Explorations in the Equity-Premium Puzzle

James Myatt *

May 7, 2012

An honors thesis submitted in partial fulfillment

of the requirements for the degree of

Bachelor of Science

Undergraduate College

Leonard N. Stern School of Business

New York University

May 2012

Faculty Advisor: Marti G. Subrahmanyam

Thesis Advisor: David Backus

*I owe many people thanks for helping me to complete this thesis. First, I owe Professors Dave Backus and Tom Sargent for helping me extend our work in the classroom to further research. They have been enormous influences on my interests in research and decision to attend graduate school. I also want to express my love and gratitude for my parents. I will never be able to completely repay them for giving me the opportunity to attend NYU.

Abstract

Historically, the average return on equity has greatly exceeded the average return on short-term, conventionally "risk-free" securities by six percent. Mehra and Prescott (1985) noted that generally accepted and widely used economic models could not produce the same empirical result, and gave this problem the name, "the equity premium puzzle." This paper explains the equity premium puzzle in further detail and highlights some of the progress made my Hansen and Jagannathan (1991) as well as Tallarini (2000) to solve it.

1 Intoduction

The equity premium puzzle is an ongoing unsolved problem in economics. The main idea of the puzzle is that commonly used economic models fail to predict the empirical premium of stock market returns over short-term, relatively risk-free debt. The average real annual yield on the Standard & Poors 500 Index from 1889-2009 was close to eight percent, while the average annual yield on short-term, relatively risk-free debt, was less than two percent. With aggregate real consumption data and a reasonable value governing risk aversion, the theory predicts an equity premium less than one percent. This paper will highlight some of the attempts to describe and remedy this discrepancy.

This problem was dubbed the equity premium puzzle by Mehra and Prescott (1985), and their paper has driven a wealth of research to solve it. The most obvious solution is to adjust the parameter governing risk aversion, the coefficient of relative risk aversion (CRRA). Mehra and Prescott noted that the only way to drive the model's predicted equity premium to the empirical equity premium is to raise the CRRA to implausible levels. For them, the CRRA should be positive, and most likely near one, but their model requires a CRRA around 30. There is some debate over what is a reasonable value for the CRRA, and Kocherlakota (1996) reviews some of the literature. On one side, Mankiw and Zeldes (1991) argue that "an individual with a coefficient of relative aversion above ten would be willing to pay unrealistically large amounts to avoid bets." While Kandel and Stambaugh (1991) show that "even values of (the CRRA) as high as thirty imply quite reasonable behavior when the bets involves a maximal potential loss of around 1% of the gambler's wealth. Kocherlakata finds the consensus to be a CRRA less than ten.

Another argument has been made about the data. The United States enjoyed a rise to becoming a world super power during the time span of the data used, 1889-2009. While there were major recessions and a depression during this time period, the United States has grown enormously in terms of output per capita. Some economists asked what is the equity premium in other countries or at other times? Dimson, Marsh, and Staunton (2006) used data from seventeen countries over a 106 year interval and find an equity premium over U.S. Treasury Bills of about four and a half percent. Mehra and Prescott (2003) also investigated the international equity premium and found similar results.

Hansen and Jagannathan (1991) made a large step forward in research by approaching the equity premium puzzle from the data. Modern asset pricing requires a stochastic discount factor which is related to an agents marginal rate of substitution between future consumption goods and current consumption. Instead of modifying the agents stochastic discount factor to fit the returns data, they took the returns data to put a bound on the agent's stochastic discount factor. This did not solve the puzzle, but it shined a light on where the answer would be. And the answer, at that time, was very off.

Since the advent of the Hansen-Jagannathan bound, there has been significant research to push models into the admissible region, which is within the bounds. Epstein and Zin (1989) and Tallarini (2000) have made significant progress by parsing out the effects of risk aversion, which comes from the CRRA, from the effects of an agent's intertemporal elasticity of substitution (IES). While the preferences used by these researchers show the model moving in the right direction, towards the Hansen-Jagannathan bounds, the results still require a large CRRA.

Current research by Barillas, Hansen, and Sargent (2009) focuses on model uncertainty where the agent is unsure of the probability distributions and the world she faces. This adds another dimension of the problem where the agent understands measurable risk and her IES, but she does not know what risks apply to her situation. This is like a card counter who understands the likelihood of cards dealt to him if he is playing with a standard deck, but he does not know if the deck is actually standard. Interestingly, this approach could answer the suggestion made by Mehra and Prescott in their original paper where they supposed that some contracts in a market might not be enforceable, adding uncertainty and increasing the required premium. This research has broader implications than finding a model to fit the data. Improved models shed light on the macroeconomic costs of unexpected changes in consumptions growth.

This paper will replicate some of the results from the above summary. First, we will derive the stochastic discount factor which will be used throughout the paper. Section 3 will discuss the commonly used power utility function with the CRRA and discuss reasonable values for the CRRA using a thought experiment. We will replicate Mehra and Prescott's result using a simple model in Section 4. Then we will derive and explain the Hansen Jagannathan bounds in Section 5. Finally we will show Tallarini's use of Epstein-Zin preferences to push our model to the admissible region in the Hansen-Jagannathan bounds.

