Suboptimal Choice and Asset Pricing

Jun Liu[1]

December 4, 2018

Abstract

I develop a model of suboptimal choice. In this model, the choice of the agent is random in the sense that any feasible choice is possible and the probability of a choice being selected increases with the utility of the choice. Under this model, the aggregate consumption does not equal to the optimal consumption of the representative agent, thus the Euler equation does not hold. The model helps to resolve the interest rate puzzle and the equity premium puzzle.

[1]Jun Liu is with the Rady School of Management, University of California San Diego. I thank Michael Brennan, Lawrence Christiano, Philip Dybvig, Mark Loewenstein, Jun Pan, Allan Timmermann, Rossen Valkanov, Hong Yan (discussant) and seminar participants at Southwestern University of Finance and Economics, Shanghai Jiao Tong University, UCSD for helpful comments.

1 Introduction

Even though agents in economic model are assumed to make choice to maximize utility, in reality, the choice of most people does not maximize their utility for various reasons. This is probably one of the reasons that consumption data of individual agents are not used to test asset pricing models, even though in theory the marginal utility of every agent in the economy should be the pricing kernel. Most tests of asset pricing models use aggregate consumption data. Underlying this approach is the assumption that errors of individual agents average out in aggregate; thus, even though the consumption of individual agents is not optimal, the aggregate consumption is optimal and satisfies the optimality condition such as the Euler equation. However, most tests of asset pricing model based on Euler's equation have failed, even though aggregate consumption is used. Examples include consumption-based CAPM, equity premium puzzle. In New Keynsian model, economists have to introduce preference shock to account for the failure of Euler equations.

In this paper, I propose a plausible model of suboptimal choice. In this model, any choice in the feasible set can be selected by the agent. The choice with higher utility has a higher probability being selected. The aggregate consumption is not optimal thus the randomness of suboptimality does not "average out." In such a model, interest rate is lower than models with optimal choice, due to the result that suboptimal choice effectively reduces the subjective discount rate, among other effects. The equity premium is higher than optimal choice models, because suboptimal choice effectively reduces the probability of good states and increases the probability of bad states.

2 The Standard Utility Theory

The standard utility theory has two key elements.

1. The space $\mathcal{S}$ of feasible choice C.
2. Agents choose C^* (the optimal choice) from $\mathcal{S}$ to maximizes utility $U(C)$

$$C^* = \text{argumax}_{C \in \mathcal{S}} U(C).$$

In this paper, we will use consumption choice over two dates as illustration. Let us consider consumption C_0 and C_1 over two dates 0 and 1. For now, we assume that there is no uncertainty. Suppose that the agent has initial wealth w_0 and the interest rate is r.

The agent's choice is constrained by the budget constraint

$$C_0 + e^{-r}C_1 = W_0.$$

The choice space is

$$\mathcal{S} = \{C_1 : e^{-r}C_1 \in [0, W_0]\},$$

with

$$C_0 = W_0 - e^{-r}C_1$$

determined by the budget constraint. We assume that the budget constraint always holds.

The objective of the agent is

$$\max_{C_0 \in \mathcal{S}} U(C_0, C_1).$$

For an additive utility for C_0 and C_1,

$$U(C_0, C_1) = u(C_0) + e^{-\beta}u(C_1),$$

with

$$u(C) = \frac{C^{1-\gamma}}{1-\gamma},$$

the objective function becomes

$$U \quad = \quad \frac{(W_0 - e^{-r}C_1)^{1-\gamma} - 1}{1-\gamma} + e^{-\beta}\frac{C_1^{1-\gamma} - 1}{1-\gamma}.$$

The Euler equation

$$C_0^{-\gamma} = e^{-\beta}C_1^{-\gamma}e^{r}$$

leads to

$$C_0 = e^{\frac{\beta-r}{\gamma}}C_1.$$

Substituting the above equation into the budget constraint, we obtain the optimal choice that maximizes utility give by

$$C_1^* = \frac{1}{e^{-r} + e^{\frac{\beta-r}{\gamma}}}W_0.$$

3 A Model of Suboptimal Choice

We make the following assumptions.

