Why Risk and Return are Uncorrelated:
A Relative Status Approach

Eric Falkenstein
Eden Prairie, Minnesota, USA
This draft: 10/7/06

Abstract

This paper presents a utility function refinement that explains the empirical irrelevance of risk to returns. The key is that in an environment where people care about relative wealth, risk is a deviation from what everyone else is doing, and therefore becomes like diversifiable risk in the CAPM, avoidable. Using an equilibrium or an arbitrage argument, a relative status utility function creates a zero risk-return correlation via a market model that implies a zero risk premium. This approach is described as being theoretically consistent, intuitive and a better description of the data.

JEL classification: D01, D81, G11, G12
Keywords: status, relative risk, CAPM, equity risk premium, cross-sectional asset pricing

Contact author at efalken@gmail.com. Comments welcome.
Introduction

Imagine a world were people only cared about their wealth relative to their neighbors. With all investors averse to deviations from the average return, a ‘risk-free’ portfolio in this context would be the same average portfolio because that would guarantee an average return. Nondiversifiable ‘risk’ can be avoided in an equilibrium where everyone holds the same market portfolio, making it similar to diversifiable risk in a world of absolute returns. The theory, intuition and empirical support for this approach are discussed below.

Most economists assume nondiversifiable volatility is positively and linearly related to returns. The application of standard utility assumptions and statistics to create the CAPM is a ‘theoretical tour de force’, yet also has virtually no empirical support (Fama and French (2006)).\(^1\) It is rare that risk, when measured directly, is even positively correlated with higher returns, let alone linearly. Nevertheless, the logic of the theory is so sound that the CAPM remains a pillar of finance curriculums, and most economic models generate a positive variance-return relation adjusted by a risk aversion coefficient. Bill Sharpe probably speaks for most economists when he states that in spite of empirical difficulties, “it would be irresponsible to assume that [the CAPM] is not true.”\(^2\)

This paper argues that there is no risk-reward relation because it is consistent with how people internalize risky decisions that comes from comparing themselves to others, as opposed to the standard approach to risk as absolute volatility of their portfolio, as in Rothschild and Stiglitz (1970). A concern for relative status has been long recognized, as anthropologist Donald Brown (1999) noted that concern for relative status is a ‘human

---

\(^1\)Quote from and statement from Fama and French (2006): “The CAPM’s more general problem is that variation in \(\beta\) unrelated to size and the value-growth characteristic goes unrewarded throughout 1926 to 2004.”

\(^2\) See his interview in Burton (1998)
universal’ documented in every known human culture.\(^3\) Richard Easterlin (1974) documents a number of studies find that self-reported happiness may be more sensitive to relative than to absolute income. In experiments, Robert Axelrod (1984) notes he puts students in game situations, instructing them “that it should not matter to them whether they score a little better or a little worse than the other player, so long as they collect as many dollars for themselves as possible. These instructions simply do not work. The students look for a standard of comparison to see if they are doing well or poorly.” In the world of money management especially relevant to the CAPM, relative risk is an institutional reality, as Siegel (2002) notes that “performance is measured not by absolute return, but the return relative to some benchmark.”

It can be argued that we should derive satisfaction from what we have, not in comparison to others, but similar normative advice could be made about people wanting more absolute wealth once some base level of material comfort is obtained.\(^4\) Envy may be petty, but perhaps no more so than greed, which economists have long understood as a mechanism for efficiency in a competitive economy. Consider that the typical poor person today has real conveniences that would make them rich relative to a century ago: the average American classified as poor is not hungry, has a car, air conditioning, a refrigerator, stove, clothes washer and dryer, microwave, two color televisions, etc. (see Rector and Johnson (2004)). Nevertheless, no one would argue that poverty has disappeared. In 1930 John Maynard Keynes wrote an optimistic essay where he foresaw a bright future in 100 years: a standard of living 4 to 8 times higher.\(^5\) In such a world interest rates might fall to zero, and

\(^3\) H.L. Mencken defined wealthy as “any income that is at least one hundred dollars more a year than the income of one's wife's sister's husband.”


