$\frac{\sigma^2 s^2}{2}y_s^2 = 0, \tag{10.25}$$

$$\frac{\sigma^2 s^2}{2} x_{ss} + rsx_s + x_t = 0. \tag{10.26}$$

Looking at (10.25) we see that y can not depend on s, if we want to transform this PDE to the heat equation. Additionally, we must satisfy the following two conditions:

$$\frac{\sigma^2 s^2}{2} x_s^2 + y_t = 0, \tag{10.27}$$

$$\frac{\sigma^2 s^2}{2} x_{ss} + rsx_s + x_t = 0. \tag{10.28}$$

Looking at (10.23), and recalling that y is only a function of t, we see immediately that sx_s must not be a function of s. Recalling that x must also be a function of t, we use the very simple change of variable:

$$x(s,t) = ln(s) + \gamma\tau. \tag{10.29}$$

This implies that(in order to transform the PDE to the heat equation):

$$\frac{\sigma^2}{2} + y_t = 0, \tag{10.30}$$

$$\frac{-\sigma^2}{2} + r - \gamma = 0. \tag{10.31}$$

This implies that $\gamma = r - \frac{\sigma^2}{2}$ and $y(t) = \frac{\sigma^2 \tau}{2}$ Therefore, we arrive at the heat equation with the boundary condition:

$$G(x(s,T), y(T)) = [e^x - K]^+. \tag{10.32}$$

(This can be seen immediately after looking at the initial boundary condition, looking at our change of variables, and then plugging in for $t = T$) We now use a separation of variables technique: We assume that G can be written like:

$$G(x,y) = X(x)Y(y). \tag{10.33}$$

This implies that we have the following differential equation:

$$X''(x)Y(y) = X(x)Y'(y). \tag{10.34}$$

This immediately implies that we have the following pair of ordinary differential equations(where k is a constant):

$$X'' + kX = 0. \tag{10.35}$$

$$Y' + kY = 0. \tag{10.36}$$

Letting $a = \sqrt{k}$, we see that :

$$X(x) = C_1 e^{iax}. \tag{10.37}$$

$$Y(y) = C_2 e^{-a^2 y}. \tag{10.38}$$

and

$$G(x,y) = ce^{iax - a^2 y}. \tag{10.39}$$

And, since c can take on any value, we know that

$$G(x,y) = \int_{-\infty}^{\infty} c(a) e^{iax - a^2 y} da. \tag{10.40}$$

where c, C_1, and C_2 are arbitrary constants.

We can see that this is true by applying the differential operator(the one for the differential equation) to that integral, and then bringing that differential operator inside the integral.

If $y = 0$, (i.e. if $T = t$) we see that

$$f(x) = \int_{-\infty}^{\infty} c(a)e^{iax}da. \tag{10.41}$$

Applying the inverse Fourier transform, we see that:

$$c(a) = \frac{1}{2\pi}\int_{-\infty}^{\infty} f(x)e^{-iax}dx. \tag{10.42}$$

Plugging this into our equation for G, we see that:

$$G(x,y) = \frac{1}{2\pi}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} f(u)e^{-i(u-x)a-a^2y}duda. \tag{10.43}$$

By Fubini's Theorem, we can reverse the order of integration (by Fubini-Tonelli). Then, we see that we have a Fourier pair:

$$G(x,y) = \frac{1}{2\pi}\int_{-\infty}^{\infty} f(u)\int_{-\infty}^{\infty} e^{-i(u-x)a-a^2y}dadu. \tag{10.44}$$

which implies that:

$$G(x,y) = \frac{1}{2\pi}\int_{-\infty}^{\infty} f(u)\int_{-\infty}^{\infty} \widehat{e^{-a^2y}}(u-x)dadu. \tag{10.45}$$

On page six of Professor Himonas' notes, we are told that, in one dimension:

$$\widehat{e^{\frac{-\lambda}{2}|x|^2}}(\zeta) = (\frac{2\pi}{\lambda})^{\frac{1}{2}}e^{-\frac{|\zeta|^2}{2\lambda}}. \tag{10.46}$$

From that, we see that (10.44) is equivalent to:

$$G(x,y) = \frac{1}{2\sqrt{\pi y}}\int_{-\infty}^{\infty} f(u)e^{\frac{-(u-x)^2}{4y}}du. \tag{10.47}$$

Plugging in our boundary condition, which only allows positive values for the option, we see that:

$$G(x,y) = \frac{1}{2\sqrt{\pi y}}\int_{ln(K)}^{\infty} (e^u - K)e^{\frac{-(u-x)^2}{4y}}du. \tag{10.48}$$

Switching back to V, and plugging in for x and y, we see that:

$$V(s,t) = \frac{e^{-r\tau}}{\sigma\sqrt{2\pi\tau}}\int_{ln(K)}^{\infty} (e^u - K)e^{\frac{-(u-ln(s)-(r-\frac{\sigma^2}{2})\tau)^2}{2\sigma^2\tau}}du. \tag{10.49}$$

By the linearity of integrals, we see that:

$$V(s,t) = \frac{e^{-r\tau}}{\sigma\sqrt{2\pi\tau}}\int_{ln(K)}^{\infty} e^u e^{\frac{-(u-ln(s)-(r-\frac{\sigma^2}{2})\tau)^2}{2\sigma^2\tau}}du - \frac{e^{-r\tau}}{\sigma\sqrt{2\pi\tau}}\int_{ln(K)}^{\infty} (K)e^{\frac{-(u-ln(s)-(r-\frac{\sigma^2}{2})\tau)^2}{2\sigma^2\tau}}du. \tag{10.50}$$

We will now try to transform these integrals to be of the normal distribution form: We add and subtract $(ln(s) + r\tau)$, to the exponent under the first integral. This makes the first integral look like(One can see this by first adding $(ln(s) + r\tau)$ to the exponential function outside the integral, then by expanding the term inside, by subtracting $(ln(s) + r\tau)$, and, finally, by combining terms):

$$\frac{s}{\sigma\sqrt{2\pi\tau}}\int_{ln(K)}^{\infty} e^{\frac{-(u-ln(s)-(r+\frac{\sigma^2}{2})\tau)^2}{2\sigma^2\tau}}du. \tag{10.51}$$

Now, if we let $p = \frac{(u-ln(s)-(r+\frac{\sigma^2}{2})\tau)}{\sqrt{\sigma^2\tau}}$, then we immediately see that the first integral is equal to:

$$\frac{s}{\sqrt{2\pi}}\int_{-\infty}^{\frac{ln(s/K)+(r+\frac{\sigma^2}{2})\tau}{\sigma\sqrt{T}}} e^{\frac{-p^2}{2}}dp. \tag{10.52}$$

Based on the definition of the normal distribution, this is equal to: $s\Phi(\frac{ln(s/K)+(r+\frac{\sigma^2}{2})\tau}{\sigma\sqrt{\tau}})$. Similarly, we see that the second integral, if we make the substitution $w = \frac{(u-ln(s)-(r-\frac{\sigma^2}{2})\tau)}{\sqrt{\sigma^2\tau}}$, is equal to:

$$\frac{Ke^{-r\tau}}{\sqrt{2\pi}}\int_{-\infty}^{\frac{ln(s/K)+(r-\frac{\sigma^2}{2})\tau}{\sigma\sqrt{\tau}}} e^{-\frac{w^2}{2}}dw. \tag{10.53}$$

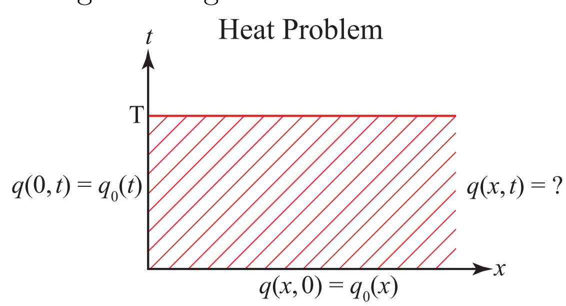
And, this is clearly equal to:

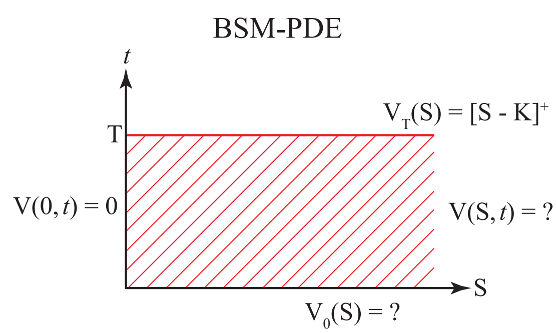
$$Ke^{-r\tau}\Phi(\frac{ln(s/K)+(r-\frac{\sigma^2}{2})\tau}{\sigma\sqrt{\tau}}). \tag{10.54}$$

Therefore, we have solved the PDE, and have our equation for the price of a call option:

$$V(s,t) = s\Phi(\frac{ln(s/K)+(r+\frac{\sigma^2}{2})\tau}{\sigma\sqrt{\tau}}) + Ke^{-r\tau}\Phi(\frac{ln(s/K)+(r-\frac{\sigma^2}{2})\tau}{\sigma\sqrt{\tau}}). \tag{10.55}$$

It is important to note the difference between the equation we just solved and the famous heat differential equation. (It looks like a flip of the initial conditions.) We can see this difference through the following two diagrams





Appendix

11. Necessary Mathematics

The two tools most often used in this paper are:

Definition 2. (Ito's Lemma in two dimensions) Let f(t,x,y) be a function whose partial derivatives f_t, f_x, f_y, f_{xx}, f_{xy}, f_{yx}, and f_{yy} are defined and continuous. Let $X(t)$ and $Y(t)$ be Ito processes (essentially of the form $dX(t) = \mu(t)dt + \sigma(t)dW(t)$, where W is a Brownian motion). The two dimensional Ito's formula in differential form is:

$$df(t, X(t), Y(t)) = f_t dt + f_x dx + f_y dy + \frac{1}{2} f_{xx} dx dx + f_{xy} dx dy + \frac{1}{2} f_{yy} dy dy \tag{11.56}$$

Please note, that this comes out of our Brownian motion rules, which tell us that

$$(dt)(dt) = 0,\ (dt)(dW) = 0,\ and (dW)(dW) = dt. \tag{11.57}$$

That is why we do not include additional terms.

We now state the Cauchy-Kovalevsky theorem, which is used throughout our research.

Theorem 2. *The initial value problem (i.v.p.) for the following m-th order nonlinear partial dierential equation in* $\mathbb{R}^{n+1}$

$$\begin{aligned} &\partial_t^m u = F(x, t, \{\partial_x^\alpha \partial_t^j\} u |\alpha| + j \le m, j < m) \\ &\partial_t^j u(x, 0) u_j(x),\ 0 \le j \le m + 1\ x \in \mathbb{R}^n, t \in \mathbb{R}, \end{aligned} \tag{11.58}$$

has a unique solution in the space of analytic functions near zero in $\mathbb{R}^{n+1}$*, if all* u_j *are analytic near zero in* $\mathbb{R}^n$*, and* F *is analytic near* $(0, 0, \{\partial_x^\alpha u_j(0)\}|\alpha| + j \le m, j < m)$

Specifically, for the main theorem, we use a special case of the Cauchy Kovalevsky Theorem. This theorem is taken from Professor Himonas' notes "Analyticity for Second-Order Linear Partial Differential Equations".

Theorem 3. *Let us consider an initial value, second-order, linear partial differential equation of the form*

$$\frac{\partial^2 u}{\partial t^2} = A(x,t)\frac{\partial^2 u}{\partial x^2} + B(x,t)\frac{\partial^2 u}{\partial x \partial t} + C(x,t)\frac{\partial u}{\partial x} + D(x,t)\frac{\partial u}{\partial t} + E(x,t)u + g(x,t) \tag{11.59}$$

$$u(x, 0) = u_0(x),\ and\ \frac{\partial du}{\partial dt}(x, 0) = u_1(x) \tag{11.60}$$

where the coefficients are analytic functions around $(0, 0)$*. Then, there is a unique analytic solution to this initial value problem near* $(0, 0)$*. If the coefficients and the forcing term are analytic in the square* $\{(x, t) \in \mathbb{R}^2 |x| < r, |t| < r\}$ *and the coefficients are bounded in absolute value by* M *and the forcing term is bounded in absolute value by* L *then the region of analyticity contains the set* $\{(x, t) \in \mathbb{R}^2 : |x + \rho t| < r(1 - \frac{M(\rho+1)}{\rho^2})\}$*, where* $\rho > 1$ *and large enough so that* $\frac{M(\rho+1)}{\rho^2}) < 1$.

Before proceeding, we will introduce a definition.

Definition 3. By a metric space, we mean a pair (X, d), where X is a set, and d is a non-negative real function $d(x, y)$ defined for all $x, y \epsilon X$ which has the following three properties:

(1) $d(x, y) = 0$ if and only if $x = y$;
(2) $d(x, y) = d(y, x)$;
(3) $d(x, z) \leq d(x, y) + d(y, z)$

The first result is quite simple, but one of the most famous results from functional analysis.

Theorem 4. *(Contraction Mapping Theorem) Let (S, d) be a complete metric space and suppose that $T : S \to S$ is a contraction. (i.e. for some modulus $\beta \in (0, 1)$, $d(Tz_1, Tz_2) \leq \beta d(z_1, z_2)$ for all $z_1, z_2 \in S$.) then T has a unique fixed point $\hat{z}$; that is, there exists a unique $\hat{z}$ such that*

$$T(\hat{z}) = \hat{z}. \tag{11.61}$$

Proof: (Existence) Choose $z_0 \in S$ and construct a sequence $\{z_n\}_{n=1}^{\infty}$ with each element in S such that $z_{n+1} = Tz_n$ so that

$$z_n = T^n z_0. \tag{11.62}$$

Remember that $T^n z = T(T^{n-1}z)$ for any $n = 1, 2, \ldots$ (with $T^0 z = z$). Since T is a contraction, we know that

$$d(z_2, z_1) = d(Tz_1, Tz_0) \leq \beta d(z_1, z_0). \tag{11.63}$$

This implies that (after repeated iteration of this argument)

$$d(z_{n+1}, z_n) \leq \beta^n d(z_1, z_0), \ n = 1, 2, \ldots \tag{11.64}$$

Therefore, for any $m > n$,

$$\begin{aligned} d(z_m, z_n) &\leq d(z_m, z_{m-1}) + \ldots + d(z_{n+2}, z_{n+1}) + d(z_{n+1}, z_n) \\ &\leq (\beta^{m-1} + \ldots + \beta^{n+1} + \beta^n) d(z_1, z_0) \\ &\leq \frac{\beta^n}{1-\beta} d(z_1, z_0), \end{aligned} \tag{11.65}$$

where the first inequality uses the triangle inequality(one of the properties of a metric), the second uses (11.64) and the third uses the geometric series and the fact that $1 + \beta + \beta^2 + \ldots > 1 + \beta + \beta^2 + \ldots + \beta^{m-n-1}$.