1. Any feasible C in $\mathcal{S}$ can be chosen.

2. The probability density function $f(C)$ for C to be chosen is proportional $\exp\left(\frac{U(C)}{h}\right)$, with $h > 0$.

In the standard utility theory, the choice by the agent is unique (if the utility is convex). Under our assumption, all feasible choice can be selected by the agent with some probability.

It is natural to assume that a feasible choice C that yields higher utility $U(C)$ should have a higher probability being selected, thus $f(C)$ should equal to $g(U(C))$, where g is a strictly increasing function. We assume that $g(U) = e^{\frac{U}{h}}$; this is strictly increasing and positive and has one parameter $h > 0$.

Example. In our previous consumption choice example, any $C_0 \in \mathcal{S} \equiv [0, W_0]$ can potentially be selected by the agent. The probability density function for C_0 to be chosen

$$f(C_0) \propto \exp\left(\frac{1}{h}\left(\frac{C_0^{1-\gamma}}{1-\gamma} + e^{-\beta+(1-\gamma)r}\frac{(W_0 - C_0)^{1-\gamma}}{1-\gamma}\right)\right).$$

Figure 1 graphs the probability function $f(C)$ for different value of h. When h is reduced, the probability is more concentrated near the optimal choice C^*.

Figure 2 shows that the mean and variance of the consumption choice and the optimal consumption choice. The mean is lower than the optimal consumption choice. When h is close to zero, the mean is close to the optimal choice. When h is large, the mean is close to $1/2$, which is the result of completely random choice. The variance of the choice increases with h; when h is close to zero, the choice is almost unique and is the optimal choice and the choice is completely random when h is large. Thus, h is a parameter that characterizes the degree of randomness in choice.

3.1 Steepest Decent and the Optimal Choice

We show that the optimal choice is a special case.

Proposition 3.1. When $h \to 0$, $f(C) \propto \delta(C - C^*)$.

This is due to the mathematical theorem of steepest decent.

When $h \to \infty$, $f(C)$ is a constant and all $C \in \mathcal{S}$ are equally likely to be chosen.

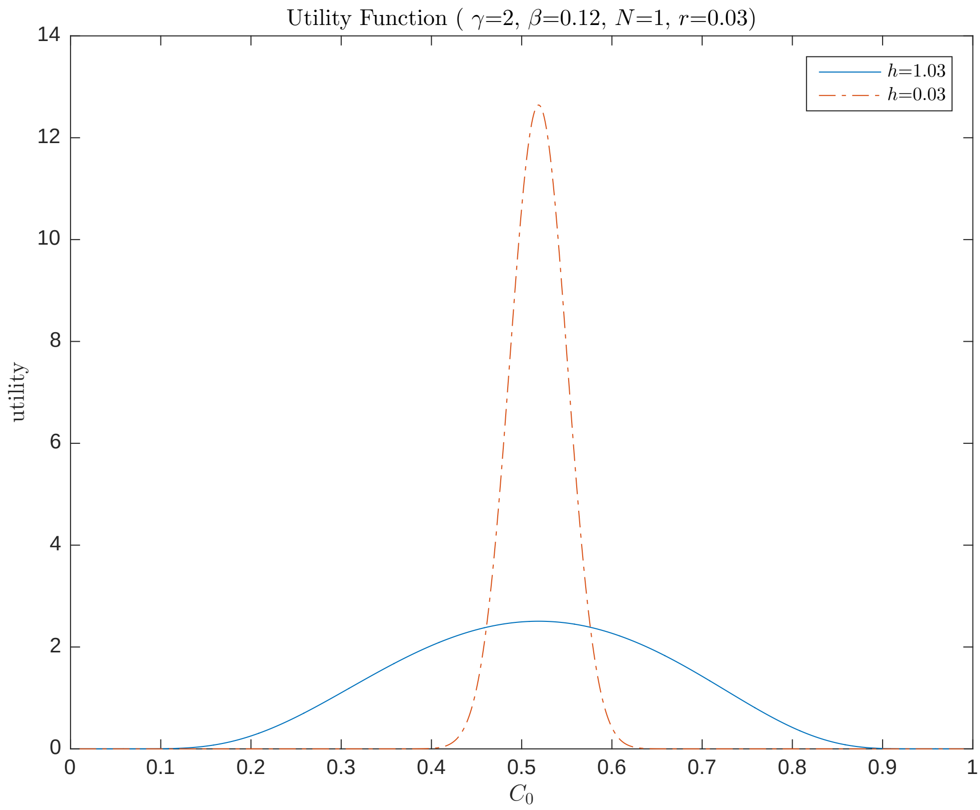

Figure 1: The probability density function for C_0: higher h leads to a more random choice.