\(^5\) Economics of our Grandchildren (1930)
goods and services would become so cheap that leisure would be the biggest challenge. Keynes was right on the material prosperity, but wrong as to what it would portend, as happiness has not trended upward during a period of several fold growth (Easterbrook (2004)). Many studies directly address the ‘problem’ of income dispersion in a growing economy, even when each group sees wealth increases (see Dew-Becker and Gordon (2005) or Kaus (1998)). That this is a problem implies relative status is important. If we take this preference seriously in modeling utility functions it has a non-obvious implication: risk does not relate to returns.

While the idea that people care about their relative status is well-acknowledged, it has not become a standard for utility functions. Nonetheless, there is a considerable growing precedence in the academic literature. Thorstein Veblen popularized the notion of utility depending on social rank in his seminal work on ‘conspicuous consumption’ back in 1899, first modeled formally by James Duesenberry (1949) as a way to explain why people in rich countries save more than people in poor countries. Positional goods, ‘saving face’, preoccupation about ‘disrespect’ are all related to status in society, and have come under increased study as a way of explaining otherwise irrational behavior (see Frank (2000), Kim (1993) and Stewart (2006), respectively). Popular utility models that have context-dependent reference points include the hedonic treadmill of Brickman and Campbell (1971), habit formation model of Constantinides (1990), and the ‘keeping up with the Joneses’ utility function of Abel (1990). Economists have shown that an incorporation of a concern for social context alters results of traditional growth models (Cole, Mailath and Postlewaite (1992), tax policy (Bagwell and Bernheim (1996)), social norms (Bernheim (1994)), charity (Cialdini and Trost (1998)) and repeated interactions (Sobel (2005)).
Why should this be so? Neurologically, since humans are quintessentially social animals, much of the brain is given over to processing social information, including in-group status. Biologists Insel and Fernald (2004) argue that because information about social status is essential for reproduction and survival, it seems likely that specialized neural mechanisms have evolved to process social information. Status orientation is hard-wired into our brains as a consequence of evolutionary selection. In economics, a status-oriented utility function may be more efficient, as optimizing status has been modeled as a useful signaling device (Pesendorfer (1995)) where status goods signal ability. Rayo and Becker (2006) model a scenario where maximizing status is efficient because, like an eyeball adjusting to brightness, one gets greater precision in their signals by measuring productivity in context, relative to a group average (think of the usefulness of evaluating mutual fund managers using ranking, rather than their absolute returns).

If utility is a status function, specifically the variation in wealth to one’s peers, only bets away from the consensus are ‘risky’. Risk is symmetrical in that it can come from taking on too small just as too great an exposure in contrast the idea that risk increases monotonically from taking no exposure. In life, philosophers have long noted that too little courage is cowardly, too much foolhardy, both vices. Similarly, too little exposure to some asset class is just as bad as too much, one cowardly and one reckless, objectively measured by how far these bets take us from the average.

---

6 see Cremers and Petajisto (2006) where they quantify active portfolio management in mutual funds in this manner
I. Risk Taking in a Relative Risk Environment

Economists have long suspected, and psychologists have amply documented that people are generally overconfident about their abilities in vocational and avocational interests (see Alpert and Raiffa (1982), or Gervais, Heaton, and Odean (2003)).

Santos-Pinto and Sobel (2005) note that contemporary psychologists agree that ‘on nearly any dimension that is both subjective and socially desirable, most people see themselves as better than average’. Overconfidence invites people to deviate from the norm—to take risk—because they think they know some fact or theory most other people do not sufficiently appreciate (Barber and Odean (1999) document excessive retail trading which they relate to overconfidence).