2 Deriving the Stochastic Discount Factor

Throughout this paper, we will use the Euler equation $E_t(m_{t+1}R_{t+1}) = 1$ to price assets where $R_{t+1} = 1 + r_{t+1}$ will be the gross rate of return on an asset and m_{t+1} is called the stochastic discount factor or sometimes the pricing kernel. Since we know information about expected interest rates from data, the key to asset pricing is determining m_{t+1}, and we will see later in Section 4 that solving the equity premium puzzle is a matter of generating sufficient dispersion in the stocastic discount factor.

The model consists of a representative agent living at time t and considers a consumption-savings decision for next period, $t+1$. The agent is endowed with $y_t > 0$ and she can use her endowment for some combination of immediate consumption or purchasing a quanitity b of a "risky" asset for price p that pays some dividend $x(\omega)$ in $t+1$. The payoff depends on the state $\omega \in \Omega$, realized in $t+1$, and ω is distributed by the probability density function $f(\omega)$. Once the state is realized, the agent consumes the entire payoff. The agent's utility of consumption is described by $u(c)$ where the only requirements are that $u' > 0$, $u'' < 0$ and $\lim_{c\to 0} u(c) = \infty$. Allowing $\beta \in (0,1)$ to be the agent's discount factor, she wants to choose c_t, $c_{t+1}(\omega)$ and b to maximize the function

$$u(c_t) + \beta \int_{\omega\in\Omega} u(c_{t+1}(\omega)) f(\omega) d\omega$$

such that

$$c_t + pb \leq y_t$$

$$c_{t+1}(\omega) \leq x(\omega)b$$

This is essentially a two period version of the model described in Lucas (1978). The agent makes consumption-savings tradeoffs to maximize discounted, expected utility. To solve the problem we employ methods from Lagrange.

$$L = u(c_t) + \beta \int_{\omega\in\Omega} u(c_{t+1}(\omega)) f(\omega) d\omega + \theta(y_t - c_t - pb) + \int_{\omega\in\Omega} \lambda(\omega)(x(\omega)b - c_{t+1}(\omega)) d\omega$$

where θ and $\lambda(\omega)$ are Lagrange multipliers.The first order necessary conditions are:

$$u'(c_t) - \theta = 0$$

$$\beta u'(c_{t+1}(\omega))f(\omega) - \lambda(\omega) = 0$$

$$-p\theta + \int_{\omega\in\Omega} \lambda(\omega)x(\omega)d\omega = 0$$

with respect to c_t, $c_{t+1}(\omega)$, b.

These equations yield the following

$$p = \int_{\omega\in\Omega} \beta\frac{u'(c_{t+1}(\omega))}{u'(c_t)}x(\omega)f(\omega)d\omega$$

We let $m_{t+1} = \beta\frac{u'(c_{t+1}(\omega))}{u'(c_t)}$ and we call m_{t+1} the stochastic discount factor. This is the discounted marginal rate of substitution between current and future consumption. Notice that we are integrating over a probability distribution so that $p = E(mx)$. And we know that $R_{t+1} = x/p$, so we can divide both sides by p to get the well known asset pricing Euler equation $E(m_{t+1}R_{t+1})$ The subscripts of time are often dropped in this paper because we are interested in the average annual, one period, payoff only.

3 Power Utility and the CRRA

Before explaining the equity premium puzzle we pause to discuss the power utility function, defined below, and reasonable values for the coefficient of relative risk aversion. In the model, the agent only has the choice of a representative composite consumption good. Economists widely use the power utility function when a concrete functional form is needed. The power utility function is an increasing, concave function which satisfies the basic axioms of utility theory. It is practical because it is familiar, easy to calculate and differentiate, so it has become an industry standard for a good "first guess" or starting point.

$$u(c) = \begin{cases} \frac{c^{1-\alpha}}{1-\alpha} & : 0 < \alpha \neq 1 \\ \log(\alpha) & : \alpha = 1 \end{cases}$$

The CRRA is α in the above equation. It governs the curvature or the utility function, and higher values of α are associated with higher levels of risk aversion. The important part of the equity-premium puzzle and asset pricing, in general, is seeking a unique and constant α that reflects the longstanding, innate views of the market. The existence and uniqueness of such a parameter is debatable, but it has a nice story of fitting the conventional wisdom delivered by most financial advisors and certain "expert" friends and relatives: "Stocks are always good in the long run."

We summarize some of the research done by economists and psycologists over the years with a simple thought experiment proposed by Mankiw (1991). Before describing the experiment, we define the certainty equivalent as $\mu = u^{-1}(E(u(c)))$ which is the amount of consumption that generates the same utility in each state. For the power utility function, the certainty equivalent is

$$\mu = [E(c^{1-\alpha})]^{\frac{1}{1-\alpha}}$$

The thought experiment is making a choice between two incomes next year. The first is a certain income of \$100,000 and the second is a "risky income" where there is a 50% chance that you will receive \$100,000 and there is a 50% percent chance that you will receive \$200,000. Surely the risky income is the better bet. A worst case scenario is that you end up with the income of \$100,000 you would have gotten anyway. Now, if the certain income is raised above $100,000, then when will you make the switch to the certain income? The income where you make the switch is the certainty equivalent, and we can use it to solve for α. Below is a table of certainty incomes and the CRRA associated with it. Clearly, people who are still willing to risk a potentially higher income will have a higher certainty equivalent, and the table reflects this as higher certainty equivalents are associated with less risk aversion. The levels of risk aversion found in the equity premium puzzle are somewhere in the range of $\mu \in [100, 110]$. Our table predicts that most people will fall somewhere near a "reasonable' α around 10 or less.