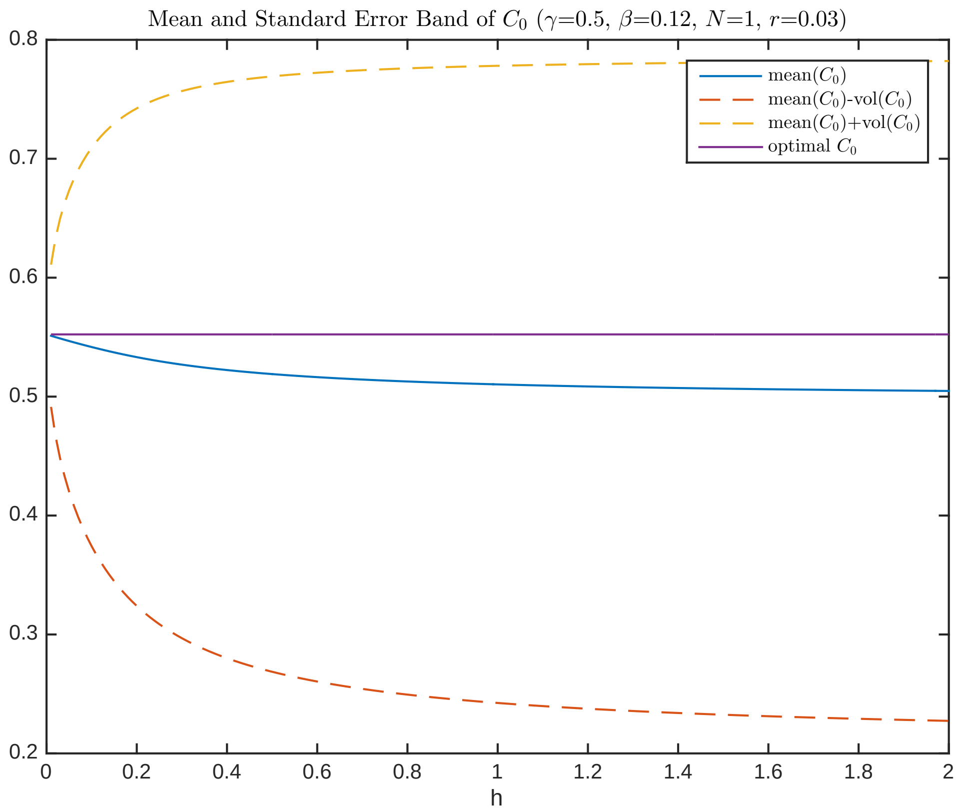

Figure 2: The optimal C_0 is higher than the mean of C_0 and is within the standard error band.

3.2 Log Utility: Probability Density Function

We can compute the probability function $f(C)$ in closed form for the the case of log-utility,

$$U = \ln C_0 + e^{-\beta} \ln C_1.$$

The probability is proportional to ($\alpha = e^{-\beta}$)

$$e^{\frac{1}{h}(\ln C_0 + \alpha \ln C_1)} = C_0^{\frac{1}{h}} C_1^{\frac{\alpha}{h}} = e^{\frac{r\alpha}{h}} (W_0 - e^{-r} C_1)^{\frac{1}{h}} C_1^{\frac{\alpha}{h}}$$

which can be written as

$$e^{\frac{r\alpha}{h}} W_0^{\frac{1+\alpha r}{h}} (1 - \hat{C}_1)^{\frac{1}{h}} \hat{C}_1^{\frac{\alpha}{h}}$$

with $\hat{C}_1 \equiv e^{-r} C_1 / W_0$. Note that $\alpha = e^{-\beta}$ represents the subjective discount rate. The discount rate disappears if $h = \infty$.