For a truly risky investment a proactive choice is made to stake a significant portion of one’s net worth in a decision the average person is not making, and therefore reflects on their judgment, not just their risk tolerance. The success of this judgment is a combination of luck and skill, so its failure is more than just bad luck. In the CAPM only the investor knows how far his portfolio was from his utility curve, and a diversified portfolio, even an extremely high beta one, is never objectively suboptimal. In contrast, risky investment decisions generate anxiety precisely because when they fail they reflect on our judgment, not merely our unobservable risk preference, because risk taking reflects a subjective belief that they know better than the market. A risky remark, career move, or fashion choice all involve something a choice along the Security Market Line does not: the potential shame for having thought it was ever a good idea. Ex post, one’s ex ante overconfidence is exposed as baseless, ignorant. Frederick (2005) shows a strong correlation between a metric of IQ and

---

7 For Aristotle summmum bonum is always a middle path between two opposites. The virtues of courage and temperance, mediated by prudence.
8 “the over-weening conceit which the greater part of men have of their own abilities, is an ancient evil remarked by the philosophers and moralists of all ages.” Adam Smith
one’s immunization to framing effects, suggesting that ability is related to proficiency in risky environments, and this probably extends to uncertain ones. The successful risk-taker is, on average, courageous, fortunate and skilled, as opposed to being merely courageous and fortunate.

It should be noted that idiosyncratic risk affects one’s wealth relative to the mean similarly to the case where one considers that people solely care about absolute wealth volatility. Insurance and risk mitigation techniques like hedging are therefore unaffected by using this approach compared with the standard approach.

II. Model of Status Risk

A. The Basic Idea

The model below shows the basis of how a concern for a relative measure of wealth generates the equivalent riskiness to low and high beta assets. This section outlines two methods of showing this result, including the corollary that in this situation, that equivalent return is equal to the risk-free return. I also show how differences of opinion relate to each other in equilibrium.
Assume there are two assets, X and Y, and two states of nature, 1 and 2. An investor faced with asset X or Y can see the following payoffs.

<table>
<thead>
<tr>
<th>State</th>
<th>Total Return</th>
<th>Avg</th>
<th>Relative Return</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>X</td>
<td>Y</td>
<td>X</td>
</tr>
<tr>
<td>State 1</td>
<td>0</td>
<td>-10</td>
<td>-5</td>
</tr>
<tr>
<td>State 2</td>
<td>+20</td>
<td>+30</td>
<td>+25</td>
</tr>
</tbody>
</table>

Y is considered riskier in standard theory as shown in the Figure 1, with a 40 point range, versus a 20 point range for X. Yet on a relative basis, each asset generates identical risk. In State 1 X is a +5 outperformer, in State 2 X is a -5 underperformer, vice versa for asset Y. In relative return space, however, the higher absolute volatility asset is not riskier. If X and Y are the only two assets in the economy, relative risk can be achieved, equivalently, by taking on an undiversified bet on X or Y. Buying ‘the market’, or allocating half of each, meanwhile, generates zero risk.

Everything really flows from this simple insight. The key is that the relative portfolio wealth is the argument in the utility function.

B. A Simple Equilibrium Model

Assume a two-period economy with two identical individuals, $i$ and $-i$. There are two types of assets, a risk-free bond that pays off 1 with certainty in period 1. Its price in period 0 is $P_{rf}^0$. There also exists an equity with payoff in period 1 of $P_E^1$, where
Here $\beta$ is a constant reflecting the equity’s factor sensitivity, and $r_m$ is distributed $N(\mu_m, \sigma_m^2)$, the only source of conventional risk in this model. Total wealth for the individual $i$ in period 0 is given by his portfolio of assets.

$$w_i^0 = \alpha_E^i P_E^0 + \alpha_f^i P_{rf}^0$$

Where $P_E^0$ and $P_{rf}^0$ are the price of the equity and the risk free bond, in period 0, respectively. $\alpha_E^i$ and $\alpha_f^i$ represent the holdings for investor $i$ on the risky and risk-free asset, respectively. Each individual is endowed with $k$ units of wealth, so the budget constraint is

$$w_i^0 \leq k$$

The agent $i$'s utility function is driven by his wealth relative to the other agent (there is no consumption) in an exponential utility function with a risk aversion coefficient $a$.