(11.65) implies that for large enough n, m z_m, z_n will approach each other. This implies that $\{z_n\}_{n=1}^{\infty}$ is a Cauchy sequence. By definition, since S is complete, every Cauchy sequence in S has a limit point in S and therefore $z_n \to \hat{z} \in S$.

We will now show that this $\hat{z}$ is a fixed point. For any natural number n we have (by the triangle inequality and then the definition of a contraction)

$$\begin{aligned} d(T\hat{z}, \hat{z}) &\leq d(T\hat{z}, T^n z_0) + d(T^n z_0, \hat{z}) \\ &\leq \beta d(\hat{z}, T^{n-1} z_0) + d(T^n z_0, \hat{z}). \end{aligned} \tag{11.66}$$

But, since we know that $z_n \to z$ it is clear that both terms on the right hand side go to zero as $n \to \infty$, which implies that $d(T\hat{z}, \hat{z}) = 0$, which, from the definition of a metric, implies that $T\hat{z} = \hat{z}$ which implies that $\hat{z}$ is a fixed point.

We now prove uniqueness. Suppose that there exist two fixed points $z_1, z_2 \in S$ with $z_1 \neq z_2$. This implies that

$$0 < d(z_1, z_2) = d(Tz_1, Tz_2) \leq \beta d(z_1, z_2), \tag{11.67}$$

where the equality uses the definition of a fixed point and the inequality uses the definition of a contraction. It is clear that we arrive at a contradiction since, from the definition of a contraction, $\beta < 1$. This implies that the fixed point is unique.

Before proceeding, we will introduce a few definitions.

Definition 4. A point x is called an interior point of a set M if x has an open neighborhood(i.e. in a metric space the points y such that $d(x, y) < r$ for some r) consisting of points in M. A set consisting entirely of interior points is called an **open set**.

Alternatively,

Definition 5. A point $x \in X$ is called a contact point of $M \subset X$ if every neighborhood of x contains at least one point of M. If a set M contains all of its contact points then it is considered closed.

We now prove two applications of the contraction mappings. We will use these applications in the upcoming economic proofs.

Theorem 5. *Let (S, d) be a complete metric space and $T : S \to S$ be a contraction mapping with $T\hat{z} = \hat{z}$.*

(1) *If S' is a closed subset of S, and $T(S') \subset S'$, then $\hat{z} \in S'$.*
(2) *Moreover, if $T(S') \subset S'' \subset S'$, then $\hat{z} \in S''$.*

Proof: Take $z_0 \in S'$, and construct the sequence $\{T^n z_0\}_{n=1}^{\infty}$. Each element of this sequence is in S', since $T(S') \subset S'$. The contraction mapping theorem implies that $T^n z_0 \to \hat{z}$. Using the definition of a closed set, this implies that $\hat{z} \in S'$. This completes the proof of the first claim.

From the first part, we know that $\hat{z} \in S'$. Then $T(S') \subset S'' \subset S'$ implies that $\hat{z} = T\hat{z} \in T(S') \subset S''$, proving the second claim.

It is often difficult to determine whether a particular operator is a contraction. The following theorem gives straightforward conditions for a contraction. This will help us in many economic problems. From now on, we will use the notation $(f + c)(x) \equiv f(x) + c$ where f is a real valued function and c is a real number.

Theorem 6. *(Blackwell's Sufficient Conditions for a Contraction) Let $X \subseteq \mathbb{R}^k$, and $B(X)$ be the space of bounded functions $f : X \to \mathbb{R}$ defined on X equipped with the sup norm $|| \cdot ||$. Suppose that $B'(X) \subset B(X)$, and let $T : B'(X) \to B'(X)$ be an operator satisfying the following two conditions:*

(1) *Monotonicity: For any $f, g \in B'(X)$, $f(x) \leq g(x)$ for all $x \in X$ implies $(Tf)(x) \leq (Tg)(x)$ for all $x \in X$; and*

(2) *Discounting: There exists* $\beta \in (0,1)$ *such that*

$$[T(f+c)](x) \leq (Tf)(x) + \beta c \text{ for all } f \in B(X), c \geq 0, \text{ and } x \in X. \tag{11.68}$$

Then T *is a contraction with modulus* β *on* $B'(X)$.

Proof: By the definition of the sup norm $||f-g|| = \max_{x \in X} |f(x) - g(x)|$. Then for any $f, g \in B'(X) \subset B(X)$,

$$\begin{aligned} f(x) &\leq g(x) + ||f-g|| && \text{for any } x \in X, \\ (Tf)(x) &\leq T[g + ||f-g||](x) && \text{for any } x \in X, \\ (Tf)(x) &\leq (Tg)(x) + \beta||f-g|| && \text{for any } x \in X, \end{aligned} \tag{11.69}$$

where the second line uses monotonicity of the T operator and the third line uses the discounting property. (Please remember that $||f-g||$ is a number.) We now apply the converse argument,

$$\begin{aligned} g(x) &\leq f(x) + ||g-f|| && \text{for any } x \in X, \\ (Tg)(x) &\leq T[f + ||g-f||](x) && \text{for any } x \in X, \\ (Tg)(x) &\leq (Tf)(x) + \beta||g-f|| && \text{for any } x \in X. \end{aligned} \tag{11.70}$$

Combining (11.69) and (11.70) we see that

$$||Tf - Tg|| \leq \beta||f-g||, \tag{11.71}$$

proving that T is a contraction on $B'(X)$.

Before stating the next theorem, we will introduce a few definitions

Definition 6. Let T be a set and X be the collection of open subsets of T. T is a topological space if

(1) The empty set and T are contained in X.
(2) The union of any collection of sets in T is also in T. (closed under unions)
(3) The intersection of a finite number of elements of T is also in T. (closed under finite intersection)

Definition 7. Consider a collection of open sets $(\bigcup_{a \in A} U_a)$ such that $C = (\bigcup_{a \in A} U_a)$. If $X \in (\bigcup_{a \in A} U_a)$, then C is considered to be an open cover of X.

Definition 8. We define a set X to be compact if whenever there is a collection of open sets $(\bigcup_{a \in A} U_a)$ such that $X = (\bigcup_{a \in A} U_a)$ there is a finite set B(i.e. B has a finite number of elements), where $B \subset A$, such that $X = \bigcup_{a \in B} U_a$.

Theorem 7. *(Weierstrass's Theorem) Consider the topological space* (X, τ) *and a continuous function* $f : X \to (R)$. *If* X' *is a compact subset of* (X, τ), *then* $\max_{x \in X'} f(x)$ *and* $\min_{x \in X'} f(x)$ *exist.*

Proof: Let $\{V_\alpha\}_{\alpha\in A'}$ be an open cover for $f(X')$. Since f is continuous, $f^{-1}(V_\alpha)$ is open for each $\alpha \in A'$. Since X' is compact, every open cover has a finite subcover. Therefore, there exists a finite $A'' \subset A'$ such that $X' \subset \bigcup_{\alpha\in A''} f^{-1}(V_\alpha)$. By definition $f(f^{-1}(Y'')) \subset Y''$ for any $Y'' \subset Y$ we have

$$f(X') \subset \bigcup_{\alpha\in A''} (V_\alpha), \tag{11.72}$$

and thus $\{V_\alpha\}_{\alpha\in A''}$ is a finite subcover of $\{V_\alpha\}_{\alpha\in A'}$, which implies that $f(X')$ is compact. But, it is well known that a compact subset of $\mathbb{R}$ contains a minimum and a maximum.

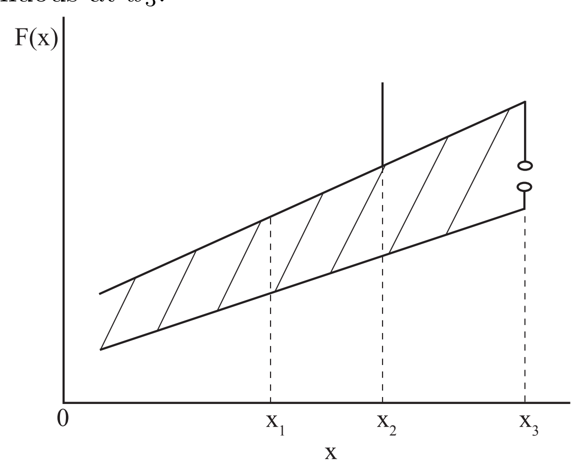
Before proceeding, we will state a few definitions.

Definition 9. The power set of X is the set of all subsets of X.

Definition 10. A correspondence F from X to the power set of Y is **upper hemicontinuous** at $x \in X$ if for every sequence $\{x_n\}_{n=1}^\infty \to x$ and every sequence $\{y_n\}_{n=1}^\infty$ with $y_n \in F(x_n)$ for every n, there exists a convergent sequence $\{y_{n_k}\}$ of $\{y_n\}_{n=1}^\infty$ such that $\{y_{n_k}\} \to y \in F(x)$.

Definition 11. F is **lower hemicontinuous** at $x \in X$ if $F(x)$ is nonempty-valued and for every $y \in F(x)$ and every sequence $\{x_n\}_{n=1}^\infty \to x$, there exists some $N \in \mathbb{N}$ and a sequence $\{y_n\}_{n=1}^\infty$ with $y_n \in F(x_n)$ for all $n \geq N$, and $\{y_n\}_{n=1}^\infty \to y$.

The following graph should help one understand the definition of upper hemicontinuous and lower hemicontinuous. The function is upper hemicontinuous and lower hemicontinuous at x_1, upper hemicontinuous but not lower hemicontinuous at x_2 and it is lower hemicontinuous but not upper hemicontinuous at x_3.



Definition 12. Let (X, d_x) and (Y, d_y) be metric spaces, and consider the correspondence F from X to the power set of Y. Then F has a closed graph at $x \in X$ if for every sequence $\{x_n, y_n\}_{n=1}^\infty \to (x, y)$ such that $y_n \in F(x_n)$ for each n, we also have $y \in F(x)$. In addition, F has a closed graph on the set X if it is closed at each $x \in X$.

The final mathematical result we will cite is Berge's Maximum Theorem.

Theorem 8. *(Berge's Maximum Theorem) Let* (X, d_x) *and* (Y, d_y) *be metric spaces. Consider the maximization problem*

$$\begin{aligned} &\sup_{y \in Y} f(x, y) \\ &\textit{subject to} \\ &y \in G(x), \end{aligned} \tag{11.73}$$

where G *is a correspondence from* X *to* Y *or a function from* X *to the power set of* Y *(not including the null set).* $f : X\, x\, Y \to \mathbb{R}$*. Suppose that* f *is continuous and* G *is compact-valued and continuous at* x*. Then*

1. $M(x) = \max_{y \in Y}\{f(x, y) : y \in G(x)\}$ *exists and is continuous at* x*, and*
2. $\pi(x) = arg \max_{y \in Y}\{f(x, y) : y \in G(x)\}$ *is nonempty-valued, compact-valued, upper hemicontinuous, and has a closed graph at* x*.*

Proof: 7 tells us that a maximum exists and that $\pi(x)$ is nonempty. (f is continuous on a compact valued set.) Now, let us consider a sequence $\{y_n\}_{n=1}^{\infty} \to y$ such that $y_n \in \pi(x)$ for each n. Since $G(x)$ is closed (remember that in a metric space compact implies closed) $y \in G(x)$. Moreover, by definition, $f(x, y_n) \in M(x)$ for each n. Since f is continuous, it follows that $f(x, y) = M(x)$. Thus, $y \in \pi(x)$ and, thus, $\pi(x)$ is closed. As a closed subset of a compact set, it is also compact. Next, we take $\{x_n\}_{n=1}^{\infty} \to x$, $\{y_n\}_{n=1}^{\infty}$, with $y_n \in G(x_n)$ for all n, with a convergent subsequence $\{y_{n_k}\} \to y$. Since $G(x)$ is upper hemicontinuous $y \in G(x)$ (continuous clearly implies upper hemicontinuous). Take any $z \in G(x)$. Since $G(x)$ is continuous and thus lower hemicontinuous, there exists $\{z_{n_k}\} \to z$ with $z_{n_k} \in G(x_{n_k})$ for all n_k. Again, $y_{n_k} \in \pi(x_{n_k})$, $M(x_{n_k}) = f(x_{n_k}, y_{n_k}) \geq f(x_{n_k}, z_{n_k})$. Moreover, since f is continuous $M(x) = f(x, y) \geq f(x, z)$. This holds for all $z \in G(x)$, $y \in \pi(x)$, and therefore $\pi(x)$ is upper hemicontinuous. Since $\pi(x)$ is upper hemicontinuous it follows immediately that at $x \in X$ every sequence $\{x_n, y_n\}_{n=1}^{\infty} \to (x, y)$ and $y_n \in F(x_n)$ for each n. It follows that $y \in F(x)$ because F has a closed set on the graph on the set X.

11.1. **Brownian Motion.** In the following section, we will be following *Real Analysis* by Folland, and *Probability Theory.* Both are graduate math textbooks.

The notation $d\omega$ is used throughout this document. This differential refers to a Brownian motion. In the upcoming section, we will attempt to define this Brownian motion in order to make sense of what we were working with. We shall consider this Brownian motion in one-dimension. We shall consider the position of a particle undergoing Brownian motion in this one-dimension at a time greater than zero to be a random variable. But, what is a random variable? In order to define a random variable, we must first define a σ-algebra, and a measurable function:

Definition 13. (σ-algebra) Given a set S, a σ-algebra F is a non-empty set of subsets of S that satisfies:

1. if $A_i \epsilon F$ for $i \geq 1$, then $\bigcup_{i=1}^{\infty} A_i \epsilon F$
2. if $A \epsilon F$, then the complement of A is in F as well

Definition 14. (Measurable Function) A function from (ω, F) to (δ, A), where ω and δ are sets and A and F are sigma algebras, is measurable if the pre-image of any set in A is in F(i.e. if the inverse function applied to any set in A is in F).