We can integrate over $\hat{C}_0$ to determined the proportional constant. The probability density is given by

$$f(\hat{C}_1) = \frac{(1 - \hat{C}_1)^{\frac{1}{h}} \hat{C}_1^{\frac{\alpha}{h}}}{\int_0^1 (1 - \hat{C}_1)^{\frac{1}{h}} \hat{C}_1^{\frac{\alpha}{h}} d\hat{C}_1} = \frac{(1 - \hat{C}_1)^{\frac{1}{h}} \hat{C}_1^{\frac{\alpha}{h}}}{B(\frac{1}{h} + 1, \frac{\alpha}{h} + 1)}$$

where $B(x, y)$ is the beta function.

We can compute the mean consumption in closed form for the case of log utility. The mean of $\hat{C}_1$ is given by

$$\mathrm{E}[\hat{C}_1] = \int_0^1 f(\hat{C}_1) \hat{C}_1 d\hat{C}_1 = \frac{B(\frac{1}{h} + 1, \frac{\alpha}{h} + 2)}{B(\frac{1}{h} + 1, \frac{\alpha}{h} + 1)}$$

Using the recursive relation of Beta function, we get

$$\mathrm{E}[\hat{C}_1] = \frac{\frac{\Gamma(\frac{1}{h}+1)\Gamma(\frac{\alpha}{h}+2)}{\Gamma(\frac{1}{h}+2+\frac{\alpha}{h}+1)}}{\frac{\Gamma(\frac{1}{h}+1)\Gamma(\frac{\alpha}{h}+1)}{\Gamma(\frac{1}{h}+1+\frac{\alpha}{h}+1)}} = \frac{\frac{\alpha}{h} + 1}{\frac{1}{h} + 1 + \frac{\alpha}{h} + 1} = \frac{\alpha + h}{1 + \alpha + 2h}.$$

When $h = 0$, we recover the optimal C_1^*

$$\hat{\mathrm{E}}[\hat{C}_1] = \frac{\alpha}{1 + \alpha}.$$

When $h = \infty$, we get

$$\mathrm{E}[\hat{C}_1] = \frac{1}{2}.$$

As long as $h > 0$, $\mathrm{E}[\hat{C}_1] > C_1^*$ thus $\mathrm{E}[\hat{C}_0] < C_0^*$. Note that wealth effect and intertemporal substitution effect cancels thus there is no r dependence. Early consumption is preferred if $\beta > 0$, $h > 0$ reduces such preference on average.

We can also compute in closed form the variance of consumption. The second moment of $\hat{C}_1$ is given by

$$\mathrm{E}[\hat{C}_1^2] = \int_0^1 f(\hat{C}_1)\hat{C}_1^2 d\hat{C}_1 = \frac{B(\frac{1}{h}+1, \frac{\alpha}{h}+3)}{B(\frac{1}{h}+1, \frac{\alpha}{h}+1)}$$

Using the recursive relation of beta function, we get

$$\mathrm{E}[\hat{C}_1^2] = \frac{(\frac{\alpha}{h}+2)(\frac{\alpha}{h}+1)}{(\frac{1}{h}+2+\frac{\alpha}{h}+1)(\frac{1}{h}+1+\frac{\alpha}{h}+1)}.$$

The variance of $\hat{C}_1$ is

$$\left(\frac{(\frac{\alpha}{h}+2)}{(\frac{\alpha}{h}+2+\frac{\alpha}{h}+1)} - \frac{(\frac{\alpha}{h}+1)}{(\frac{1}{h}+1+\frac{\alpha}{h}+1)}\right)\frac{(\frac{\alpha}{h}+1)}{(\frac{1}{h}+1+\frac{\alpha}{h}+1)}$$

which equals to

$$\frac{(\frac{\alpha}{h}+1)(\frac{1}{h}+1)}{(\frac{1}{h}+2+\frac{\alpha}{h}+1)(\frac{1}{h}+1+\frac{\alpha}{h}+1)^2}.$$

The variance goes to 0 as $h \to 0$ and 1/2 as $h \to \infty$. As long as $h > 0$, the variance of $\hat{C}_1$ is greater than 0.