$$U_i(w_{1-i}^1 - w_i^1) = -\exp\left(-a\left(w_{1-i}^1 - w_i^1\right)\right)$$

Thus the individual maximizes the utility function

$$\text{Max}_{\alpha_E^i, \alpha_f^i} E\left(w_i^1 - w_{1-i}^1\right) - \frac{a}{2} \text{Var}\left(w_i^1 - w_{1-i}^1\right)$$

Note that his utility is strictly increasing in his own wealth, $w_i^1$, and strictly decreasing in the wealth the other investor, $w_{1-i}^1$. Also, the variance of his difference from this other investor negatively impacts his utility, in an amount proportional to his coefficient of risk aversion, $a$. Substituting for $w_i^1$ and $w_{1-i}^1$ and applying the expectations operator, this problem expands to
\[
\begin{align*}
\text{Max} \quad & \alpha_E^i (1 + \beta \mu_m) + \alpha_f^i - \alpha_E^{-i} (1 + \beta \mu_m) - \alpha_f^{-i} - \frac{a \sigma_m^2}{2} ((\alpha_E^i - \alpha_E^{-i}) \beta)^2 \\
\text{s.t.} \quad & w_i^0 \leq k
\end{align*}
\]

As there is no consumption, the agent can do nothing to affect his period 0 wealth, so his relevant decision only concerns optimizing for period 1. His utility is strictly increasing in \(\alpha_E^i\) and \(\alpha_f^i\), so his budget constraint holds with equality, that is

\[
\alpha_f^i = \frac{k - \alpha_E^i P_E^0}{P_{rf}^0}
\]

Substituting for \(\alpha_f^i\), and through some algebra, generates

\[
\begin{align*}
\text{Max} \quad & (\alpha_E^i - \alpha_E^{-i}) (1 + \beta \mu_m) + \frac{k - (\alpha_E^i - \alpha_E^{-i}) P_E^0}{P_{rf}^0} - \frac{a \sigma_m^2}{2} ((\alpha_E^i - \alpha_E^{-i}) \beta)^2 \\
\text{Taking the first order conditions with respect to} \quad & \alpha_E^i, \text{we have}
\end{align*}
\]

\[
P_E^0 = \frac{1 + \beta \mu_m - a \sigma_m^2 \beta^2 (\alpha_E^i - \alpha_E^{-i})}{1/P_{rf}^0}
\]

**Proposition 1**: In a rational expectations equilibrium, the expected return on all assets is identical.

**Proof**: Since each agent is identical, in equilibrium each agent holds the same amount

\[
\alpha_E^i = \alpha_E^{-i}
\]

Given equations (9) and (10), we have

\[
P_E^0 = P_{rf}^0 (1 + \beta \mu_m)
\]

As \(E(P_E^1) = 1 + \beta \mu_m\), the expected gross return on the equity is thus

\[
\frac{E(P_E^1)}{P_E^0} = \frac{1 + \beta \mu_m}{P_{rf}^0 (1 + \beta \mu_m)} = \frac{1}{P_{rf}^0}
\]
This is the same as the expected gross return on the risk free bond, by definition, $\frac{1}{P_{rf}}$.

Q.E.D.

And as there is no $\beta$, no risk aversion coefficient $a$, in this equation, the same expected return holds regardless of its sensitivity to the market factor $r_m$, or the volatility of the market, $\sigma_m$. Thus the returns on all assets, risk-free and risky, are the same. Risk does not affect return in equilibrium.

C. The Arbitrage Model

We can also see how this result holds through arbitrage if investors are concerned about relative performance. Assume an economy with risky assets that are a function of a ‘market’ factor, $r_m$. For any investor $i$ who chooses an asset with a specific beta $\beta_i$, his returns are generated via the factor model

$$r_i = \mu_{\beta_i} + \beta_i r_m$$  \hspace{1cm} (13)

Where $\mu_{\beta_i}$ is a constant for a portfolio with the specific beta $\beta_i$, and $r_m \sim N(\mu_m, \sigma_m^2)$. The return on the risk free asset is just the constant $r_f$. The prices of both the risky and risky free asset are both equal to one, so that we are solving for a $r_f$, $\mu$ and $\mu_m$ such that this is an equilibrium.