Definition 15. (Random Variable) A random variable is a real-valued measurable function from (ω, F) to (δ, R), where ω and δ are sets, and A and F are sigma algebras.

Our Brownian motion random variable must satisfy the following conditions:

(1) We will assume that $X_{t=0} = 0$. This is just a matter of normalization. (e.g. if $X_0 = \alpha$, we could look at $X_0 - \alpha = 0$ for our Brownian motion.
(2) We shall assume that the motion of the particle after some time t should only depend on X_t. (i.e. There should be no lingering affects). Therefore, we make the assumption that if $0 \leq t_0 < t_1 < \ldots < t_n$, then the random variables $X_{t_j} - X_{t_{j-1}}$ for $(1 \leq j \leq n)$ are independent.
(3) There is a constant $C > 0$ such that for $0 < s < t, X_t - X_s$ has a normal distribution with mean zero (does not have a tendency to go in either direction) and variance C(t-s). (This will be explained.)

The variance of a random variable X is $E[(X - E[X])^2]$, or essentially the expected valued of the squared differences of the mean. Expectations are essentially integration functions over every possible realization of a random variable. Therefore, it follows immediately that there is a linearity of expectations(by the linearity of integrals). This implies that the variance is equal to $E[X^2] - 2E[X]E[X] + E[X]E[X]$. Therefore, it follows immediately that the variance of a random variable X is $E[X^2] - (E[X])^2$.

Where did this last condition come from? First, we assume that the motion of X_t is homogeneous in time. (i.e. the distribution of $X_t - X_s$ should only be dependent on the length of $t - s$, not on what s and t are specifically.) Let us divide the interval [s,t] into n equal subintervals $[t_0, t_1], \ldots, [t_{n-1}, t_n]$, where $t_0 = s, t_n = t$. We can therefore write $X_t - X_s = \Sigma_{j=1}^{n}(X_{t_j} - X_{t_{j-1}})$. Then, it follows that $X_t - X_s$ is a sum of n independent identically distributed random variables (This is clear from the second condition above, and our homogeneous in time assumption). Since n is arbitrary, the central limit theorem (to be stated and proved rigorously below) tells us that the distribution of $X_t - X_s$ is normal.

11.2. **Central Limit Theorem.**

Definition 16. The Central Limit Theorem: Let $X_1, X_2, \ldots$ be independent identically distributed with an expected value of μ, and a variance of σ^2. If $S_n = X_1 + X_2 + \ldots + X_n$, then $\frac{S_n - n\mu}{\sigma n^{\frac{1}{2}}}$ converges (as $n \to \infty$) to the normal distribution.

We will attempt to prove this result. First, we will define the characteristic function, which, after some massaging, will give us all of the moments of the distribution. The characteristic function can be defined as $E[e^{itX}]$, where X is the random variable. We prove that the characteristic functions of $\frac{S_n - n\mu}{\sigma n^{\frac{1}{2}}}$ converges to the characteristic function of the normal distribution $e^{\frac{-t^2}{2}}$. Finally, we prove why this implies a convergence of the distributions.

It suffices to prove this result when $\mu = 0$. (Note: We can just consider $X_i' = X_i - \mu$) Let us look at the characteristic function $a_x(t) = E[e^{itX}]$. We know that

$$a_x(0) = E[e^{i(0)X}] = 1 \tag{11.1}$$

Due to the assumption that $\mu = 0$

$$a'_x(0) = E[iXe^{i(0)X}] = 0, \tag{11.2}$$

and

$$a''_x(0) = E[i^2X^2e^{i(0)X}] = -\sigma^2 \tag{11.3}$$

Plugging (11.1), (11.2), (11.3) into our Taylor expansion for $a_x(t)$, we see that:

$$a_{X_1(t)} = E[e^{itX_1}] = 1 - \frac{\sigma^2 t^2}{2} + o(t^2) \tag{11.4}$$

Independence implies that $E[XY] = E[X]E[Y]$. This implies that (essentially just look at what happens to the Taylor expansion if X is replaced by $X_1/(\sigma\sqrt(n))$:

$$E[e^{itS_n/\sigma n^{\frac{1}{2}}}] = E[e^{it(X_1+X_2+...+X_n)/\sigma n^{\frac{1}{2}}} = (1 - \frac{t^2}{2n} + o(\frac{1}{n}))^n \tag{11.5}$$

We shall now try to prove that if $c_n \to c$, where c is an element of the complex numbers, then $(1 + \frac{c_n}{n})^n \to e^c$. To do this, we shall first prove two lemmas.

Theorem 9. *Lemma Let $z_1, \ldots, z_n$ and $w_1, \ldots, w_n$ be complex numbers of modulus less than or equal to one. Then* $\mid \prod_{m=1}^n z_m - \prod_{m=1}^n w_m \mid \leq \sum_{m=1}^n \mid (z_m - w_m) \mid$

We attempt a proof by induction. This is clearly true for $n = 1$. Now assume that this is true for $n = k$. We will prove that this is true for $n = k + 1$

$$\mid \prod_{m=1}^{k+1} z_m - \prod_{m=1}^{k+1} w_m \mid = \mid z_{k+1} \prod_{m=1}^{k} z_m - z_{k+1} \prod_{m=1}^{k} w_m + z_{k+1} \prod_{m=1}^{k} w_m - w_{k+1} \prod_{m=1}^{k} w_m \mid \tag{11.6}$$

But, from the triangle inequality, this is less than

$$\mid z_{k+1} \prod_{m=1}^{k} z_m - z_{k+1} \prod_{m=1}^{k} w_m \mid + \mid z_{k+1} \prod_{m=1}^{k} w_m - w_{k+1} \prod_{m=1}^{k} w_m \mid \tag{11.7}$$

Factoring out, (11.7), this is clearly equivalent to:

$$\mid z_{k+1} \mid\mid (\prod_{m=1}^{k} z_m - \prod_{m=1}^{k} w_m) \mid + \mid z_{k+1} \prod_{m=1}^{k} w_m - w_{k+1} \prod_{m=1}^{k} w_m \mid \tag{11.8}$$

But, each z_i, w_i has modulus less than one. Factoring out, this implies that (11.7) is less than (remember that the multiplication of numbers less than one is less than one):

$$\mid \prod_{m=1}^{k} z_m - \prod_{m=1}^{k} w_m \mid + \mid z_{k+1} - w_{k+1} \mid \tag{11.9}$$

But, equipped with our assumption, it follows that:

$$\mid \prod_{m=1}^{k+1} z_m - \prod_{m=1}^{k+1} w_m \mid \leq \sum_{m=1}^{k+1} \mid (z_m - w_m) \mid \tag{11.10}$$

And, therefore, we have proved the lemma by induction.

Additionally, we require the assistance of another lemmma in order to prove this theorem:

Theorem 10. *Lemma If b, a complex number, with* $\mid b \mid < 1$ *then* $\mid e^{-b} - (1 - b) \mid \leq \mid b \mid^2$

To prove this we first provide a Taylor expansion of $e^{-b} - (1-b)$

$$e^{-b} - (1-b) = \frac{b^2}{2} - \frac{b^3}{3!} + \frac{b^4}{4!} - \dots \tag{11.11}$$

This immediately implies that

$$\mid e^{-b} - (1-b) \mid \leq \frac{\mid b \mid}{2}(1 + \frac{1}{2} + \frac{1}{2^2} + \dots) \tag{11.12}$$

But, we quickly recognize $(1+\frac{1}{2}+\frac{1}{2^2}+\dots)$ as a geometric series equal to 2(factor of $\frac{1}{2}$), which implies that:

$$\mid e^{-b} - (1-b) \mid \leq \mid b \mid^2 \tag{11.13}$$

Let us go back to the first lemma. Let $z_m = (1 - \frac{c_n}{n})$ and $w_m = e^{\frac{-c_n}{n}}$. Then, with the assistance of the first lemma we know that:

$$\mid (1 - \frac{c_n}{n})^n - e^{-c_n} \mid \leq \sum_{m=1}^{n} \mid (1 - \frac{c_m}{m}) - e^{\frac{-c_m}{m}} \mid \tag{11.14}$$

Now, employing the second lemma, we see immediately that (since $\sum_{m=1}^{n} \mid (1 - \frac{c_m}{m}) - e^{\frac{-c_m}{m}} \mid = \sum_{m=1}^{n} \mid (e^{\frac{-c_m}{m}} - (1 - \frac{c_m}{m})) \mid$):

$$\mid (1 - \frac{c_n}{n})^n - e^{-c_n} \mid \leq n \mid \frac{c_n}{n} \mid^2 \tag{11.15}$$

As $n \to \infty$, it is clear that this goes to zero, and therefore, we have proved the theorem. We know that the characteristic functions converge to the normal distribution characteristic function, but how does this imply that the distributions converge to the normal distribution? In order to do this, we start with a definition of tight and then state, prove, and show the implications of such a theorem:

Definition 17. If a sequence of distribution is tight, it means that for any $\epsilon > 0$, there exists an M such that $\lim_{n \to \infty} sup(1 - F_n(M) + F_n(-M)) \leq \epsilon$ where the F_n are cumulative probability distribution.

Theorem 11. *(Theorem) Let $\mu_n, 1 \leq n \leq \infty$, be probability measures with characteristic functions a_n.*

1. *If $\mu_n \to \mu_\infty$, then $a_n(t) \to a_\infty(t)$ for all t*
2. *If $a_n(t)$ converges pointwise to a limit a that is continuous at 0, then the associated sequence of distribution μ_n is tight and converges to a limit μ with characteristic function a*

$a = E[e^{itX}]$ This expectation is based on the probability measures, which converge (by assumption), and e^{itX} is bounded and continuous, therefore the characteristic functions converge.

The second part of the proof is substantially longer. A distribution is defined by its density function f. The cumulative probability distribution function $F(x)$ tells us the probability that a random variable is less than x. Written explicitly:

$$F(x) = \int_{-\infty}^{x} f(y)dy \tag{11.16}$$

Now that we have an understanding of what tight is, we should fully understand the theorem, and understand why a proof of this theorem would imply a convergence of the distributions, and, therefore, a proof of the central limit theorem. We start by noting that:

$$\int_{-u}^{u}(1-e^{itX})dt = 2u - \int_{-u}^{u}(cos(tx)+isin(tx))dt \tag{11.17}$$

We know that $[-u,u]$ is a symmetric interval, and that $sin(x)$ is an odd function. Therefore, we see immediately that:

$$\int_{-u}^{u}(1-e^{itX})dt = 2u - \frac{2sin(ux)}{x} \tag{11.18}$$

We now divide both sides by u, and integrate with respect to $\mu_n(dx)$(We remember that $\int_{-\infty}^{\infty}\mu_n(dx) = 1$, and that $a_n(t) = \int_{-\infty}^{\infty} e^{itX}\mu(dx)$

$$\frac{\int_{-u}^{u}(1-a_n(t))dt}{u} = 2(\int_{-\infty}^{\infty}(1-\frac{sin(ux)}{ux})\mu_n(dx)) \tag{11.19}$$

Let us now look at the Taylor expansion for $\frac{sin(x)}{x}$

$$\frac{sin(x)}{x} = 1 - \frac{x^2}{3!} + \frac{x^4}{5!} - \dots \tag{11.20}$$

It is immediately obvious that when $x<1$, $\frac{sin(x)}{x} < 1$. Additionally, we know that $sin(x) \leq 1$ so $\frac{sin(x)}{x} \leq 1$ for $x \geq 1$. This implies that:

$$2(\int_{-\infty}^{\infty}(1-\frac{sin(ux)}{ux})\mu_n(dx)) \geq 2(\int_{|x|\geq\frac{2}{u}}(1-\frac{1}{\mid ux \mid})\mu_n(dx)) \geq \mu_n(x :\mid x \mid > \frac{2}{u}) \tag{11.21}$$

Now let us label a_∞ as the characteristic function of μ_∞ We know that $a_\infty(t) \to 1$ as $t \to 0$, Therefore we can easily see (apply a quick L'Hopital rule) that:

$$\frac{\int_{-u}^{u}(1-a_\infty)dt}{u} \to 0 \text{ as } t \to 0 \tag{11.22}$$

The Dominated Convergence Theorem(which we will prove below) tells us that:

Theorem 12. *(Dominated Convergence Theorem) If $f_n \to f$ almost everywhere, $\mid f_n \mid \leq g$ for all n, and g is integrable, then $\int f_n d\mu \to \int f d\mu$*

With the Dominated Convergence Theorem, the fact that $a_n(t) \to a_\infty(t)$ for each t, and the fact that a_∞ is integrable, we know that there exists some N, such that for $n \geq N$:

$$3\epsilon \geq \frac{\int_{-u}^{u}(1-a_n(t))}{u}dt \geq \mu_n x :\mid x \mid > \frac{2}{u} \tag{11.23}$$

The second inequality comes from (11.19) and (11.21). This immediately implies that the sequence is tight. I am a little unsure about the remainder of the proof, but my understanding is this: The fact that the sequence is tight implies that it converges to some distribution (say μ), and from the first part of the proof we know that this distribution is the one that has characteristic function a_∞. Since, any subsequence converges to this limit we know that the entire sequence converges to μ.

In the interest of completeness, we must now prove the Dominated Convergence Theorem. To prove this, we first prove the Monotone Convergence Theorem, and Fatou's lemma.

Theorem 13. *(Monotone Convergence Theorem) If f_n is a sequence in L^+ such that $f_j \leq f_{j+1}$ for all j, and $f = \lim_{n\to\infty} f_n$ then $\int f = \lim_{n\to\infty} \int f_n$*

Before we proceed to a proof, recall that an L^p space can be defined as(where X is a set M is a σ-algebra and μ is the probability measure:

$$L^p(X, M, \mu) = f : X \to C : f \text{is measurable and } \parallel f \parallel < \infty \tag{11.24}$$

where $\parallel f \parallel$ can be defined as:

$$\parallel f \parallel = [\int \mid f \mid^p \; d\mu]^{\frac{1}{p}} \tag{11.25}$$

Let us also quickly define an indicator function. An indicator function on a set C with $E \in C$, where $x\epsilon C$ will be denoted $f_E(x)$, and will take the value 1 if x is an element of E and will take the value zero if x is not an element of E. A simple function is simply(no pun intended) a linear combination of indicator functions.