3.3 Equilibrium

To study equilibrium implications, I make the following assumptions.

1. Economy has N agents.
2. All agents have the same γ and β and h and initial wealth W_0 (all agents are identical "ex ante.")
3. Different agents represent different independent draws from the distribution.

We assume that aggregate per-capita endowment at time t is D_t, $t = 0, 1$. Market clearing requires that

$$\bar{C}_t \equiv \frac{1}{N}\sum_{n=1}^{N} C_{tn} = D_t.$$

From our assumption that different agents represent different independent draws from the distribution,

$$\bar{C}_t = \mathrm{E}[C_t] = W_0 \mathrm{E}[\hat{C}_t],$$

as long as N is large enough.

The above two equations yield

$$(D_0 + e^{-r} D_1)\frac{1+h}{1+\alpha+2h} = D_0.$$

Thus, the equilibrium interest rate is given by

$$e^r = \frac{D_1}{D_0}\frac{1+h}{e^{-\beta}+h}.$$

The equilibrium interest rate r has following properties.

1. When $h = 0$, it reduces to that of the standard model, $r = \beta + \ln(D_1/D_0)$.
2. When $h = \infty$, $r = \ln(D_1/D_0)$.
3. The equilibrium interest rate is lower than the interest rate in the standard model. Thus the suboptimality helps resolve interest rate puzzle.
4. When $\gamma \neq 1$, the decrease in interest rate depends also on the inverse of the elasticity of intertemporal substitution γ.

4 Risky Case with Two States

We assume that there are two dates, $t = 0$ and $t = 1$ and there are two states at $t = 1$, u and d. There is a riskless asset and a stock that is the claim to the aggregate dividend, thus the markets are complete.

We assume that the utility of the agents are the additive expected utility given by

$$U = \ln C_0 + e^{-\beta}(\pi_u \ln C_u + \pi_d \ln C_d),$$

where π_u and π_d are probability of u and d state respectively.

Note that C_0 is preferred over C_u and C_d if $\beta > 0$ and C_u is preferred over C_d if $\pi_u > \pi_d$. Under the standard framework, there are optimal and unique solution C_0^*, C_u^*, and C_d^*.

In the random choice framework, any feasible choice that satisfies the budget constraint are possible. Because markets are complete, the budget constraint is given by

$$C_0 + p_u C_u + p_d C_d = W_0,$$

where p_u and p_d are state prices for u and d state respectively. Given C_u and C_d, C_0 is determined by the constraint. So the feasible space $\mathcal{S}$ is given by

$$\mathcal{S} = \{(C_u, C_d) : p_u C_u + p_d C_d \leq W_0, C_u \geq 0; C_d \geq 0\}.$$

Given the above expected utility, the probability for (C_u, C_d) is proportional to

$$e^{U/h} = C_0^{\frac{1}{h}} C_u^{\frac{\alpha \pi_u}{h}} C_d^{\frac{\alpha_d \pi_d}{h}} = (W_0 - p_u C_u - p_d C_d)^{\frac{1}{h}} C_u^{\frac{\alpha \pi_u}{h}} C_d^{\frac{\alpha_d \pi_d}{h}}$$

which can be written as

$$W_0^{\frac{1+\alpha\pi_u+\alpha\pi_d}{h}} p_u^{\frac{\alpha\pi_u}{h}} p_d^{\frac{\alpha\pi_d}{h}} (1 - \hat{C}_u - \hat{C}_d)^{\frac{1}{h}} \hat{C}_u^{\frac{\alpha\pi_u}{h}} \hat{C}_d^{\frac{\alpha_d \pi_d}{h}}$$

where $\hat{C}_u \equiv \frac{p_u C_u}{W_0}$ and $\hat{C}_d \equiv \frac{p_d C_d}{W_0}$.

Note that $\alpha\pi_u$ and $\alpha\pi_d$ represent utility differences for C_u and C_d.