Assume that the market return is the benchmark to which investors compare themselves, just as mutual fund managers typically try to outperform their benchmark. Their objective is to maximize their outperformance subject to minimizing its variance. Define $r_{out}^i$, which is the relative performance of investor $i$ to the market return.
\[ r^i_{out} = r_i - r_m \quad (14) \]

\( r_i \) is the return on the investor’s portfolio with its particular factor loading \( \beta_i \), and \( r_m \) is the return on the market. Investors all have the simple objective of maximizing \( r^i_{out} \) while minimizing a proportion of its variance, as in

\[ \text{Max} \quad r^i_{out} - \frac{a}{2} \sigma_i^2 \quad (15) \]

Where \( \sigma_i^2 = \text{Var}(r_i - r_m) \). Substituting equation (13) into (14) generates

\[ r^i_{out} = \mu_i + (\beta_i - 1)r_m \quad (16) \]

Since \( r_m \) is the only random variable, the variance of outperformance is just

\[ \sigma_i^2 = (\beta_i - 1)^2 \sigma_m^2 \quad (17) \]

We can replicate the relevant risk of a stock with a beta of \( \beta \), \( \sigma_i^2 \), via a portfolio consisting of \( \beta \) units of the market portfolio, borrowing of \((1-\beta)\) units of the risk free asset (cost is \( \beta - (1-\beta) = 1 \), same as for the stock). Arbitrage then implies these have the same expected returns, so

\[ E(\beta r_m + (1-\beta)r_f) = E(\mu_\beta + \beta r_m) \quad (18) \]

Here the LHS of equation (18) is leveraged \( \beta \) times in the market portfolio and borrowing \((1-\beta)\) in the risk free asset in financing, while the RHS is the unlevered \( \beta \) portfolio via equation (14). They have the same factor exposure, and cost the same (since prices for all assets are 1 by assumption), so they should have the same return in equilibrium. Equation (18) implies

\[ \mu_\beta = (1-\beta)r_f \quad (19) \]

This allows us to replace the \( \mu_\beta \) with \((1-\beta)r_f\) in equation (16) and leads to the factor model.
If the riskiness of relevance to investors is their outperformance, $\sigma_i^2$, the expected return for assets with $\beta = x$ should be the same as those with $\beta = 2 - x$, because they have the same risk in this environment: $(x - 1)^2 = ((2 - x) - 1)^2$. The risk of a $\beta = 2$ asset is identical in magnitude to that of a $\beta = 0$ asset. The expected returns must be the same

$$E(r_{out} | \beta = x) = E(r_{out} | \beta = 2 - x)$$

(21)

Since the expected return for a $\beta = x$ asset equals the expected return for a $\beta = 2 - x$ asset, using equation (20) and applying the expectations operator, we have

$$(1 - x)r_f + (x - 1)E(r_m) = (1 - (2 - x))r_f + ((2 - x) - 1)E(r_m)$$

(22)

Where the LHS of equation (22) is the expected return on the $\beta = x$ asset, and the RHS is the expected return on the $\beta = 2 - x$ asset. Solving for $E(r_m)$ we get

$$E(r_m) = r_f$$

(23)

Equations (13), (19) and (23) imply

$$r_i = r_f + \beta_i (r_m - r_f)$$

$$r_m = N(r_f, \sigma_m)$$

(24)

Thus no arbitrage, in the sense that things equivalent in risk are priced the same (as risk is defined here), generates the traditional CAPM with the significant difference that the market return is equivalent to the risk free rate. Just as the equilibrium model in prior section implies, the expected return on all assets is the same, because $E\beta(r_m - r_f) = 0 \forall \beta$.