Proof: $\int f_n$ is an increasing sequence of numbers, so the limit exists($\int f$) and $\int f_n \leq \int f$ for all n, so $\lim_{n\to\infty} \int f_n \leq \int f$. Now let α be a number between zero and one., and let ϕ be a simple function with $0 \leq \phi \leq f$. Additionally, let $E_n = x : f_n(x) \geq \alpha\phi(x)$. It is clear that E_n is an increasing sequence of measurable sets whose union is the entire set (say X). This is clear because the f_ns have a limit of f. From this definition, we can immediately see that $\int f_n \geq \int_{E_n} f_n \geq \alpha \int_{E_n} \phi$. And since $E_n \in E_{n+1}$, it follows that the measure of the union of all the E_ns is the same as the measure of the $\lim_{n\to\infty} E_n$. This implies that:

$$\lim \int_{E_n} \phi = \int_X \phi \tag{11.26}$$

This immediately implies that:

$$\lim \int f_n \geq \alpha \int \phi \tag{11.27}$$

This is true for all α, and, hence, true for $\alpha = 1$. And, taking the supremum of all of the simple $\phi \leq f$, we see that $\lim \int f_n \geq \int f$, which immediately implies that $\lim \int f_n = \int f$.

Theorem 14. *(Fatou) If f_n is a sequence of nonnegative measurable functions on a measure space (X, M, μ) then*

$$\int \lim inf\, f_n \leq \lim inf \int f_n \tag{11.28}$$

Now we shall try to prove this. We know that:

$$\liminf f_n = \lim_{k\to\infty} (\inf_{n\geq k} f_n) \tag{11.29}$$

Now, it is clear that the infirmas of these f_ns are increasing as the k is increasing. This is due to the fact that the groups are getting smaller, and thus potentially smaller values(infirmas) are being thrown out of the group. Therefore, since this sequence $(\inf_{n\geq k} f_n)_{k=1}^{\infty}$ is monotone increasing, by the Monotone Convergence Theorem we have:

$$\int \liminf f_n = \int \lim_{k\to\infty} \inf_{n\geq k} f_n = \lim_{k\to\infty} \int \inf_{n\geq k} f_n \tag{11.30}$$

From the definition of infirma, we know for any $n \geq k$:

$$\int \inf {}_{n\geq k} f_n \leq \int f_n \tag{11.31}$$

Therefore, we have:

$$\int \inf {}_{n\geq k} f_n \leq \inf {}_{n\geq k} \int f_n \tag{11.32}$$

Therefore, it follows from the Monotone Convergence Theorem that:

$$\int \liminf f_n \leq \liminf \int f_n \tag{11.33}$$

and we have proved Fatou's theorem.

Theorem 15. *(Dominated Convergence Theorem) Let $f_n : X \to C$ (X, C are non-empty sets) be a sequence in L^1 such that:*
a. $f_n \to f$ almost everywhere
b.$\mid f_n \mid \leq g$, $n = 1, 2, ...,$ for for some $g \in L^1$ Then

$$\lim_{n\to\infty} \int f_n = \int \lim_{n\to\infty} f_n = \int f \tag{11.34}$$

Proof: We know that f is in L^1 because g is in L^1, and $\mid f_n \mid \leq g$. Now by taking real and imaginary parts, we cam assume that the f_ns are real-valued. Thus,

$$\mid f_n \mid \leq g \iff -g \leq f_n \leq g \iff g + f_n \geq 0 \text{and } g - f_n \geq 0 \tag{11.35}$$

We now apply Fatou's lemma to both $g + f_n$ and $g - f_n$ and we obtain:

$$\begin{aligned} \int (g+f) &= \int \liminf (g+f_n) \leq \liminf \int (g+f_n) = \int g + \liminf \int f_n \\ \int (g-f) &= \int \liminf (g-f_n) \leq \liminf \int (g-f_n) = \int g - \limsup \int f_n \end{aligned} \tag{11.36}$$

The initial equalities follow from the fact that $f_n \to f$ almost everywhere. The inequalities follow from Fatou's lemma. The second equalities follow from the fact that g is independent of n, and in order to achieve an infirma we must add the smallest value and subtract the largest one.

Anyways, these two inequalities imply that(by subtracting $\int g$, and dividing by negative one in the second inequality):

$$\limsup \int f_n \leq \int f \leq \liminf \int f_n \tag{11.37}$$

And, this clearly implies that:

$$\lim_{n\to\infty} \int f_n = \int \lim_{n\to\infty} f_n = \int f \tag{11.38}$$

Thus we have rigorously proved the Central Limit Theorem.

We would now like to explain the variance of $X_t - X_s$. The v represent the variance function

$$v(X_t - X_s) = v(\sum_1^n (X_{t_j} - X_{t_{j-1}})) \tag{11.39}$$

$$v(X_t - X_s) = v(\sum_1^n (X_{t_j} - X_{t_{j-1}})) \tag{11.40}$$

By the linearity of expectations we see that: (which is essentially the same thing as the rule in calculus that $\int (a+b)dx = \int adx + \int bdx$, when the integrals are finite)

$$v(\sum_1^n (X_{t_j} - X_{t_{j-1}})) = \sum_{j=1}^n v(X_{t_j} - X_{t_{j-1}}) \tag{11.41}$$

But the $X_{t_j} - X_{t_{j-1}}$ are identically distributed which implies that:

$$\sum_{j=1}^n v(X_{t_j} - X_{t_{j-1}}) = nv(X_{t_1} - X_{t_0}) \tag{11.42}$$

It immediately follows that if $t - s = b(t' - s')$, and b is rational(remember it needs to be broken down into intervals), then:

$$v(X_t - X_s) = bv(X_{t'} - X_{s'}) \tag{11.43}$$

And, this, of course, implies that the variance of $X_t - X_s$ is proportional to $t - s$

Additionally, we have successfully defined an abstract Brownian motion process. We will now take a break from Brownian motion, and set up some important economics concepts which will be important in the formation of my problem.

12. Economic Appendix

12.1. **Characterizing the Solution to the Utility-Maximization Problem.** For this description we follow Acemoglu's text: Introduction to Modern Economic Growth

We begin with the problem:

$$\begin{aligned} V(x(0)) = &\sup_{\{x(t+1)\}_{t=0}^{\infty}} \sum_{t=0}^{\infty} \beta^t U(x(t), x(t+1)) \\ &\text{subject to} \\ &x(t+1) \in G(x(t)), \text{ for all } t \geq 0 \\ &x(0) \text{ given} \\ &\text{where } \beta \in [0,1) \text{ and } x \in X \subset \mathbb{R}^k. \end{aligned} \tag{12.1}$$

Additionally, we let G be a correspondence from X to X. Please recall that the difference between a correspondence and a function is that a correspondence maps elements to sets while a function maps elements to elements. U is a function from X x $X \to \mathbb{R}$. U is an ordinal function that measures the utility, or happiness, that a representative feels. β is a discount factor that captures the fact that people do not care as much about future periods as the current period. We consider $x(t)$ to be a state variable, or something that describes the state of the world, and $x(t+1)$ to be a control variable, or something that a representative chooses based on the state of the world.(i.e. something he or she controls). An example of a state variable is current wealth, while an example of a control variable is how much an investor chooses to invest in for the next period. G is a correspondence that tells us what choice variables are available given a certain state variable. This is clearly an optimization problem. We want to maximize the utility over all time periods, which is represented by the infinite sum. We are trying to find the optimal sequence, $\sup\{x(t+1)\}_{t=0}^{\infty} \in X^{\infty} \in \ell^{\infty}$, in the set X^{∞} which is in the set of bounded sequences, ℓ^{∞}. To ease notation in the future, we define

$$\phi(x(t)) = \{\{x(s)\}_{s=t}^{\infty} : x(s+1) \in G(x(s)) \text{ for } s = t, t+1, \ldots\}. \tag{12.2}$$

We use the supremum notation in place of the maximum notation because there is no guarantee that we will find a feasible sequence that maximizes the sum. (i.e. we may just find feasible sequences that get arbitrarily close to an optimal sequence). This problem is very difficult to solve. We will show that given certain assumptions this optimization is the same as

$$V(x) = \sup_{y \in G(x)} \{U(x, y) + \beta V(y)\}. \tag{12.3}$$

It is clear that in this example we try to find a policy function, or a function that relates the state variable with the choice variable. That is, a policy function π will be such that $y = \pi(x)$,

To prove this equivalence and additional theorems we first state five assumptions.

(1) $G(x)$ is nonempty for all $x \in X$; and for all $x(0) \in X$ and $\mathbf{x} \in \phi(x(0))$, $\lim_{n\to\infty} \sum_{t=0}^{n} \beta^t U(x(t), x(t+1))$ exists and is finite. Please note that bold font will be used to indicate vectors in this document.

(2) X is a compact subset of $\mathbb{R}^k$, G is nonempty-valued, compact-valued, and continuous. Moreover, $U : X_G \to \mathbb{R}$ is continuous, where $X_G = \{(x, y) \in X \text{ x } X : y \in G(x)\}$.

(3) U is concave and G is convex. Additionally, if $x \neq x_1$,

$$U(\alpha x + (1-\alpha)x_1, \alpha y + (1-\alpha)y_1) > \alpha U(x, y) + (1-\alpha)U(x_1, y_1). \tag{12.4}$$

(4) For each $y \in X$, $U(., y)$ is strictly increasing in each of its first k arguments, and G is monotone in the sense that $x \leq x_1$ implies that $G(x) \subset G(x_1)$.

(5) U is continuously differentiable on the interior of its domain X_G.

Let us first define $\bar{U}$

$$\bar{U}(\mathbf{x}) = \sum_{t=0}^{\infty} \beta^t U(x(t), x(t+1)), \tag{12.5}$$

where $\mathbf{x}$ is any feasible vector in the constraint set, $\phi(x(0))$

Lemma 5. *We suppose that the first assumption holds. Then, for any $x(0) \in X$ and any $\boldsymbol{x}$ in $\phi(x(0))$,*

$$\bar{U}(\boldsymbol{x}) = U(x(0), x(1)) + \beta \bar{U}(\boldsymbol{x}'), \tag{12.6}$$

where $\boldsymbol{x}' = (x(1), x(2), \ldots)$.

Proof: Since $\bar{U}(\mathbf{x})$ exists and is finite (the first assumption)

$$\begin{aligned}\bar{U}(\mathbf{x}) &= \sum_{t=0}^{\infty} \beta^t U(x(t), x(t+1)) \\ &= U(x(0), x(1)) + \beta \sum_{s=0}^{\infty} \beta^s U(x(s+1), x(s+2)) \\ &= U(x(0), x(1)) + \beta \bar{U}(\mathbf{x}').\end{aligned} \tag{12.7}$$

Remark: It is important to note that this proof is substantially different than the one we will be considering later because we are dealing with feasible sequences and, thus, are not using sup notation.

Theorem 16. *Suppose that the first assumption holds. Then for any $x \in X$, $V(x)$ is a solution to* (12.1) *if and only if it is a solution to* (12.3).

Proof: (12.1)$\Longrightarrow$ (12.3) We begin by noting that if $\beta = 0$ the result is trivial. Therefore, it suffices to assume $\beta > 0$. Let $x(0) \in X$ and $x(1) \in G(x(0))$. Let us first assume that V is a solution to (12.1). We shall try to prove that this implies (12.3) Now, by the definition of supremum, which means least upper bound, we know that for any $\epsilon > 0$ there exists an $x'_\epsilon \in \phi(x(1))$ such that $\bar{U}(x'_\epsilon) \geq V(x(1)) - \epsilon$. Additionally, for $x_\epsilon = (x(0), x'_\epsilon, \ldots) \in \phi(x(0))$ we know from the definition of a supremum that $\bar{U}(x_\epsilon) \leq V(x(0))$. By lemma 5, we know that

$$\begin{aligned}V(x(0)) &\geq U(x(0), x(1)) + \beta \bar{U}(x'_\epsilon) \\ &\geq U(x(0), x(1)) + \beta V(x(1)) - \beta\epsilon.\end{aligned} \tag{12.8}$$

But, ϵ is arbitrary which implies that

$$V(x(0)) \geq U(x(0), x(1)) + \beta V(x(1)). \tag{12.9}$$

By the definition of supremum and the fact that we assumed V is a solution to (12.1), there exists a feasible infinite sequence, for any $\epsilon > 0$, $x'_\epsilon = (x(0), x'_\epsilon(1), x'_\epsilon(2), \ldots) \in \phi(x(0))$ such that:

$$\bar{U}(x'_\epsilon) \geq V(x(0)) - \epsilon. \tag{12.10}$$

Define $x''_\epsilon = (x'_\epsilon(1), x'_\epsilon(2), \ldots) \in \phi(x'_\epsilon(1))$ By lemma 5, we know that

$$\begin{aligned} V(x(0)) - \epsilon &\leq U(x(0), x'_\epsilon(1)) + \beta \bar{U}(x''_\epsilon) \\ &\leq U(x(0), x'_\epsilon(1)) + \beta V(x'_\epsilon(1)). \end{aligned} \tag{12.11}$$

This is true for any $\epsilon > 0$. Therefore, (12.9), (12.11) and the definition of supremum implies that V is a solution to (12.3).

(12.3)$\Longrightarrow$ (12.1) We now go in the reverse direction. We assume that V is a solution to (12.3). By the definition of supremum and under the assumption that $x(1) \in G(x(0))$,

$$V(x(0)) \geq U(x(0), x(1)) + \beta V(x(1)). \tag{12.12}$$

By recursively substituting for $V(x(1)), V(x(2)), \ldots$ we have that

$$V(x(0)) \geq \sum_{t=0}^{n} U(x(t), x(t+1)) + \beta^{n+1} V(x(n+1)). \tag{12.13}$$

Additionally, if we define $\mathbf{x} = (x(0), x(1), \ldots)$, we have

$$\bar{U}(\mathbf{x}) = \lim_{n\to\infty} \sum_{t=0}^{n} \beta^t U(x(t), x(t+1)). \tag{12.14}$$

Additionally, by the first assumption we know that

$$\lim_{n\to\infty} \beta^{n+1} V(x(n+1)) = \lim_{n\to\infty} [\beta^{n+1} \lim_{m\to\infty} \sum_{t=n}^{m} \beta^t U(x(t), x(t+1))] = 0. \tag{12.15}$$

This implies that:

$$V(x(0)) \geq \bar{U}(\mathbf{x}). \tag{12.16}$$

for any $\mathbf{x} \in \phi(x(0))$.