We can compute the proportional constant in closed form,

$$\int_{\{\hat{C}_u + \hat{C}_d \leq 1\}} (1 - \hat{C}_u - \hat{C}_d)^{\frac{1}{h}} \hat{C}_u^{\frac{\alpha\pi_u}{h}} \hat{C}_d^{\frac{\alpha\pi_d}{h}} d\hat{C}_u d\hat{C}_d = B(1 + \frac{1}{h}, 1 + \frac{\alpha\pi_u}{h} 1 + \frac{\alpha\pi_d}{h}).$$

where $B(x_1, x_2, x_3)$ is the Dirichlet function. Thus the probability density

$$\frac{(1 - \hat{C}_u - \hat{C}_d)^{\frac{1}{h}} \hat{C}_u^{\frac{\alpha\pi_u}{h}} \hat{C}_d^{\frac{\alpha_d \pi_d}{h}}}{\int_{\{\hat{C}_u + \hat{C}_d \leq 1\}} (1 - \hat{C}_u - \hat{C}_d)^{\frac{1}{h}} \hat{C}_u^{\frac{\alpha\pi_u}{h}} \hat{C}_d^{\frac{\alpha_d \pi_d}{h}} d\hat{C}_u d\hat{C}_d} = \frac{(1 - \hat{C}_u - \hat{C}_d)^{\frac{1}{h}} \hat{C}_u^{\frac{\alpha\pi_u}{h}} \hat{C}_d^{\frac{\alpha_d \pi_d}{h}}}{B(1 + \frac{1}{h}, 1 + \frac{\alpha\pi_u}{h} 1 + \frac{\alpha\pi_d}{h})}.$$

As for the case without risk, we can compute the equilibrium. The market clearing condition is

$$\mathrm{E}[C_u] = D_u$$

and

$$\mathrm{E}[C_d] = D_d$$

which can be written as

$$\mathrm{E}[\hat{C}_u] = \hat{D}_u$$

and

$$\mathrm{E}[\hat{C}_d] = \hat{D}_d$$

where $\hat{D}_u = \frac{p_u D_u}{W_0}$ and $\hat{D}_d = \frac{p_d D_d}{W_0}$.

We can compute the expectation

$$\mathrm{E}[\hat{C}_u] = \frac{1 + \frac{\pi_u \alpha}{h}}{3 + \frac{1+\alpha(\pi_u+\pi_d)}{h}} = \frac{h + \alpha\pi_u}{3h + 1 + \alpha}.$$

So the market clearing leads to

$$\frac{h+\alpha\pi_u}{3h+1+\alpha} = \frac{p_u D_u}{W_0}.$$

Note that $W_0 = D_0 + p_u D_u + p_d D_d$, we get

$$(h+\alpha\pi_u)(D_0 + p_u D_u + p_d D_d) = (3h+1+\alpha) p_u D_u.$$

Similarly,

$$(h+\alpha\pi_d)(D_0 + p_u D_u + p_d D_d) = (3h+1+\alpha) p_d D_d.$$

Adding the two equations, we get

$$(2h+\alpha)(D_0 + p_u D_u + p_d D_d) = (3h+1+\alpha)(p_u D_u + p_d D_d).$$

We have

$$p_u D_u + p_d D_d = \frac{(2h+\alpha) D_0}{1+h}.$$

From the above equation, we can solve the state prices

$$p_u = \frac{h+\alpha\pi_u}{1+h}\frac{D_0}{D_u} = \frac{1+\frac{h}{\alpha\pi_u}}{1+h}\alpha\pi_u\frac{D_0}{D_u}$$

and

$$p_d = \frac{h+\alpha\pi_d}{1+h}\frac{D_0}{D_d} = \frac{1+\frac{h}{\alpha\pi_d}}{1+h}\alpha\pi_d\frac{D_0}{D_d}.$$

As a comparison, in the standard model,

$$p_u = \alpha\pi_u\frac{D_0}{D_u}$$

and

$$p_d = \alpha\pi_d\frac{D_0}{D_d}.$$

The pricing kernel is given by

$$M_u = \frac{p_u}{\pi_u} = \frac{h+\alpha\pi_u}{(1+h)\pi_u}\frac{D_0}{D_u} = \frac{1+\frac{h}{\alpha\pi_u}}{1+h}\alpha\frac{D_0}{D_u}$$

and

$$M_d = \frac{p_d}{\pi_d} = \frac{h+\alpha\pi_d}{(1+h)\pi_d}\frac{D_0}{D_d} = \frac{1+\frac{h}{\alpha\pi_d}}{1+h}\alpha\frac{D_0}{D_d}.$$

In the standard model,

$$M_u = \alpha\frac{D_0}{D_u}$$

and

$$M_d = \alpha \frac{D_0}{D_d}.$$

The u state weight factor is $1 + \frac{h}{\alpha \pi_u}$ which is smaller than the d weight factor $1 + \frac{h}{\alpha \pi_d}$ if $\pi_u > \pi_d$. So d state has even more weight relative to the standard model.