Table 1:

$\mu =$ certainty equivalent	$\alpha = CRRA$
100	≈ 70.660
110	8.2015
120	4.2556
130	2.406
140	1.1683
150	≈ 0

4 The Equity Premium Puzzle

Now that we have reviewed the stochastic discount factor and reasonable parameters for the power utility function, we can explain Mehra and Prescott's original result. They used the same Euler equation derived above, power utility, and data about aggregate personal consumption, and they found that the model predicted too small of a risk premium. The data they used is found in Table 2, but it has been updated with recent more returns. However, more observations does not significantly change their original result.

Table 2: United States 1889-2009

	Mean	Standard Deviation	Variance Covariance c_{t+1}/c_t	$1+r^e$	$1+r^1$
c_{t+1}/c_t	1.0207	0.0356	0.001266	0.003595	-0.000126
$1+r^e$	1.0769	0.1846		0.034083	0.001177
$1+r^1$	1.0198	0.0580			0.003362
r^e-r^1	0.0571	0.1873			
			$\log(c_{t+1}/c_t)$	$\log(1+r^e)$	$\log(1+r^1)$
$\log(c_{t+1}/c_t)$	0.0198	0.0350	0.001228	0.003463	-0.000101
$\log(1+r^e)$	0.0587	0.1795		0.032221	0.001289
$\log(1+r^1)$	0.0180	0.0566			0.003206
$\log(1+r^e)-\log(1+r^1)$	0.0407	0.1811			

We will allow dividends of the risk-free bond to be unity, $x^1(\omega) = 1$ for all ω, and therefore $p^1 < 1$ for discounting. Since the agent must consume all returns from investing at $t+1$ we allow the dividend from equity to be exactly consumption growth : $x^e_{t+1} = \frac{c_{t+1}(\omega)}{c_t}$ and we define $g_{t+1} \equiv \frac{c_{t+1}(\omega)}{c_t}$. We will also say that consumption growth is distributed log-normally: $\log g_{t+1} \sim N(\kappa_1, \kappa_2)$. Table 2 provides us with estimates for the arguments: $\kappa_1 = 0.0198$ and $\kappa_2 = (0.0350)^2$. Since the agent has power utility we can rewrite the stochastic discount factor with a concrete utility function

$$m = \beta \frac{u'(c_{t+1}(\omega))}{u'(c_t)} = \beta \left(\frac{c_{t+1}(\omega)}{c_t} \right)^{-\alpha}$$

and the Euler equation, $E(mr) = 1$, can be written for each asset i.

$$1 = \beta E_t \left((1 + r^i_{t+1}) \left(\frac{c_{t+1}(\omega)}{c_t} \right)^{-\alpha} \right)$$

Now $(x^1_{t+1}(\omega)) = 1$ so can derive the price of a risk free asset.

$$\begin{aligned} p^1_t =& E(m_{t+1} 1) \\ =& E(\exp(\log m_{t+1})) \\ =& E(\exp(\log \beta - \alpha \log g_{t+1})) \\ =& \exp(\log \beta - \alpha \kappa_1 + \alpha^2 \kappa_2 / 2) \end{aligned}$$

where the last step follows from the moment generating function of a log-normal random variable. Finally we can find the required return of the risk free asset by noting that $1 + r^1_{t+1} = \frac{1}{p^1_t}$

$$\log(1 + r^1_{t+1}) = -\log \beta + \alpha \kappa_1 - \alpha^2 \kappa_2 / 2 \tag{1}$$

The price for equity, p^e_{t+1}, can be calculated in the same fashion, but the dividend is uncertain and related to consumption growth.

$$
\begin{aligned}
p_t^e =& E_t(m_{t+1} g_{t+1}(\omega)) \\
=& E_t(\beta g_{t+1}(\omega)^{1-\alpha}) \\
=& E_t(\exp(\log\beta + (1-\alpha)\log g_{t+1}(\omega)) \\
=& \exp(\log\beta + (1-\alpha)\kappa_1 + (1-\alpha)^2\kappa_2/2
\end{aligned}
$$

The last line follows from the moment generating function of the log normal random variable $g_{t+1}(\omega)$, again. And we can also calculate our agent's required return on equity by $E(\log(1+r^e_{t+1}) = E(\log g_{t+1}(\omega)) - \log(p^e_t)$

$$
\begin{aligned}
E(\log(1+r^e_{t+1}(\omega)) =& E_t(\log(g_{t+1}(\omega))) - \log p^e_t \\
=& -\log\beta + \alpha\kappa_1 - (1-\alpha)^2\kappa_2/2
\end{aligned}
$$

Now the equity premium is calculated by taking the difference of the returns. Here we note that $\log(1+r) \approx r$.

$$
E_t(\log(1+r^e_{t+1}) - \log(1+r^1_{t+1}) = (2\alpha - 1)\kappa_2/2
$$

If we plug in for $\kappa_2 = (.0350)^2$ and "guess" that α is 10, which is a generous guess according to Mankiw and Kocherlakota, then we only a acheive an equity premium of 1.16%. A benefit of the above result is that the equity premium is increasing in α, so that we may adjust the CRRA to fit the observed equity premium. Here, the observed "log" equity premium is 0.0407. When we fix this and solve for the CRRA, we get $\alpha = 33.7$. This is a reasonable result for some economists, but it relates to Mankiw's experiment from section 3 as a certainty equivalent of $102.14, which is extremely risk averse.