From the definition of supremum and the fact that V is a solution to (12.3) we know that for $\epsilon > 0$ and for any $\epsilon' = \epsilon(1-\beta) > 0$, there exists a $x_\epsilon \in G(x(0))$ such that

$$V(x(0)) \leq U(x(0), x_\epsilon(1)) + \beta V(x_\epsilon(1)) + \epsilon'. \tag{12.17}$$

We now choose $x_\epsilon(t) \in G(x(t-1))$ with $x_\epsilon(0) = x(0)$. We define $x_\epsilon = (x(0), x_\epsilon(1), x_\epsilon(2), \ldots)$. By substituting recursively, we again see that

$$\begin{aligned} V(x(0)) &\leq \sum_{t=0}^{n} U(x_\epsilon(t), x_\epsilon(t+1)) + \beta^{n+1} V(x(n+1)) + \epsilon'(1+\beta+\ldots+\beta^n) \\ &\leq \bar{U}(\mathbf{x}_\epsilon) + \epsilon. \end{aligned} \tag{12.18}$$

where the second inequality follows from the geometric series and because the $\lim_{n\to\infty} \sum_{t=0}^{n} U(x_\epsilon(t), x_\epsilon(t+1)) = \bar{U}(\mathbf{x}_\epsilon)$ and $\lim_{n\to\infty} \beta^{n+1} V(x(n+1)) = 0$. This shows that if V is a solution to (12.3) then it is a solution to (12.1). Because we have gone in both directions, we have shown the equivalence of the two problems.

Therefore, the proof is complete.

It is far more interesting to discuss the optimal paths to (12.1) and (12.3) than the optimal solutions to (12.1) and (12.3). Therefore, we prove the following theorem.

Theorem 17. *Given a feasible solution, $\hat{x} \in \phi(x(0))$, it is an optimal plan for (12.1) if and only if it is an optimal solution for (12.3).*

Proof:(12.1)$\Longrightarrow$ (12.3) Let us assume that $\hat{x}$ is a solution to (12.1). We define $\hat{x}_t = (\hat{x}(t), \hat{x}(t+1), \ldots)$. We want to show that for any $t \geq 0$

$$\bar{U}(\hat{x}_t) = V(\hat{x}(t)) \tag{12.19}$$

$t = 0$ is true for our assumption that $\hat{x}$ is a solution to (12.1). As is standard in induction proofs, we assume that (12.19) for $t = k$ and prove that it is true for $t = k+1$. From lemma 5 we know that

$$\begin{aligned} V(\hat{x}(k)) &= \bar{U}(\hat{x}_k) \\ &= U(\hat{x}(k), \hat{x}(k+1)) + \beta \bar{U}(\hat{x}_{k+1}). \end{aligned} \tag{12.20}$$

We define $x_{k+1} = (\hat{x}(k+1), x(k+2), \ldots) \in \phi(\hat{x}(k+1))$. By definition, $x_k = (\hat{x}(k), x_{k+1}) \in \phi(\hat{x}(k))$. By the definition of supremum, we know that

$$\begin{aligned} V(\hat{x}(k)) &\geq \bar{U}(x_k) \\ &= U(\hat{x}(k), \hat{x}(k+1)) + \beta \bar{U}(x_{k+1}). \end{aligned} \tag{12.21}$$

Comparing (12.20) with (12.21), we immediately see that

$$\bar{U}(\hat{x}_{k+1}) \geq \bar{U}(x_{k+1}) \tag{12.22}$$

for all $x_{k+1} \in \phi(\hat{x}_{k+1})$. Therefore, $\hat{x}_{k+1}$ attains the supremum starting from $\hat{x}(k+1)$ and we have proved (12.19) by induction.

Using lemma 5, (12.19) implies that

$$\begin{aligned} V(\hat{x}(t)) &= \bar{U}(\hat{x}_t) \\ &= U(\hat{x}(t), \hat{x}(t+1)) + \beta \bar{U}(\hat{x}_{t+1}) \\ &= U(\hat{x}(t), \hat{x}(t+1)) + \beta V(\hat{x}(t+1)). \end{aligned} \tag{12.23}$$

which implies that $\hat{x}$ is a solution to (12.3).

(12.3)$\Longrightarrow$ (12.1) We now go in the other direction. We assume that $\hat{x} \in \phi(x(0))$ is a solution to (12.3). Substituting for $V(\hat{x}(1)), V(\hat{x}(2)), \dots, V(\hat{x}(n))$ yields

$$V(x(0)) = \sum_{t=0}^{n} \beta^t U(\hat{x}, \hat{x}(t+1)) + \beta^{n+1} V(\hat{x}(n+1)). \tag{12.24}$$

Additionally, by the first assumption we know that

$$\lim_{n\to\infty} \beta^{n+1} V(x(n+1)) = \lim_{n\to\infty} [\beta^{n+1} \lim_{m\to\infty} \sum_{t=n}^{m} \beta^t U(x(t), x(t+1))] = 0. \tag{12.25}$$

which implies that

$$\begin{aligned} V(x(0)) &= \lim_{n\to\infty} \sum_{t=0}^{n} \beta^t U(\hat{x}(t), \hat{x}(t+1)) \\ &= \bar{U}(\hat{x}). \end{aligned} \tag{12.26}$$

which implies that $\hat{x}$ solves (12.1) which completes the proof.

We now want to show that this value function V is unique and that there exists an optimal plan.

Theorem 18. *(Existence of Solutions) We now assume that the first and the second assumptions hold. Then there exists a unique continuous bounded function $V : X \to \mathbb{R}$ that satisfies (12.3). Furthermore, for any $x(0) \in X$ an optimal plan $\hat{x} \in \phi(x(0))$ exists.*

Proof: We define $\mathbf{C}(X)$ to be the set of continuous functions defined on X. They have the sup norm, $||f(x)|| = \sup_{x\in X} |f(x)|$. By assumption, we know that X is compact, and therefore all functions in $C(X)$ are bounded. For $V \in C(X)$, we define the operator T as

$$TV(x) = \max_{y\in G(x)} \{U(x,y) + \beta V(y)\}. \tag{12.27}$$

A fixed point of this operator will clearly be a solution to (12.3). First, we will prove the existence of a fixed point. The right-hand side of (12.27) requires a maximization of a continuous function(the sum of two continuous functions is continuous) over a compact set, which by Weierstrass's Theorem (7) implies that there exists a solution. Thus, our operator T is well defined. Now, because $G(x)$ is a nonempty-valued continuous correspondence by our first assumption and $U(x,y)$ and $V(y)$ are continuous by assumption, Theorem (8) implies that

$$\max_{y\in G(x)} \{U(x,y) + \beta V(y)\} \tag{12.28}$$

is continuous in x. Therefore, $TV(x) \in C(X)$, and $T : C(X) \to C(X)$.

We will now try to prove that T is a contraction using Blackwell's sufficient conditions for a contraction. Theorem (6) We first show monotonicity. Let us assume that $V_1(x) \leq V_2(x)$ for any $x \in X$.

$$\begin{aligned} TV_1(x) &= \max_{y \in G(x)} \{U(x,y) + \beta V_1(y)\} \\ &\leq \max_{y \in G(x)} \{U(x,y) + \beta V_2(y)\} \\ &= TV_2(x), \end{aligned} \tag{12.29}$$

where the second line follows from the assumption. We shall now prove discounting. Let $f \in C(X)$

$$\begin{aligned} [T(f+c)](x) &= \max_{y \in G(x)} \{U(x,y) + \beta(f(y) + c)\} \\ &= \max_{y \in G(x)} \{U(x,y) + \beta f(y)\} + \beta c \\ &= Tf(x) + \beta c. \end{aligned} \tag{12.30}$$

The second line follows from the fact that $\max_{x \in X}(f(x) + c) = \max_{x \in X}(f(x)) + c$ where c is a constant.

Therefore, we know that T is a contraction. This implies that there exists a unique fixed point $V \in C(X)$, that is the solution to (12.3). By Theorem (16), this implies that it is also a solution to (12.1). We shall now show the existence of an optimal plan. By Weierstrass's Theorem, we know that since U and V are continuous and $G(x)$ is compact-valued that there exists $y \in G(x)$ that achieves the maximum. We can define the set of maximizers for (12.3) to be

$$\pi(x) = arg \max_{y \in G(x)} \{U(x,y) + \beta V(y)\}. \tag{12.31}$$

If we define $\hat{x} = (x(0), \hat{x}(1), \hat{x}(2), \ldots)$, with $\hat{x}(t+1) \in \pi(\hat{x}(t))$ for all $t \geq 0$. Then from Theorem 17, we know that this plan is also an optimal plan for (12.1).

Next, we want to show when this optimal plan is unique. To do this we first must show that the value function is strictly concave.

Theorem 19. *(Concavity of the Value Function) Suppose that our first three assumptions hold. Then the unique $V : X \to \mathbb{R}$ that solves* (12.3) *is strictly concave.*

Proof: Let $C'(X) \subset C(X)$ be the set of bounded, continuous, concave functions on X and let $C''(X) \subset C'(X)$ be the set of strictly concave functions. It is clear that $C'(X)$ is a closed subset of the complete metric space $C(X)$, but $C''(X)$ is not a closed subset (consider a set of strictly concave functions approaching a linear functional). Now, let T be defined as follows

$$TV(x) = \max_{y \in G(x)} \{U(x,y) + \beta V(y)\}. \tag{12.32}$$

As was proven in the previous theorem, T is a contraction and therefore has a unique fixed point in $C(X)$. By Theorem 5, we know that proving that $T(C'(X)) \subset C''(X) \subset C'(X)$ would be sufficient to establish that the unique fixed point is in $C''(X)$, which would, of course, imply that the value function is strictly concave. Let $V \in C'(X)$ and for $x_1 \neq x_2$ and $\alpha \in (0,1)$, let

$$x_\alpha \equiv \alpha x_1 + (1-\alpha)x_2. \tag{12.33}$$

Let $y_1 \in G(x_1)$ and $y_2 \in G(x_2)$ be solutions to problem (12.3) with the state vectors x_1 and x_2, respectively. Then

$$\begin{aligned} TV(x_1) &= U(x_1, y_1) + \beta V(y_1), \text{ and} \\ TV(x_2) &= U(x_2, y_2) + \beta V(y_2). \end{aligned} \tag{12.34}$$

The third assumption tells us that $G(x)$ is convex. Therefore, $y_\alpha \equiv \alpha y_1 + (1-\alpha)y_2 \in G(x_\alpha)$, so that

$$\begin{aligned} TV(x_\alpha) &\geq U(x_\alpha, y_\alpha) + \beta V(y_\alpha) \\ &> \alpha[U(x_1, y_1) + \beta V(y_1)] + (1-\alpha)[U(x_2, y_2) + \beta V(y_2)] \\ &= \alpha TV(x_1) + (1-\alpha)TV(x_2), \end{aligned} \tag{12.35}$$

where the first line follows from the fact that $y_\alpha \in G(x_\alpha)$ is not necessarily the maximizer starting with the state x_α. The second line uses the third assumption (the strict concavity of U and the concavity of V) and the third line uses the definition from (12.34). This implies that for any $V \in C(X')$, TV is strictly concave, and thus $T[C'(X)] \subset C''(X)$. Thus, we know that the unique fixed point, say $\hat{V}$, is in $C''(X)$, and, thus, is strictly concave.

We now state and prove the important result that the optimal plan is unique.

Theorem 20. *Suppose the first three assumptions hold. Then there exists a unique optimal plan $\hat{x} \in \phi(x(0))$ for any $x(0) \in X$. Moreover, the optimal plan can be expressed as $\hat{x}(t+1) = \pi(\hat{x}(t))$, where $\pi : X \to X$ is a continuous policy function.*

Proof: The third assumption implies that $U(x, y)$ is concave in y, and thus we know from Theorem 19 that $V(y)$ is strictly concave in y. Therefore, the right-hand side of (12.3) is strictly concave. Since $G(x)$ is convex for each $x \in X$(the third assumption), there exists a unique maximizer $y \in G(x)$ for each $x \in X$. Therefore, the policy correspondence, $\pi(x)$, is a function (it is single valued). Since it is hemicontinuous this immediately implies that it is continuous.

Before stating the next theorem it is important to review the definition of a subgradient.

First, please recall that a function $f : C \to \mathbb{R}$ is convex if for any two points x, y in the convex set C and for any $t \in [0,1]$, we have that

$$f(tx + (1-t)y) \leq tf(x) + (1-t)f(y). \tag{12.36}$$

Please recall that a set is convex if for any two points x, y in a convex set, the line joining the two points is composed of points in that convex set.

It may be helpful to first describe a subderivative. This is essentially the case of a subgradient in $\mathbb{R}$. It may be easier to understand. We will later make an extension. A subderivative of a convex function $f : U \to \mathbb{R}$ at x_0, where x_0 is a point in the open interval U, is a real number a such that

$$f(x) - f(x_0) \geq a(x - x_0). \tag{12.37}$$

Since f is a convex function, it is clear that $a \in [b, c]$ where b, c are defined as follows

$$b = \lim_{x \to x_0^-} \frac{f(x) - f(x_0)}{x - x_0}, \tag{12.38}$$

$$c = \lim_{x \to x_0^+} \frac{f(x) - f(x_0)}{x - x_0}. \tag{12.39}$$

From this it is easier to see the definition of a subgradient. A vector $\hat{x}$ is said to be a subgradient of a convex function f at a point x if

$$f(z) \geq f(x) + < \hat{x}, z - x > \tag{12.40}$$

for any z. (when f is a function on the real line) There is a geometric meaning to this as well. Consider the affine transformation(a dilation plus translation)

$$h(z) = f(x) + < \hat{x}, z - x > \tag{12.41}$$

This is a hyperplane to the convex set epi f at the point $x, f(x)$.

The next theorem gives us an important property of the value function. This will allow for important qualitative analysis.