The riskless rate is

$$R_f^{-1}(h) = \frac{h + \alpha \pi_u}{1+h} \frac{D_0}{D_u} + \frac{h + \alpha \pi_d}{1+h} \frac{D_0}{D_d}.$$

When $h = 0$, we recover the riskless rate for the standard model

$$R_f^{-1}(0) = \alpha \pi_u \frac{D_0}{D_u} + \alpha \pi_d \frac{D_0}{D_d}.$$

When $h = \infty$, we have

$$R_f^{-1}(\infty) = \frac{D_0}{D_u} + \frac{D_0}{D_d}.$$

Note that there is "double counting" because probability weights have disappeared. Note that risk decreases R_f and randomness of choice also reduces the riskless rate, $R_f(\infty) < R_f(0)$.

The price of the stock, which is the claim to the aggregate dividend, is

$$S = p_u D_u + p_d D_d = \frac{2h + \alpha}{1+h} D_0.$$

The expected return of the stock is

$$\mu = \frac{\pi_u D_u + \pi_d D_d}{\frac{(2h+\alpha)D_0}{1+h}} = \frac{\pi_u D_u + \pi_d D_d}{\alpha D_0} \frac{1+h}{1+2h/\alpha} < \frac{\pi_u D_u + \pi_d D_d}{\alpha D_0}.$$

The risk premium is

$$\frac{1+h}{D_0} \left(\frac{\pi_u D_u + \pi_d D_d}{2h + \alpha} - \left(\frac{h + \alpha \pi_u}{D_u} + \frac{h + \alpha \pi_d}{D_d} \right)^{-1} \right).$$

This can be re-written as

$$\frac{1}{R_f} \frac{\pi_u \pi_d (D_u - D_d)^2}{D_u D_d} \frac{\alpha}{2h + \alpha} \left(1 + \frac{(\pi_u D_u - \pi_d D_d) h}{\alpha \pi_u \pi_d (D_u - D_d)} \right).$$

where

$$\frac{1}{R_f} \frac{\pi_u \pi_d (D_u - D_d)^2}{D_u D_d}$$

is the equity premium in the standard CCAPM. Assuming that $2\pi_d < 1$ and $D_d < D_u$, the factor

$$\frac{\alpha}{2h+\alpha}\left(1+\frac{(\pi_u D_u - \pi_d D_d)h}{\alpha\pi_u\pi_d(D_u - D_d)}\right)$$

is greater than 1, thus the equity premium is higher than the standard case. So the suboptimality may help resolve equity premium puzzle.

The equilibrium for general γ cannot be solved in closed form but can be numerically evaluated. We can also extend the model to than 1 period and more than 2 risky states.

5 Conclusions

In this paper, I build a plausible model of suboptimal choice. In this setup, the choice is random in the sense that any feasible choice can be selected by the agent. We show that the randomness in suboptimality does not "average out" in the sense that the average consumption does not equal to the optimal consumption, thus the Euler equation is violated. This may provide an explanation for failure of many asset price model tests based on Euler equation. The model also generate lower interest rate and higher risk premium. The expected return is lower than the model without randomness due to lower interest rate.

Appendix: Beta Function

The definition of Beta function is

$$B(x,y) = \int_0^1 C^{x-1}(1-C)^{y-1}dC.$$

One property of Beta function is

$$B(x,y) = \frac{\Gamma(x)\Gamma(y)}{\Gamma(x+y)}.$$

where Γ function is defined by

$$\Gamma(x) = \int_0^\infty C^{x-1}e^{-C}dC.$$

The Γ function has the following property

$$\Gamma(x+1) = x\Gamma(x).$$