Whether or not the CRRA implied above is plausible, the result has other problems as well. Notice that the expected value of log returns for both assets is quadratic in α. This means that we can raise α to allow for a large equity premium, but it comes at the cost of decreasing returns. At some point, we will raise α and the return on an asset will be *decreasing*, an unrealistic consequence of the model. This is shown in Figure 1. Clearly the role of α is constrained in this context.

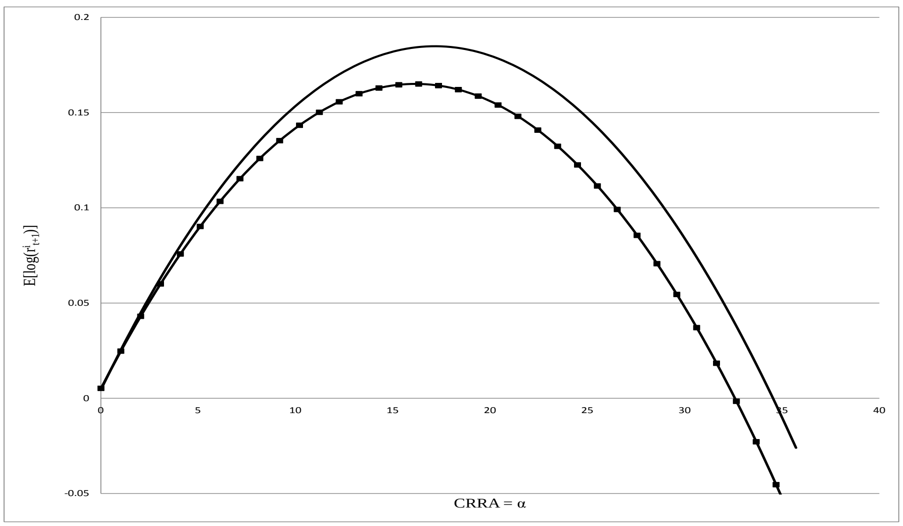

Figure 1: Expected log return on an asset with log normal consumtion growth and power utility

The solid line is the required return of a risky asset. The solid line with squares is the required return of a riskfree asset.

5 Hansen-Jagannathan Bounds

The previous section showed how power utility failed to explain the risk premium unless we used a “high” α = 33.7. The question is why did the CRRA behave so poorly? Power utility is imposed on the representative agent and comes into the Euler equation, $E(mr) = 1$, through the stochastic discount factor. If we cannot explain why power utility makes a poor stochastic discount factor, can we explain more about what a stocastic discount factor should be?

Hansen and Jagganthan (1991) addressed this question. Instead of focusing on measuring expected consumption growth or utility of consumption, they allowed the data to speak for itself. Since agents are pricing assets in the market everyday, they took market returns data and used it to put a boundary on m. While they do not solve the equity premium puzzle by identifying the correct m they find that a plausible stocastic discount factor will need sufficient dispersion, or a higher standard deviation than we can get from the power utility model.

The derivation comes from econometrics and we start with the data we observe. For I assets in a market, we can observe their prices, $p = (p_1, ...p_i, ..., p_I)$ and dividends for any state, $x = (x_1(\omega), ...x_i(\omega), ..., x_I(\omega))$. We once again use the Euler equation to describe the relationship between asset prices, dividends and the stochastic discount factor.

$$p = E(mx) = \int_{\omega} m(\omega)x(\omega)f(\omega)d\omega$$

Where p is an Ix1 vector, m is a scalar, and x is Ix1. We use the least squares method to regress m on x. While we do not have data on m, we can use the Gauss-Markov assumptions in least squares to massage the regression model and yield useful information on m.

$$m = a + x'b + e \tag{2}$$

$$E(ex) = 0 \tag{3}$$

$$E(e) = 0 \tag{4}$$

The least squares model is equation (2), equation (3) is the Gauss-Markov assumption that the unobserved effects on m are uncorrelated with x, and equation (3) says that the average effect of the unobserved terms on m are zero. We now take the expectation of equation (2) and subtract it from itself. Then we multiply the result by $(x - E(x))$.

$$(x - E(x))(m - E(m)) = (x - E(x))(x' - E(x'))b' + (x' - E(x'))e$$

When we take the expectation of both sides, we notice that the left hand side of the result is the definition of the covariance between x and m and the right hand side is the variance-covariance matrix between all our secutities and multiplied by our regression coeffiecient b'. The second term on the right hand side has zero expectation when we use our Gauss-Markov assumptions. Therefore, we are a step closer to solving for b our estimator and for a the constant.

$$\begin{aligned} E(x - E(x))(m - E(m)) &= E(x - E(x))(x' - E(x'))b' \\ \text{Cov}(x, m) &= \text{Cov}(x, x)b' \\ b &= [\text{Cov}(x, x)]^{-1}\text{Cov}(x, m) \\ a &= E(m) - E(x'b) \end{aligned}$$

Again, we have a problem because we have no data on m. However, we can use the modified Euler equation, $p = E(mx)$, to make a subsitution for the Cov(x, m) term

$$\begin{aligned}\text{Cov}(x, m) =& E(mx) - E(m)E(x)\\ =& p - E(m)E(x)\\ b =& [\text{Cov}(x, x)]^{-1}[p - E(m)E(x)]\end{aligned}$$

If we were to make a guess about $E(m)$ the above equation would give us the estimator b to get a value for m. Instead, Hansen and Jagannathan used the least squares model to get an estimate on the standard devation of m.

$$\begin{aligned}m &= a + x'b + e\\ \text{Var}(m) &= \text{Var}(x'b) + \text{Var}(e)\\ \Rightarrow \text{Var}(m) &\geq \text{Var}(x'b)\\ \sigma(m) &\geq (\text{Var}(x'b))^{1/2}\end{aligned}$$

This last equation puts a lower bound on the standard deviation of any stochastic discount factor. Now we amend these formulas to cater to our interest in asset returns. We note that the gross return of an asset is $R = x/p$, and the return of the risk free asset is 1. Therefore $p^1 = E(m)$ is known. We could plug this value in and then we know the exact value of m. However, if we have doubts about the risk free asset or do not have a risk free asset available in the data, we can use a set of returns to trace out the lower bound on $\sigma(m)$.