**D. Incorporating Risk Takers**

In equilibrium the return on all assets is the same because holding the same portfolio is unique risk minimizing portfolio and no one is forced to take risk. However, clearly people
often do take risks, and in the process create a situation where they hold a more-than-average amount of an asset, implying others to hold a less-than-average amount.

Using the arbitrage approach of II.C, let us assume an economy with two risky assets, A and B, and a risk free asset

$$r_m = \pi_B r_A + \pi_B r_B + (1 - \pi_A - \pi_B) r_f$$

(25)

Here $\pi_A$ represents the share of the market occupied by portfolio A, $\pi_B$ the share for portfolio B, and $(1 - \pi_A - \pi_B)$ represents the share of wealth occupied by the risk free asset, which has a return of $r_f$.

Assume there are two investors (or, equivalently, two sets of investors), investor $i$ and investor -$i$, who represent all of the investors. Specifically, let $\alpha$ represent the share of wealth owned by investor $i$, so that

$$\pi_A = \alpha z^i_A + (1 - \alpha) z^i_B$$

(26)

Where $z^i_A$ is the proportion of investor $i$'s portfolio invested in asset A, and similarly for $z^i_B$.

A similar identity holds for portfolio B. Investor $i$ looks at his portfolio return as

$$r_i = z^i_A r_A + z^i_B r_B + (1 - z^i_A - z^i_B) r_f$$

(27)

In this case, his outperformance, define as above as $r^{i}_{out} = r_i - r_m$ is thus

$$r^{i}_{out} = z^i_A r_A + z^i_B r_B + (1 - z^i_A + z^i_B) r_f - \pi_A r_A - \pi_B r_B - (1 - \pi_A r_A - \pi_B r_B) r_f$$

(28)

Let the returns for the portfolios be represented by

$$r_A = r_f + \beta (r_m - r_f) + \theta$$

$$r_B = r_f + \beta (r_m - r_f)$$

(29)

Here $\theta$ is a value that allows investor $i$ to have optimistic or pessimistic expectations for portfolio A, an 'abnormal return factor'. We will assume investors perceive $\theta$ differently so that each investor, $i$ and -$i$, believe in different means for this abnormal return factor. Thus
\[ \theta - N(\mu, \sigma^2) \text{, and } E[\theta] = \mu_i. \]  
If \( \mu_i \) is positive, \( i \) is an optimist, if negative a pessimist.

We will assume \( \sigma^2 \) is the same for each investor and \( \beta \) is the same for each asset for ease of exposition.

Some algebra simplifies equation (29) to

\[ r_{out}^i = (z_A^i - \bar{z_A})r_A + (z_B^i - \bar{z_B})r_B + (\bar{z_A} + \bar{z_B} - z_A^i - z_B^i) r_f \]  

Total variance of \( r_{out}^i \) is thus

\[ \sigma_i^2 = \left[ \sigma_m^2 \left( (z_A^i - \bar{z_A}) \beta + (z_B^i - \bar{z_B}) \beta \right)^2 + \left( z_A^i - \bar{z_A} \right)^2 \sigma_\theta^2 \right] \]  

In this situation, investor \( i \) maximizes \( r_{out}^i - \theta^2 \sigma_i^2 \), generating

\[ \text{Max} \ E_i \left( (z_A^i - \bar{z_A})r_A + (z_B^i - \bar{z_B})r_B + (\bar{z_A} + \bar{z_B} - z_A^i - z_B^i) r_f \right. \]  

\[ - \frac{a}{2} \left[ \sigma_m^2 \left( (z_A^i - \bar{z_A}) \beta + (z_B^i - \bar{z_B}) \beta \right)^2 + \left( z_A^i - \bar{z_A} \right)^2 \sigma_\theta^2 \right] \]  

Applying the expectations operator, his first order condition for \( z_A^i \) generates

\[ \beta (r_m - r_f) + \mu_i - a \sigma_m^2 \beta \left( (z_A^i - \bar{z_A}) \beta + (z_B^i - \bar{z_B}) \beta \right) - a \sigma_\theta^2 (z_A^i - \bar{z_A}) = 0 \]  