Theorem 21. *(Monotonicity of the Value Function) Suppose that the first, second and fourth assumptions hold. Let $V : X \to \mathbb{R}$ be the unique solution to (12.3). Then V is strictly increasing in its arguments.*

Proof. Let $C'(X) \subset C(X)$ be the set of bounded, continuous, nondecreasing functions on X, and let $C''(X) \subset C(X)$ be the set of strictly increasing functions. $C'(X)$ is clearly a closed subset of the complete metric space, but, alternatively, $C''(X)$ is clearly not closed. We again define T by

$$TV(x) = \max_{y \in G(x)} \{U(x, y) + \beta V(y)\}. \tag{12.42}$$

We have already proved that this is a contraction. By Theorem 5, we know that proving that $T(C'(X)) \subset C''(X) \subset C'(X)$ would be sufficient to establish that the unique fixed point is in $C''(X)$, which would, of course, imply that the value function is a strictly increasing function. Let $V \in C'(X)$ be any nondecreasing function. In view of the fourth assumption and the fact that V is non-decreasing, $\max_{y \in G(x)}\{U(x, y) + \beta V(y)\}$ is strictly increasing which implies that $TV \in C''(X)$.

The next proof is very important because it will enable us to solve for the value function and the optimal plan.

Theorem 22. *(Differentiability of the Value Function) Suppose that the first, second, and fifth assumptions hold. Let $\pi(\cdot)$ be the policy function defined in Theorem 20 and assume that $x \in Int\ X$ and $\pi(x) \in Int\ G(x)$. Then $V(\cdot)$ is differentiable at x, with gradient given by*

$$DV(x) = D_x U(x, \pi(x)). \tag{12.43}$$

From Theorem 20 we know that the policy correspondence $\pi(x)$ is a continuous function. By hypothesis $\pi(x) \in G(x)$ and from the second assumption $G(x)$ is continuous. Thus, there exists a neighborhood $N(x)$ of x such that $\pi(x) \in Int\ G(x)$ for all $x \in N(x)$. We now define $W(\cdot)$ on $N(x)$ by

$$W(x_1) = U(x_1, \pi(x)) + \beta V(\pi(x)) \text{ for all } x_1 \in N(x). \tag{12.44}$$

We know that $V(\pi(x))$ is a fixed number independent of x_1. Additionally, by the third and fifth assumptions, we know that U is concave and differentiable. This immediately implies that $W(\cdot)$ is concave and differentiable. Furthermore, since $\pi(x) \in G(x_1)$ for all $x_1 \in N(x)$, it follows that

$$W(x_1) \leq \max_{y \in G(x_1)} \{U(x_1, y) + \beta V(y)\} = V(x_1) \text{ for all } x_1 \in N(x), \tag{12.45}$$

with equality at $x_1 = x$.

Since, $V(\cdot)$ is concave, $-V(\cdot)$ is convex, and by a standard result in convex analysis, it possesses subgradients. Moreover, any subgradient $-p$ of $-V$ at x must satisfy

$$p(x_1 - x) \geq V(x_1) - V(x) \geq W(x_1) - W(x) \text{ for all } x_1 \in N(x), \tag{12.46}$$

The first inequality uses the definition of a subgradient and the second inequality uses the fact that $W(x_1) \leq V(x_1)$, with equality at x. Therefore, every subgradient $-p$ of $-V$ is also a subgradient of $-W$. Since W is differentiable at x, its subgradient p must be unique, and we also know that any convex function with a unique subgradient at an interior point x is differentiable at x. This reasoning establishes that $-V$ is differentiable, which implies that V is differentiable. We shall show why the gradient takes that particular form in the upcoming discussion.

We shall now try to show the conditions that the solution to (12.3) must satisfy certain conditions. We start with (12.3)

$$V(x) = \sup_{y \in G(x)} \{U(x, y) + \beta V(y)\}.$$

From Theorem 19 we know that the right-hand side is strictly concave, and from Theorem 22 we know that it is also differentiable. This implies that for a $y \in Int\ G(x)$ to be an interior solution it is necessary and sufficient for it to satisfy the first-order conditions. In particular, $\hat{y}$ solutions, must solve the Euler conditions:

$$D_y U(x, \hat{y}) + \beta D(\hat{y}), \tag{12.47}$$

where the subscript y indicates that we are differentiating with respect to the second vector argument. We then differentiate with respect to the first argument. This leads us to

$$\begin{aligned} DV(x) &= D_xU(x,\hat{y}) + (D_yU(x,\hat{y}) + \beta D(\hat{y}))\frac{dy}{dx} \\ &= D_xU(x,\hat{y}). \end{aligned} \tag{12.48}$$

The second line clearly comes from (12.48).

This is not sufficient to solve the problem. We must also impose the transversality condition.

$$\lim_{t\to\infty} \beta^t \frac{\partial U(\hat{x}(t),\hat{x}(t+1))}{\partial x}\hat{x}(t) = 0. \tag{12.49}$$

This condition is absolutely necessary. Consider a Ponzi scheme. It clearly will satisfy the first order conditions in every period, but is clearly not feasible. We must impose some growth constraint. (i.e. a Ponzi scheme will blow up for t large enough.)

The final theorem we will prove tells us that if the two conditions discussed above are satisfied then we have an optimal solution.

Theorem 23. *(Euler Equations and the Transversality Condition) Let $X \subset \mathbb{R}^k_+$, and suppose that the first thru fifth assumptions hold. Then a sequence $\{\hat{x}(t)\}_{t=0}^{\infty}$ such that $\hat{x}(t+1) \in Int\ G(\hat{x}(t))$, $t = 0, 1, \ldots,$ is optimal for (12.3) (and thus (12.1)) given $x(0)$ if and only if it satisfies (12.47) and (12.49).*

Proof. (12.47) and (12.49) $\Longrightarrow$ optimality. First we shall prove sufficiency. Consider an arbitrary $x(0)$ and let $\hat{x} \equiv (x(0), \hat{x}(1), \ldots) \in \phi(x(0))$ be a feasible sequence satisfying (12.47) and (12.49). First, we will show that $\hat{x}$ yields a higher value than any other $x = (x(0), x(1), ...) \in \phi(x(0))$. For any $x \in \phi(x(0))$, define

$$\delta_x \equiv \lim_{T\to\infty} \inf \sum_{t=0}^{T} \beta^t[U(\hat{x}(t),\hat{x}(t+1)) - U(x(t),x(t+1))] \tag{12.50}$$

as the lim inf as the difference of the values of the objective function evaluated at the feasible sequences $\hat{x}$ and x as T goes to infinity. There is no guarantee that the limit exists (e.g. the limit could approach multiple values). That is why we use the limit of the infirmas (because, by definition, this will be the smallest of the limits).

From the second and fifth assumptions, U is continuous, concave and differentiable. Since U is concave

$$\begin{aligned} \delta_x \geq \lim_{T\to\infty} \inf_{\mathbf{X}} \sum_{t=0}^{T} \beta^t[&D_xU(\hat{x}(t),\hat{x}(t+1))(\hat{x}(t) - x(t)) \\ &+ D_yU(\hat{x}(t),\hat{x}(t+1))(\hat{x}(t+1) - x(t+1))] \end{aligned} \tag{12.51}$$

for any $\mathbf{x}\in \phi(x(0))$. Since $\hat{x}(0) = x(0)$, $D_xU(\hat{x}(0), \hat{x}(1))(\hat{x} - x(0)) = 0$. Using this and the fact that

$$\liminf(x_n + y_n) \geq \liminf x_n + \liminf y_n$$
$$\liminf(x_n - y_n) \geq \liminf x_n - \limsup y_n,$$

we arrive at the following inequalities

$$\begin{aligned}
\delta_x \geq & \lim_{T\to\infty} \inf \sum_{t=0}^{T} \beta^t [D_yU(\hat{x}(t), \hat{x}(t+1)) + \\
& \beta D_xU(\hat{x}(t+1), \hat{x}(t+2))](\hat{x}(t+1) - x(t+1)) \\
& - \lim_{T\to\infty} \sup \beta^T D_xU(\hat{x}(T+1), \hat{x}(T+2))\hat{x}(T+1) \\
& + \lim_{T\to\infty} \inf \beta^T D_xU(\hat{x}(T+1), \hat{x}(T+2))x(T+1).
\end{aligned}$$

Since $\hat{x}$ satisfies 12.47, all the terms in the first line are equal to zero. Additionally, since it satisfies the transversality condition, the third line equals zero. From our fourth assumption, we know that U is increasing in x, which implies that $D_xU \geq 0$. Moreover, $x \geq 0$ by hypothesis, so the last term is nonnegative, which implies that $\delta_x \geq 0$ for any $x \in \phi(x(0))$. Therefore, $\hat{x}$ yields a higher value than any other feasible sequence.

We shall now prove necessity. We define

$$\delta'_x \equiv \lim_{T\to\infty} \sup \sum_{t=0}^{T} \beta^t [U(\hat{x}(t), \hat{x}(t+1)) - U(x(t), x(t+1))] \geq 0. \tag{12.52}$$

Suppose that $\{\hat{x}(t+1)\}_{t=0}^{\infty}$, with $\hat{x}(t+1) \in Int\ G(\hat{x}(t))$ for all t constitutes an optimal plan, which implies that δ'_x is nonnegative for any $\mathbf{x}\in \phi(x(0))$. Consider $\mathbf{x}\in \phi(x(0))$ such that $x(t) = \hat{x}(t) - \epsilon z(t)$, where $z(t) \in \mathbb{R}^k$ for each t and ϵ is a real number. For ϵ sufficiently small, such an $\mathbf{x}\in \phi(x(0))$ can be found because $\hat{x}(t+1) \in G(\hat{x}(t))$ for all t and G is concave and continuous. Using Taylor's theorem, little o-notation and the fact that

$$\limsup_{\mathbf{x}}(x_n + y_n) \leq \limsup x_n + \limsup y_n, \tag{12.53}$$

we see that

$$\begin{aligned}
\delta'_x \equiv & \lim_{T\to\infty} \sup \sum_{t=0}^{T} \beta^t [D_xU(\hat{x}(t), \hat{x}(t+1))\epsilon z(t) \\
& + D_yU(\hat{x}(t), \hat{x}(t+1))\epsilon z(t+1)] \\
& + \lim_{T\to\infty} \sup \sum_{t=0}^{T} \beta^t o(\epsilon, t).
\end{aligned}$$

Please remember that $o(\epsilon, t)$ goes to zero faster than $\epsilon \to 0$ as $\epsilon \to 0$ for any t. If 12.47 is violated at some t_1, then we could take $y(t) = 0$ for all $t \neq t_1$ and choose ϵ and $z(t_1)$ such that $D_xU(\hat{x}(t_1), \hat{x}(t_1+1))\epsilon z(t_1) < 0$ and $\epsilon \to 0$. This will guarantee that $\delta'_x < 0$, which is a contradiction. Therefore, 12.47 must be satisfied.

Next, we suppose that 12.47 is satisfied but that 12.49 is violated. We choose $x(t) = (1-\epsilon)\hat{x}(t)$ and we repeat the same steps as above. This leads us to

$$\begin{aligned}\delta'_x \leq -\epsilon \lim_{T\to\infty} \inf \beta^T D_x U(\hat{x}(T), \hat{x}(T+1))\hat{x}(T+1) \\ + \lim_{T\to\infty} \sup \sum_{t=0}^{T} \beta^t o(\epsilon, t),\end{aligned} \tag{12.54}$$

where all the other terms have been canceled by 12.47. We now want to prove that

$$\lim_{\epsilon\to 0} \lim_{T\to\infty} \sup \sum_{t=0}^{T} \beta^t \frac{o(\epsilon,t)}{\epsilon} = 0 \tag{12.55}$$

By definition, $\lim_{\epsilon\to 0} \frac{o(\epsilon,t)}{\epsilon} = 0$ for each t and there exists $M < \infty$ such that for ϵ sufficiently small, $|o(\epsilon,t)/\epsilon| < M$ for each t. For any $a > 0$, choose $\bar{T}$ such that $\frac{M\beta^{T+1}}{1-\beta} \leq \frac{a}{2}$ for all $T > \bar{T}$. Then

$$\lim_{T\to\infty} \sup \sum_{t=0}^{T} \beta^t |\frac{o(\epsilon,t)}{\epsilon}| \leq \sum_{t=0}^{\bar{T}} \beta^t |\frac{o(\epsilon,t)}{\epsilon}| + \frac{a}{2} \tag{12.56}$$

for ϵ sufficiently small. The inequality follows from our choices and the geometric series.

Furthermore, since $\sum_{t=0}^{\bar{T}} \beta^t |o(\epsilon,t)/\epsilon|$ is a finite sum, there exists a $\bar{\epsilon}$ such that for $\epsilon \leq \bar{\epsilon}$, $\sum_{t=0}^{\bar{T}} \beta^t |o(\epsilon,t)/\epsilon| < \frac{a}{2}$. This implies that the left-hand side of (12.56) is less than a. Since a is arbitrary, we see that (12.55) follows. Note that if the transversality condition is violated then the first term in (12.54) can be made either negative or positive by choosing ϵ to be either positive or negative. This combined with (12.55) implies that $\delta'_x < 0$, which is a contradiction. Therefore we have finished the necessity of the two conditions and have therefore finished the proof.

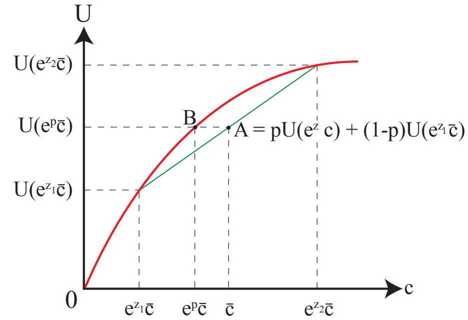
12.2. **Elasticity of Intertemporal Substitution and The Coefficient of Relative Risk Aversion.** Throughout this section, we will be following Professor Cosimano's notes, and several economic textbooks including *Asset Pricing*, *Macroeconomics* and *Intermediate Microeconomics*. We start by making the assumption that our utility function is defined recursively. Let U represent utility, and c represent consumption. We call W an aggregator, because it combines current consumption with future expected utility.

$$U(c_0, c_1, \ldots) = W(c_0, U(c_1, c_2, \ldots)) \tag{12.1}$$

We want to talk about the coefficient of risk aversion.