For a set of returns, $p = \mathbf{1}$, an I x 1 vector of ones, and we use the following two equations to calculate the lower bound. Note that we use the variance-covariance matrix between equity returns and risk free returns in Table 1 for this computation.

$$b = [\text{Cov}(x,x)]^{-1}[\mathbf{1} - E(m)E(x)] \tag{5}$$

$$\sigma(m) \geq \sqrt{b'\text{Cov}(x,x)b} \tag{6}$$

There is some extra intuition on Hansen-Jagannathan bounds in the case of a scalar equity premium. If we allow $x = r^e - r^1$, then $E(mx) = 0$. Therefore we can use the definition of covariance in the following relation:

$$E(mx) = E(m)E(x) + \text{Cov}(m,x) = E(m)E(x) + \rho_{mx}\sigma_m\sigma_x$$

The correlation between m and x is denoted by $|\rho_{mx}| \leq 1$. So therefore the following relation must hold

$$\frac{|E(x)|}{\sigma_x} \leq \frac{\sigma_m}{E(m)} \tag{7}$$

The right hand side is the ratio of the standard deviation of m to its mean, which is sometimes called the market price of risk. The right hand side is the Sharpe ratio of the equity premium. Thus we can conclude that the market price of risk must be at least the Sharpe ratio of the equity premium.

Since Hansen and Jagannathan discovered a relationship between the expected value of m and its standard deviation, we can find the corresponding values for m using power utility function.

$$m_{t+1} = \beta \frac{c_{t+1}}{c_t}^{-\alpha}$$
$$= \beta \exp(-\alpha(\log(c_{t+1}) - \log(c_t))$$

Ljungqvist-Sargent (2000) write that United States data on per capita consumption of nondurables and services can be suitably approximated by a random walk with drift:

$$\log(c_t) = \mu + \log(c_{t-1}) + \sigma_\varepsilon \varepsilon_t$$

where ε_t i.i.d. $\sim N(0,1)$. Estimates from the data are $\mu = 0.004952$ and $\sigma_\varepsilon = 0.005050$. Therefore our model for m becomes

$$m_{t+1} = \beta \exp(-\alpha\mu - \alpha\sigma_\varepsilon\varepsilon_t)$$
$$\log(m_{t+1}) = \log(\beta) - \alpha\mu - \alpha\sigma_\varepsilon\varepsilon_t$$

Since $\log(m_{t+1})$ is a normally distributed random variable with mean log β - $\alpha\mu$ and variance $\alpha^2\sigma_\varepsilon^2$, we can calculate $E(m)$ and $\sigma(m)$.

$$E(m_{t+1}) = E(\exp(\log(m_{t+1}))) = \beta \exp(-\alpha\mu + \alpha^2\sigma_\varepsilon^2/2)$$
$$\sigma(m_{t+1}) = E(m)[\exp(\alpha^2\sigma_\varepsilon^2) - 1]^{.5}$$

These relationships suggest plotting the Hansen-Jagannathan bounds for many guesses of $E(m)$ in the $E(m)$ vs. $\sigma(m)$ plane, and then also including analogous results for different guesses of α as done in the following figures. Figure 2 shows a parabolic Hansen-Jagannathan bound, with the supergraph shaded. Stochastic discount factors must lie in the shaded region. The discrete diamond points show the coordinates of $(E(m), \sigma(m))$ implied by different, “reasonable,” values of α. Figure 3 shows those same coordinate pairs as α increases to “unreasonable” levels. The CRRA must be around 302 to be in the supergraph of the HJ bound.

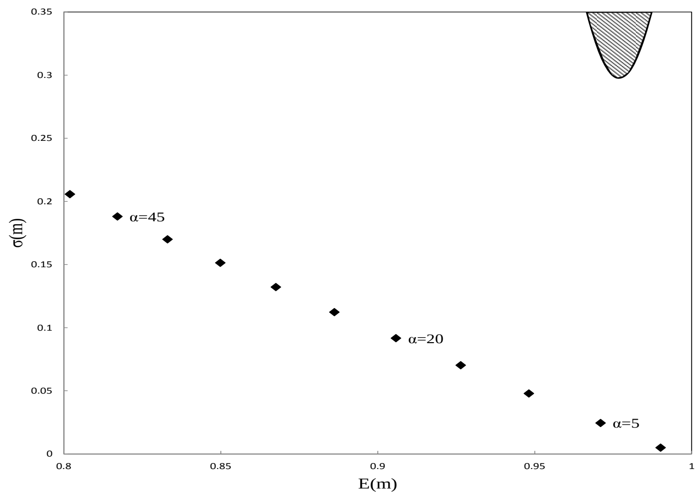

Figure 2: The mean of an SDF vs its standard deviation

The parabola at the top right is the Hansen-Jagannathan bound, and the shaded region is the permissible area for any stochastic discount factor. The diamonds are coordinate pairs of different SDF’s implied by guesses for the CRRA