Definition 18. (Risk Averse) We say that an investor is risk averse if they will not accept an actuarially fair (one with an expected value of zero) gamble.

Now, let U be the utility function. A utility function is a measure of happiness that a representative investor (an investor that represents an average investor in the economy) receives from consumption. For now, we will just state three properties of this utility function. First, we assume that the first derivative is positive (more consumption equals more happiness). Second, we assume that the second derivative is negative.(The second hamburger does not provide as much utility as the first) Therefore, the utility function is concave. Finally, we assume that the utility function satisfies the transitive property. That is, if bundle A provides more utility than bundle B, and bundle B provides more utility than bundle C, then bundle A provides more utility than bundle C.



We define the risk premium p as the amount an investor must pay so that (where a is a probability, and thus an element of $[0, 1]$):

$$U(\bar{C}e^p) = aU(e^{z_1}\bar{C}) + (1-a)U(e^{z_2}\bar{C}). \tag{12.2}$$

Definition 19. p will allow us to define a certainty equivalent.(one where you adjust the C values in order to make the expected utility equal to a certain utility.)

We will now try to derive an expression for p:

First, we note that (this approximation, which comes directly from Taylor series, will be used throughout the proof):

$$e^x \approx 1 + x, \tag{12.3}$$

We will now take a first-order Taylor approximation of the left side of (12.2)

$$U(\bar{C}e^p) \approx U(\bar{C}) + U'(\bar{C})(\bar{C} - e^p\bar{C}). \tag{12.4}$$

Using (12.3), we see that:

$$U(\bar{C}e^p) \approx U(\bar{C}) + U'(\bar{C})(-p\bar{C}). \tag{12.5}$$

We now take a second-order Taylor approximation to the right side of (12.2):

$$aU(e^{z_1}\bar{C}) + (1-a)U(e^{z_2}\bar{C}) \approx a[U(\bar{C}) + U'(\bar{C})(e^{z_1}\bar{C} - \bar{C}) + \frac{1}{2}U''(\bar{C})(e^{z_1}\bar{C} - \bar{C})^2] \tag{12.6}$$

$$+ (1-a)[U(\bar{C}) + U'(\bar{C})(e^{z_2}\bar{C} - \bar{C}) + \frac{1}{2}U''(\bar{C})(e^{z_2}\bar{C} - \bar{C})^2]. \tag{12.7}$$

Using (12.3), we see that:

$$aU(e^{z_1}\bar{C}) + (1-a)U(e^{z_2}\bar{C}) \approx a[U(\bar{C}) + U'(\bar{C})(z_1\bar{C}) + \frac{1}{2}U''(\bar{C})(z_1\bar{C})^2] \tag{12.8}$$

$$+ (1-a)[U(\bar{C}) + U'(\bar{C})(z_2\bar{C}) + \frac{1}{2}U''(\bar{C})(z_2\bar{C})^2]. \tag{12.9}$$

This can be rewritten as:

$$U(\bar{C}) + U'(\bar{C})(az_1\bar{C} + (1-a)(z_2\bar{C})) + \frac{1}{2}U''(\bar{C})(a(z_1\bar{C})^2 + (1-a)(z_2\bar{C})^2), \tag{12.10}$$

which is equivalent to:

$$U(\bar{C}) + E[z](\bar{C})U'(\bar{C}) + \frac{1}{2}(\bar{C})^2E[z^2]U''(\bar{C}). \tag{12.11}$$

But, we know that $E[z] = 0$, and that $E[z^2] = V(z)$, where V represents the variance. Therefore, we know that:

$$U(\bar{C}) + E[z]\bar{C}U'(\bar{C}) + \frac{1}{2}E[z^2](\bar{C})^2U''(\bar{C}) \approx U(\bar{C}) + \frac{1}{2}(\bar{C})^2V(z)U''(\bar{C}) \tag{12.12}$$

Equating (12.12) and (12.5), we arrive at an expression for the risk premium p

$$p \approx \frac{-1}{2}\bar{C}V(z)\frac{U''(\bar{C})}{U'(\bar{C})}. \tag{12.13}$$

We immediately see that this is directly proportional to the coefficient of relative risk aversion.

Definition 20. The coefficient of relative risk aversion $R(C)$ can be defined as $R(C) = \frac{-CU''(C)}{U'(C)}$.

Definition 21. A utility function exhibits constant relative risk aversion if $R(C)$ is constant.

The coefficient of absolute risk aversion can be derived by employing the same exact method, and substituting an additive random variable $\bar{C} + z$ for the multiplicative one $e^z\bar{C}$. One will arrive at a risk premium $p = \frac{-1}{2}V(z)\frac{-U''(C)}{U'(C)}$.

Definition 22. The coefficient of absolute risk aversion $A(C)$ can be defined as $A(C) = \frac{-U''(C)}{U'(C)}$.

We will now try to derive a function that has a constant relative risk aversion. We start by assuming that absolute risk aversion has the following functional form ($B > 0$):

$$A(C) = \frac{1}{A + BC}. \tag{12.14}$$

This functional form is convenient and exhibits decreasing absolute risk aversion(this is a necessary property since people like Bill Gates are more willing to take on risk than your average middle class earner):

$$A'(C) = \frac{-B}{(A + BC)^2}. \tag{12.15}$$

It is clear that if there is to be constant relative risk aversion then A must equal zero. Remembering that $A(C) = \frac{-U''(C)}{U'(C)}$, we now integrate $A(C)$, and the functional form we assumed for it (with $A = 0$) and arrive at the expression:

$$\ln(U'(C)) = \frac{-\ln(BC)}{B} + D. \tag{12.16}$$

This implies that:

$$U'(C) = e^D(BC)^{\frac{-1}{B}}. \tag{12.17}$$

Integrating again, we arrive at an expression for $U(C)$

$$U(C) = e^D\frac{(BC)^{1-\frac{1}{B}}}{1 - \frac{1}{B}} + R, \tag{12.18}$$

Where R is a constant. Note, R is not affected by a change in C, and therefore we can ignore it when looking at our functional form. Essentially, the functional form we have derived for a utility function with a constant coefficient of relative risk aversion is:

$$U(C) = AC^{\alpha} \tag{12.19}$$

Where A, and α are constants.

Before we continue, let us lay some more economic foundation.

Definition 23. Elasticity is the ratio of the percent change in one variable to the percent change in another variable.

For example, the price elasticity of demand tells us the percent change of demand with a one-percent increase in the price. In this example, it is clear that the magnitude of the price elasticity of demand for water will be lower than the magnitude of the price elasticity of demand for ketchup. People need water to survive, but they may be willing to substitute away from ketchup.

We shall be looking at the elasticity of substitution.

Definition 24. The elasticity of substitution describes how the ratio of two inputs changes with respect to a change in their marginal utilities, or the derivative of utility (in this case with respect to consumption).

In differential notation this can be written as:

$$d(\frac{c}{z})/\frac{c}{z}/d(\frac{U_c}{U_z})/\frac{U_c}{U_z}. \tag{12.20}$$

One can quickly see that this is equivalent to:

$$\frac{d(ln(\frac{c}{z}))}{d(ln(\frac{U_c}{U_z}))}. \tag{12.21}$$

This is not sufficient to define the intertemporal elasticity of substitution, however. It is clear that we can increase the ratio of $\frac{c}{z}$, by changing z or c. The way in which we change this (e.g. increasing c, or decreasing z), could effect the marginal utility ratio in different ways. This problem is solved by restricting our class of utility functions to homothetic functions.

Definition 25. A homothetic function is one such that (where U is a utility function, λ is a scalar, and c is a consumption process) $U(\lambda c) \geq U(\lambda c')$ implies $U(c) \geq U(c')$.

Let us now consider time separable homothetic utility functions. For example, $U(c_1, c_2, \ldots, c_n) = \Sigma_i^n(\beta^i u(c_i))$ where i represents a specific time period, and u could, for example, take the form of (12.19). This is a very common and popular class of utility functions. We will show that the intertemporal elasticity of substitution and the coefficient of relative risk aversion are inter-related for this utility function. To do this we set out to compute the intertemporal elasticity of substitution.

First, since the utility function is time separable, changing a c_i will have no impact on $u(c_j)$ for $i \neq j$. We will just consider a change in c_i to change $d(\frac{c_i}{c_j})$. Therefore, $d(\frac{c_i}{c_j}) = \frac{d(c_i)}{c_j}$, and $d(\frac{U_{c_i}}{U_{c_j}}) = \frac{d(u'(c_i))}{u'(c_2)}$. This implies that the intertemporal elasticity of substitution is equal to

$$\begin{aligned} & d(\frac{c_i}{c_j})/\frac{c_i}{c_j}/d(\frac{(U_{c_i})}{U_{c_j}})\frac{U_{c_i}}{U_{c_j}} = \\ & \frac{\frac{d(c_i)}{c_j}}{\frac{c_j}{c_i}}/\frac{d(U_{c_i})}{U_{c_j}}\frac{U_{c_j}}{U_{c_i}} = \frac{d(c_i)}{c_i}/\frac{d(U_{c_i})}{U_{c_i}} = \\ & = \frac{d(c_i)}{d(U_{c_i})}\frac{U_{c_i}}{c_i} = \frac{1}{u''(c_i)}\frac{u'(c_i)}{c_i} \end{aligned} \tag{12.22}$$

From the last equality, we can clearly see that the intertemporal elasticity of substitution, in this situation, is equal to the negative of the inverse of the coefficient of relative risk aversion. This is not true empirically.

Many economists argue that the fact that the CRRA and the intertemporal elasticity of substitution are tied together places too large of a restriction on the model. Therefore, Epstein and Zin used a utility function that separated these two variables. (We will use the same utility function.)

Epstein and Zin wanted their utility function to exhibit some sort of constant elasticity of substitution.(This is a result of convenience and data-fitting). Therefore, they picked the following aggregator W.

$$W(c,z) = [(1-\beta)c^p + \beta z^p]^{\frac{1}{p}} \tag{12.23}$$

Please note that p and β are constants.

Taking partial derivatives, first with respect to c, and then with respect to z we see that:

$$\begin{aligned} W_c &= \frac{1}{p}[(1-\beta)c^p + \beta z^p]^{\frac{1}{p}-1}(1-\beta)p(c^{p-1}) \\ &= [(1-\beta)c^p + \beta z^p]^{\frac{1}{p}-1}(1-\beta)(c^{p-1}). \end{aligned} \tag{12.24}$$

$$\begin{aligned} W_z &= \frac{1}{p}[(1-\beta)c^p + \beta z^p]^{\frac{1}{p}-1}(\beta)p(z^{p-1}) \\ &= [(1-\beta)c^p + \beta z^p]^{\frac{1}{p}-1}(\beta)(z^{p-1}). \end{aligned} \tag{12.25}$$

This implies that:

$$\frac{W_c}{W_z} = \frac{1-\beta}{\beta}(\frac{c}{z})^{p-1}. \tag{12.26}$$

Taking the natural log of both sides we see that:

$$\ln(\frac{W_c}{W_z}) - \ln(\frac{1-\beta}{\beta}) = (p-1)\ln(\frac{c}{z}) \tag{12.27}$$

This immediately implies that the elasticity of substitution is constant since:

$$\frac{d(\ln(\frac{c}{z})}{d(ln(\frac{W_c}{W_z}))} = \frac{1}{p-1} \tag{12.28}$$

And, as stated earlier, p is a constant.

Additionally, they used an expected utility function with a constant relative risk aversion as a certainty equivalent.(They raised the function to a power to preserve homotheticity. They let $z = E[x^a]^{\frac{1}{a}}$, where E is the expectation operator.(Please note that this z brings in a coefficient of relative risk aversion that can be separated from the intertemporal elasticity of substitution.) This immediately leads us to the utility function

$$U_t = [(1-\beta)c_t^p + \beta(E_t[U_{t+1}^a])^{\frac{p}{a}}]^{\frac{1}{p}} \tag{12.29}$$

This utility is clearly in discrete time. In continuous time, this process assumes the form:

$$dU_t = (-f(c_t, U_t) - A(U_t)\frac{\sigma_V^2(t)}{2})dt + \sigma_v(t)dB_t \tag{12.30}$$

We know that we can apply the implicit function theorem to (12.29), in order to get a solution(Please note: We do not know the boundary points of this solution, and that is an unresolved issue.)

$$m(bU_{t+1} \mid F_t) = G(c_t, U_t) \tag{12.31}$$

where the b represents the distribution(all it's possible values and probabilities) of U, and $\mid F_t$ essentially means given the information at time t. (e.g. we know that $m(\ U_t \mid F_t) = U_t$ because it is a known quantity.) Subtracting this quantity from both sides, we see that:

$$m(bU_{t+1} \mid F_t) - m(bU_t \mid F_t) = G(c_t, U_t) - U_t \tag{12.32}$$

Changing the interval from one to dt we see that

$$m(bU_{t+dt} \mid F_t) - m(bU_t \mid F_t) = G(c_t, U_t, dt) - G(c_t, U_t, 0) \tag{12.33}$$

We now assume differentiability, to show that:

$$\frac{d(m(bU_{t+s} \mid F_t)) \mid_{s=0}}{ds} = -f(c, U) \tag{12.34}$$

To make this transformation, Duffie and Epstein assume that U has a stochastic differential representation of the form

$$dU_t = \mu_t dt + \sigma_t dB_t \tag{12.35}$$

Somehow, Duffie and Epstein differentiate $m(\ U_{t+s} \mid F_t) \mid_{s=0}$, and show that:

$$m(bU_{t+s} \mid F_t) \mid_{s=0} = \mu_t + \frac{1}{2}A(U_t)\sigma_t\sigma_t \tag{12.36}$$

which immediately implies that:

$$\mu_t = -f(c_t, U_t) - A(U_t)\sigma_t\sigma_t \tag{12.37}$$

They then define, where $v = u$:

$$f(c, U) = \frac{\beta(c^p - U^p)}{pU^{p-1}} \text{ and } m(\ U) = [E(U^\alpha)]^{\frac{1}{\alpha}} \text{ and } A(U) = \frac{-\gamma}{U} \tag{12.38}$$

and the equivalent aggregator $(\bar{f}, \bar{A})$

$$\bar{f}(C, U) = \frac{\beta}{p}\frac{C^p - ((1-\gamma)U)^{\frac{1-p}{1-\gamma}}}{((1-\gamma)U)^{\frac{p}{1-\gamma}-1}} \tag{12.39}$$

$$\text{and } \bar{A} = 0 \tag{12.40}$$

To see why this is what we might expect, and to derive an important condition, which will be used later in the paper, let us go back and analyze the recursive utility function (this time with an interval of dt, and with $1-\gamma$ substituted for a, and $1-p$ substituted for p). We will be algebraically manipulating this equation to arrive at a condition:

12.3. **Condition for the Utility Function.**

$$U(t) = [(1-e^{-\beta dt})c_t^{1-p} + e^{-\beta dt}(E_t[U(t+dt)^{1-\gamma}])^{\frac{1-p}{1-\gamma}}]^{\frac{1}{1-p}} \tag{12.41}$$

This is equivalent to:

$$\frac{U(t)^{1-p}}{1-p} = (1-e^{-\beta dt})\frac{c_t^{1-p}}{1-p} + e^{-\beta dt}\frac{(E_t[U(t+dt)^{1-\gamma}])^{\frac{1-p}{1-\gamma}}}{1-p} \tag{12.42}$$

We will now define some functions to make the algebra easier:

$$u_\alpha(x) \equiv \begin{cases} \frac{x^{1-\alpha}}{1-\alpha} & \text{if } 0 < \alpha \neq 1\,, \\ \ln(x) & \text{if } \alpha = 1\,, \end{cases} \tag{12.43}$$

and

$$g(x) = u_p(u_\gamma^{-1}(x)) \equiv \begin{cases} \frac{((1-\gamma)x)^{1/\theta}}{1-p} & \text{if } \gamma, p \neq 1\,, \\ u_\rho(e^x) & \text{if } \gamma = 1, p \neq 1 \\ \ln((1-\gamma)x)/(1-\gamma) & \text{if } p = 1, \gamma \neq 1\,. \end{cases} \tag{12.44}$$

where

$$\theta = \frac{1-\gamma}{1-p}.$$

Additionally,

$$J(C(t),t) = u_\gamma(U(C(t),t)) = \begin{cases} \frac{U(C(t),t)^{1-\gamma}}{1-\gamma} & \text{if } 0 < \gamma \neq 1\,, \\ \ln(U(t)) & \text{if } \gamma = 1\,, \end{cases}$$

which implies that:

$$g(J(C(t),t)) = u_p(u_\gamma^{-1}(u_\gamma(U(C(t),t)))) = u_p(U(C(t),t)) = \begin{cases} \frac{U(t)^{1-\rho}}{1-\rho} & \text{if } 0 < \rho \neq 1\,, \\ \ln(U(C(t),t)) & \text{if } \rho = 1\,. \end{cases}$$

We also calculate:

$$g(E_t[J(t+dt)]) = u_p(u_\gamma^{-1}(E_t(J(t+dt)))) = \begin{matrix} \frac{((1-\gamma)E_t[J(t+dt)])^{1/\theta}}{1-\theta} & \text{if } \gamma, p \neq 1\,, \\ u_\rho(e^{E_t[J(t+dt)]}) & \text{if } \gamma = 1, p \neq 1 \\ \ln((1-\gamma)E_t[J(t+dt)])/(1-\gamma) & \text{if } p = 1, \gamma \neq 1\,. \end{matrix} \tag{12.45}$$

This implies that (12.42) is equivalent to:

$$g(J(t)) = (1-e^{-\beta dt})u_p(c(t)) + e^{-\beta dt}g(E_t[J(t+dt)]). \tag{12.46}$$

This can be rewritten as:

$$g(J(t)) = (1-e^{-\beta dt})u_p(c(t)) + e^{-\beta dt}g(E_t[J(t)+dJ(t)]). \tag{12.47}$$

From our Taylor formula for e^x, we know that around $dt = 0$ (Please note that we use the fact that $(dt)^2 = 0$)

$$e^{-\beta dt} = e^{-\beta 0} - \beta e^{-\beta 0}dt = 1 - \beta dt \tag{12.48}$$

This implies that:

$$g(J(t)) = (\beta dt)u_p(c(t)) + e^{-\beta dt}g(E_t[J(t)+dJ(t)]) \tag{12.49}$$

We now take a Taylor approximation of $e^{-\beta dt}g(E_t[J(t)+dJ(t)])$ around $dt=0$ and $dJ(t)$:

$$e^{-\beta dt}g(E_t[J(t)+dJ(t)]) = e^{-\beta 0}g(J(t)+0)+e^{-\beta 0}g'(J(t))(E_t[dJ(t)]-0)+(-\beta e^{-\beta 0})g(J(t))(dt-0). \tag{12.50}$$

Please remember that $E_t[J(t)] = J(t)$ because the expectation of the present is what is actually occurring! Plugging (12.50) into (12.49), we see that:

$$g(J(t)) = (\beta dt)u_p(c(t)) + e^{-\beta 0}g(J(t)) + g'(J(t))E_t[dJ(t)] + (-\beta dt)g(J(t)) \tag{12.51}$$

Canceling, we see that:

$$(\beta dt)u_p(c(t)) + g'(J(t))E_t[dJ(t)] + (-\beta dt)g(J(t)) = 0 \tag{12.52}$$

Solving for $E_t[dJ(t)]$, we see that:

$$E_t[dJ(t)] = \frac{-\beta(u_p(c(t))dt - g(J(t)))dt}{g'(J(t))} \tag{12.53}$$

This agrees with our equivalent aggregator((12.39)), because

$$E_t[dJ(t)] + \bar{f}(c(t),U(t)) + \frac{1}{2}\bar{A}(U(t))\sigma_v(t)\sigma_v(t) = 0 \tag{12.54}$$

except that p is used in place of $1-p$.

From this condition, it is clearer to see where f and A come from.

12.4. **Euler Condition.** Throughout this section, we will be following the work of *Asset Pricing* I will use (12.53) in my research. The other big equation that I will derive and use is the Euler equation in continuous time. This is a more intuitive description of the Euler condition than the one presented earlier. We shall first try to derive the discrete time version of this formula:

$$P_t = E_t[\beta\frac{u'(C_{t+1})}{u'(C_t)}x_{t+1}] \tag{12.1}$$

where P_t is the price of the asset at time t, u is utility, C is consumption, and x_{t+1} is the payoff at the asset at the payoff of the asset at time $t+1$. Intuitively this means that the price of an asset today is equal to the expected discounted value of the payoff of the asset in the next time period.

We define a utility function over both times t and $t+1$

$$U(C_t, C_{t+1}) = u(C_t) + \beta E_t[u(C_{t+1})] \tag{12.2}$$

Naturally, we try to maximize this function. But, we impose two restrictions:

$$C_t = y_t - P_t a \tag{12.3}$$

and

$$C_{t+1} = y_{t+1} + x_{t+1}a. \tag{12.4}$$

where a represents the amount of an asset purchased, and y represents the consumption level if the investor bought none of the asset.

Thus, we are left with the objective to maximize (we will choose a):

$$u(y_t - P_t a) + E_t[\beta u(y_{t+1} + x_{t+1}a)]. \tag{12.5}$$

This leads to the first order condition:

$$P_t u'(C_t) = E_t[\beta u'(C_{t+1}) x_{t+1}] \tag{12.6}$$

But, this is equivalent to (12.1).(Please note that P_t can be brought inside the expectation operator because $E[P_t] = P_t$)

We now let $x_{t+1} = P_{t+1} + D_{t+1}$. That is, we let the payoff at time $t+1$ be equal to the price of the asset at that time plus the dividends received at that time. Thus, our equation becomes

$$u'(C_t)P_t = E_t[\beta u'(C_{t+1})D_{t+1} + P_{t+1}] \tag{12.7}$$

Decreasing the time interval from 1 to Δt so that we arrive at:

$$\begin{aligned} u'(C(t))P(t) =& e^{-\delta \Delta t} E_t[u'(C(t+\Delta t)) \cdot (D(t+\Delta t) + P(t+\Delta t))] \\ =& e^{-\delta \Delta t} E_t[u'(C(t+\Delta t))D(t+\Delta t)] + E_t[u'(C(t+\Delta t))P(t+\Delta t)]. \end{aligned} \tag{12.8}$$

Now multiply by $e^{-\delta t}$ to obtain

$$e^{-\delta t} u'(C(t))P(t) = e^{-\delta(t+\Delta t)} E_t[u'(C(t+\Delta t))D(t+\Delta t)] + E_t[u'(C(t+\Delta t))P(t+\Delta t)]. \tag{12.9}$$

Bring all terms to the right hand side of the equal side to yield

$$\begin{aligned} 0 = & E_t[e^{-\delta(t+\Delta t)} u'(C(t+\Delta t))D(t+\Delta t)] \\ & + E_t[e^{-\delta(t+\Delta t)} u'(C(t+\Delta t))P(t+\Delta t) - e^{-\delta t} u'(C(t))P(t)]. \end{aligned} \tag{12.10}$$

In continuous time dividends are paid at a rate $D(t)$ per unit of time subject to a stochastic process. For example, assume

$$dD(t) = D(t)dt + D(t)\sigma d\omega(t),$$

where $d\omega$ is Brownian motion. Integrating this relation we obtain

$$D(t+\Delta t) = \int_t^{t+\Delta t} D(s)ds + \int_t^{t+\Delta t} D(s)\sigma d\omega(s).$$

Here, we use the fact that the investor does not receive a dividend payment $D(t)$ at time t, since they are not the owner until the next instant. Now multiply by $e^{-\delta(t+\Delta t)} u'(C(t+\Delta t))$ to find

$$\begin{aligned} e^{-\delta(t+\Delta t)} u'(C(t+\Delta t))D(t+\Delta t) =& e^{-\delta(t+\Delta t)} u'(C(t+\Delta t))(\int_t^{t+\Delta t} D(s)ds \\ & + \int_t^{t+\Delta t} D(s)\sigma d\omega(s)). \end{aligned}$$

Finally, take condition expectations

$$\begin{aligned} E_t\left[e^{-\delta(t+\Delta t)} u'(C(t+\Delta t))D(t+\Delta t)\right] =& E_t\left[e^{-\delta(t+\Delta t)} u'(C(t+\Delta t)) \int_t^{t+\Delta t} D(s)ds\right] \\ & + E_t\left[e^{-\delta(t+\Delta t)} u'(C(t+\Delta t)) \int_t^{t+\Delta t} D(s)\sigma d\omega(s)\right]. \end{aligned} \tag{12.11}$$

The last term is zero since $E_t\left[d\omega(s)\right] = 0$.

In these circumstances, we can substitute (12.11) into (12.10) to find

$$0 = E_t\left[e^{-\delta(t+\Delta t)}u'(C(t+\Delta t))\int_t^{t+\Delta t} D(s)ds\right] + E_t[e^{-\delta(t+\Delta t)}u'(C(t+\Delta t))P(t+\Delta t) - e^{-\delta t}u'(C(t))P(t)] \tag{12.12}$$

Let $\Delta t \to 0^+$ so that

$$0 = e^{-\delta t}u'(C(t))D(t)\,dt + E_t[d(e^{-\delta t}u'(C(t))P(t))]. \tag{12.13}$$

By the definition of $\Lambda(t) \equiv e^{-\delta t}u'(C(t))$,

$$0 = \Lambda(t)D(t)dt + E_t[d(\Lambda(t)P(t))],$$

Definition 26. We refer to Λ as the continuous time stochastic discount factor.

Works Cited

Acemoglu, Daron. Introduction to Modern Economic Growth. Princeton: Princeton University Press, 2009.

Apostol, Tom M. Calculus. John Wiley and Sons, 1967.

Bansal, R., Yaron, Amir. *Risks for the Long Run: A Potential Resolution of Asset Pricing Puzzles.* Journal of Finance. 59, August 2004: 1481-1509.

Beyer, Marianne. Portfolio Decisions in Continuous Time. Senior Thesis. April 14, 2008.

O. Calin, Y. Chen, T. Cosimano, and A. Himonas. *Solving asset pricing models when the price-dividend function is analytic. Econometrica* 73 (2005), no. 3, 961–982.

Campbell, J. Y., Cochrane, J., 1999, *By Force of Habit, a Consumption-Based Explanation of Aggregate Stock Market Behavior. Journal of Political Economy.* 107, 205-251.

Chappell, David. *On the Derivation and Solution of the Black-Scholes Option Pricing Model. Spoudai.* Vol. 42, no. 3-4.

Chen, Y., Cosimano, T., Himonas, A.. *Asset Pricing Model with Duffie-Epstien Preferences.* Notes. September 2008.

Chow, Gregory. Dynamic Economics: Optimization by the Lagrange Method. New York: Oxford University Press, 1997.

Cochrane, John H. Asset Pricing Revised Edition. Princeton: Princeton University Press, 2005.

Cvitanic, J., and Zapatero, F.. Introduction to the Economics and Mathematics of Financial Markets. Cambridge: The MIT Press, 2004.

Duffie, D. and Epstein, L.. *Stochastic Differential Utility. Econometrica* 60 (1992), no. 2, 353-394.

Folland, Gerald B. Real Analysis: Modern Techniques and Their Applications. New York: John Wiley and Sons, 1999.

Himonas, A. Analyticity for Second-order Linear Partial Differential Equations. Notes.

October 2006.

Himonas, A. Complex Methods in Partial Differential Equations. Notes. Fall 2008.

Kolmogorov, A.N.. Introductory Real Analysis. New York: Dover Publications, Inc., 1975.

Kreps, D. and Porteus, E.. *Temporal Resolution of Uncertainty and Dynamic Choice Theory. Econometrica* 46 (1978), no.1, 185-200.

Ljungqvist, Lars. Recursive Macroeconomic Theory. Cambridge: The MIT Press., 2004.

Lucas, Robert E. Jr. Recursive Methods in Economic Dynamics. Cambridge: Harvard University Press, 1989.

Rockafellar, Tyrrell R. Convex Analysis. Princeton: Princeton University Press, 1970.

Ross, Sheldon. A First Course in Probability. Upper Saddle River: Prentice Hall, Inc., 1998.

Shreve, Steven E. Stochastic Calculus for Finance II: Continuous Time Models. New York: Springer Science+Business Media, Inc., 2004.

Wachter, J. A., 2006, A Consumption Model of the Term Structure of Interest Rates. *Journal of Financial Economics 79* 365-399.