Inequality 7, in the context of the power utility function representations of $E(m)$ and $\sigma(m)$, can help explain why the power utility function will fail to achieve the observed expected equity premium. The undesirable part of the model is the expression for E(m). The exponent has countering forces. First, the term, $\exp(-\mu\alpha)$ reflects the agents dislike of deviations from smooth consumption paths across time. The parameter, μ, governs deviations from intertemporal consumption smoothness in the random walk. The agent must be compensated for these deviations regardless of risk, and the amount of compensation, in the

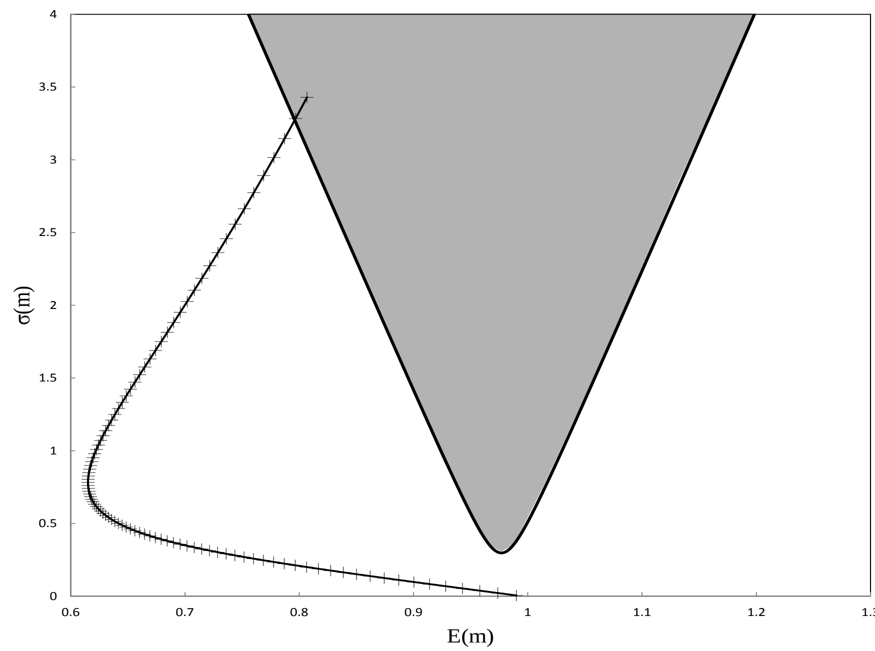

Figure 3: The mean of an SDF vs its standard deviation

The upright parabola shows the same Hansen-Jagannathan bound, and the curve with crosses is the coordinate pairs of different SDFs implied by guesses of the CRRA that range from 1 to 302.

model, is governed by $\exp(-\mu\alpha)$. However, α is meant to control required compensation for bearing risk, and it does so in the second term $\exp(\alpha^2\sigma_\varepsilon^2/2)$. Here a higher α makes the agent require a higher return.

The main problem is that α is being used for two purposes in $E(m)$: compensation for intertemporal consumption smoothing and atemporal risk aversion. In the model, for lower values of α, $E(m)$ is pushed downward because the effect of $\exp(-\mu\alpha)$ dominates the effect of $\exp(\alpha^2\sigma_\varepsilon^2/2)$. This explains the leftward movement of the curve in Figure 3. For higher values of α, $\exp(\alpha^2\sigma_\varepsilon^2/2)$ dominates $\exp(-\mu\alpha)$, so $E(m)$ eventually moves right as in Figure 3. The intertemporal elasticity of substitution for consumption goods and the aversion from risky consumption paths are clearly two different characteristics of a consumer and should not be governed by the same value. Since the stochastic discount factor from the power utility function will yield this undesirable property, it will fail to explain the empirical equity premium.

6 Nonexpected Utility

The problem with the model in the previous section was that the CRRA is used for risk-aversion and intertemporal substitution, which need not be same or even well correlated. Consider the example of a college student and her parents. Both can have the same level of risk aversion but different intertemporal elasticities of substitution. A college student, on a presumably low income, cannot substitute consumption goods today for consumption goods later because saving today (and forgoing consumption) could mean almost no consumption today and lots in the future. This is a huge deviation away from smooth consumption. Even very high interest rates may not persuade a poor student to save. Conversely, her parents, presumably of higher income and saving for retirement, are more sensitive to the interest rate and intertemporal substitution. A small increase in the interest rate is likely to make them save more now without much deviation from a smooth consumption path. While the models discussed here do not consider heterogeneous agents, the example reflects the idea that aggregate market participants need not have a level of risk aversion that is identical to their elasticity of substitution.

To parse out the effects of the intertemporal elasticity of substitution from the CRRA, we use recursive preferences from Epstein and Zin (1989) which are "not necessarily expected."

$$V_t = W(C_t, \mu(V_{t+1}))$$

Here, V is a value function which maps todays consumption and a certainty equivalent of next period's value function onto a value today through W, an aggregator function. The certainty equivalent, as discussed in Section 3 must be of a form:

$$\mu(V_{t+1}) = f^{-1}(E_t f(V_{t+1}))$$

We will give f a concrete function that describes atemporal risk in a form similar to the power utility function:

$$f(z) = \begin{cases} z^{1-\alpha} & : 0 < \alpha \neq 1 \\ \log(z) & : \alpha = 1 \end{cases}$$

Epstein and Zin used the CES aggregator function to give V_t a concrete form

$$W(c,\mu)=\begin{cases}[(1-\beta)c^{1-\eta}+\beta\mu^{1-\eta}]^{\frac{1}{1-\eta}} & : 0<\eta\neq 1\\ c^{1-\beta}\mu^{\beta} & : \eta=1\end{cases}$$

Tallarini allowed $\eta = 1$ as a special case of this model.

$$V_t = c_t^{1-\beta}[(E_t(V_{t+1}^{1-\alpha}))^{1/(1-\alpha)}]^{\beta}$$

If we take logs of both sides and divide by $(1-\beta)$ we get

$$\frac{\log V_t}{1-\beta} = \log(c_t) + \frac{\beta}{(1-\alpha)(1-\beta)}\log E_t(V_{t+1}^{1-\alpha})$$

Now we allow $U_t = \log V_t/(1-\beta)$ and $\theta = \frac{-1}{(1-\beta)(1-\alpha)}$ and we end up with

$$U_t = \log(c_t) - \beta\theta\log E_t[\exp(\frac{-U_{t+1}}{\theta})] \tag{8}$$

With a recursive utility function in hand, we solve for the stochastic discount factor which will be

$$\begin{aligned} m_{t+1} &= \beta\frac{\frac{\partial U_t}{\partial c_{t+1}}}{\frac{\partial U_t}{\partial c_t}}\\ &= \beta\frac{c_t}{c_{t+1}}\frac{\exp(-\theta^{-1}U_{t+1})}{E_t[\exp(-\theta^{-1}U_{t+1})]}\\ &= \beta\frac{c_t}{c_{t+1}}\frac{\exp((1-\beta)(1-\alpha)U_{t+1})}{E_t[\exp((1-\beta)(1-\alpha)U_{t+1})]} \end{aligned}$$

The term $g(\varepsilon_{t+1}) = \frac{\exp((1-\beta)(1-\alpha)U_{t+1})}{E_t[\exp((1-\beta)(1-\alpha)U_{t+1})]}$ is a liklihood ratio, the ratio of one probability density to another. Therefore, we can find

$$g(\varepsilon_{t+1}) = \frac{\exp((1-\beta)(1-\alpha)U_{t+1})}{E_t[\exp((1-\beta)(1-\alpha)U_{t+1})]} = \exp(w\varepsilon_{t+1} - \frac{1}{2}w^2)$$

where $w = -\frac{\sigma_\varepsilon}{\theta(1-\beta)} = \sigma_\varepsilon(1-\alpha)$. So now we can use the same methods from Section 5 to solve for the stochastic discount factor.

$$\begin{aligned} m_{t+1} &= \beta \exp(-(\log(c_{t+1}) - \log(c_t))) \exp(w\varepsilon_{t+1} - \frac{1}{2}w^2) \\ &= \beta \exp(-(\mu + \sigma_\varepsilon \varepsilon_{t+1} - \sigma_\varepsilon(1-\alpha)\varepsilon_{t+1} + \frac{1}{2}\sigma_\varepsilon^2(1-\alpha)^2)) \end{aligned}$$

Therefore

$$\begin{aligned} \log(m_{t+1}) &= \log\beta - \mu - \alpha\sigma_\varepsilon\varepsilon_{t+1} - \frac{1}{2}\sigma_\varepsilon^2(1-\alpha)^2 \\ \log(m_{t+1}) &\sim N(\log\beta - \mu - \frac{1}{2}\sigma_\varepsilon^2(1-\alpha)^2, \alpha^2\sigma_\varepsilon^2 \end{aligned}$$

We can use the above equation to solve for coordinate pairs of $(E(m), \sigma(m))$ implied by non-expected utility. Figure 4 shows these coordinate pairs along with the coordinate pairs derived from a power utility function as in Section 5. Clearly, our new coordinates are heading in the right direction, but we still do not acheive the Hansen-Jagannathan bound till about a CRRA of over 80.

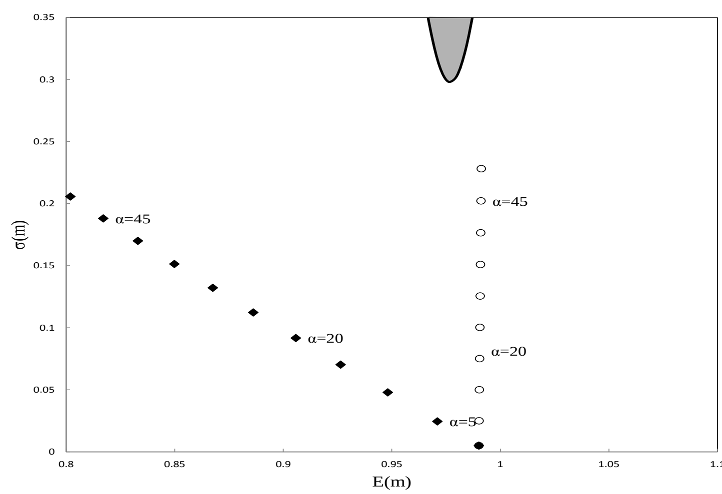

Figure 4: The mean of an SDF vs its standard deviation

The parabola at the top right is the Hansen-Jagannathan bound, and the shaded region is the permissible area for any stochastic discount factor. The diamonds are coordinate pairs of different SDF's implied by guesses for the CRRA. The circles show the coordinate pairs using the same values for the CRRA implied by Tallarini's SDF.

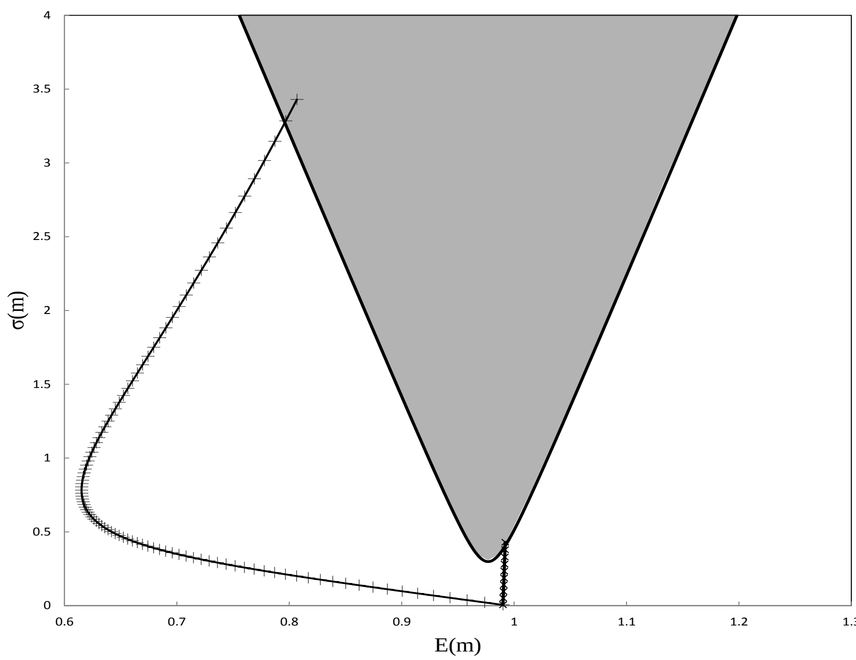

Figure 5: The mean of an SDF vs its standard deviation

The upright parabola shows the same Hansen-Jagannathan bound, and the curve with crosses is the coordinate pairs of different SDFs implied by guesses of the CRRA that range from 1 to 302. The plus signs give the Tallarini coordinate pairs for CRRAs that range from 1 to 82

7 Conclusion

The equity premium puzzle has evolved from the observation that common models do not explain simple empirical results to a fruitful exercise in developing macroeconomic theory. We have seen how the key to asset pricing lies in the stochastic discount factor m. Hansen and Jagannathan showed how we can use data to put a lower bound on any m in an economy. Epstein and Zin and Tallarini showed how deeper preferences can make a stochastic discount factor more reasonable. Yet the puzzle is still not solved, and research continues to make more realistic preferences and give more explanatory power to macroeconomic models.

The research that goes into the equity premium puzzle is more important than making the model fit the data. It allows us to make more conclusions about the broader well-being of the economy. How can capital markets better suit the needs of consumers and what are the costs of risk and uncertainty? The puzzle is a great motivator of research and seeking a conclusion has opened up more interesting questions which makes it invaluable to the field.

References

Barillas, F., L. P. Hansen and T. J. Sargent (2009) “Doubts or variability” *Journal of Economic Theory*, Vol 144, Iss 6, 2388-2418

Dimson, E., P. Marsh and M. Staunton (2006), “The Worldwide Equity Premium: A Smaller Puzzle” *Zurich Meetings Paper, AFA 2008*

Epstein, L. G. and Zin, S. E. (1989), “Substitution, Risk Avesion, and the Temporal Behavior of Consumption and Asset Returns: A Theoretical Framework” *Econometrica*, Vol. 57, Iss. 4, 937-969.

Hansen, L.P. an R. Jagannathan (1991), “Implications of Security Market Data for Models of Dynamic Economies” *Journal of Political Economy*, 99, 225-262

Kandel, S and R. F. Stambaugh (1991), “Asset Returns and Intertemporal Preferences” *Journal of Monetary Economics*, 1, 39-71

Kocherlakota, N. R. (1996), “The Equity Premium: It’s Still a Puzzle” *Journal of Economic Literature* Vol. 34(1), 42-71

Ljungqvist, L. and T. J. Sargent (2000) “Recursive Macroeconomic Theory” *The MIT Press*, 1st Edition.

Lucas, R. E. (1978), “Asset Prices in an Exchange Economy” *Econometrica*, Vol. 46, Iss. 6, 1429-1445

Mankiw, G. and S. P. Zeldes (1991), “The consumption of stockholders and nonstockholders” *Journal of Financial Economics*, 29, 97-112

Mehra, R. and E. C. Prescott (1985) “The Equity Premium: A Puzzle” *Journal of Monetary Economics* 15, 145-161

Mehra, R. and E. C. Prescott (2003) “The equity premium in retrospect” *Handbook of the*

Economics of Finance, edition 1, volume 1, chapter 14, pages 889-938 Elsevier

Shiller, R. (2009). “Stock Market Data” http://www.econ.yale.edu/ shiller/data.htm April 20, 2012.

Tallarini Jr., T. D. (2000)“Risk-sensitive Real Business Cycles” *Journal of Monetary Economics*, 45, 507-532