While his first order condition for \( z_B^i \) generates

\[ \beta (r_m - r_f) - a \sigma_m^2 \beta \left( (z_A^i - \bar{z_A}) + (z_B^i - \bar{z_B}) \right) = 0 \]  

Substituting equation (34) into (33), yields

\[ \beta_A (r_m - r_f) + \mu_i - \beta_B (r_m - r_f) - a_i \sigma_\theta^2 (z_A^i - \bar{z_A}) = 0 \]  

Which simplifies to

\[ z_A^i = \frac{\mu_i}{a_i \sigma_\theta^2} + \bar{z_A} \]  

Note that, as in section C above, if the agents are identical (i.e., \( z_A^i = \bar{z_A} = z_A \), and \( \mu_i = \mu_{-i} = 0 \)), the first order conditions (33) and (34) imply \( E[r_m] = r_f \).
Given investor \( -i \), the complement to investor \( i \), optimizes a symmetric problem, and substituting for \( z^i \) and \( z^{-i} \) given equation (26), we have
\[
\bar{z}_A = \alpha \left( \frac{\mu_i}{a \sigma_{\theta}^2} + \bar{z}_A \right) + (1 - \alpha) \left( \frac{\mu_{-i}}{a \sigma_{\theta}^2} + \bar{z}_A \right)
\]
(37)
\[
\mu_i = \frac{(1 - \alpha)}{\alpha} \mu_{-i}
\]
(38)
What equation (38) says is that this can only be an equilibrium when the optimism of investor \( i \) has an equal, countervailing pessimism for investor \( -i \). \( \frac{(1 - \alpha)}{\alpha} \) represents the inverse of the relative proportion of investor \( i \) in the economy, so that if he is small, \( \frac{(1 - \alpha)}{\alpha} \) is large, meaning in equilibrium the optimism (pessimism) of investor \( i \), \( \mu_i \), is much larger than the pessimism (optimism) of investor \( -i \), \( \mu_{-i} \). A lone active investor taking a single position would necessitate an insignificant amount of corresponding pessimism by everyone else.

If \( \sigma_{\theta} \) is zero, and the subjective precision of their estimate of the abnormal return is perfect, there is no uncertainty as to the abnormal return in portfolio, and so there is no equilibrium. This is because a perfect precision investor would then be motivated to take an infinitely leveraged position of portfolio A vs. B and create an arbitrage, and his complement would be taking an infinite position on the other side.

If the optimistic investor is delusional, the pessimist is correct and the return on the stock is less than for the stock where there are no optimists. The fact that a subset of investors has accumulated a disproportionate amount of an asset tells you only that someone sees extra value here that another class does not, but it does not imply that the optimist or pessimist is correct.
III. The Empirical Case for All Assets Having the Same Expected Return

A. Evidence of a zero Risk-Return correlation

For 40 years researchers have looked for corroborating evidence for the idea that risk and reward have a positive and linear relationship, initially looking at betas derived against the US market indices versus realized returns. Since Fama and French’s seminal 1992 study documenting that the relation between beta and returns is essentially flat, many have accepted that there simply is no correlation between beta and cross-sectional returns as a practical matter. If risk is positively related to expected return, people’s perception of risk must be like facial recognition, something intuitive and common yet exceedingly difficult to model. Attempts to resurrect the CAPM have involved time-varying betas, market premiums, and augmenting indices with factors that represent human capital, sometimes simultaneously (Jagannathan and Wang (1996)). Lewellen and Nagel (2006) find time-varying betas and risk factors necessitate implausible variation in betas and equity premiums. Other attempts have focused on separating a beta into two parts, representing covariance with distinct entities that can be related to cash flows and discount rates (Campbell and Vuolteenaho (2004) or idiosyncratic consumption risk and aggregate consumption risk (Jacobs and Wang (2001)), or a risk component and a hedge component (Guo and Whitelaw (2003)). That adding degrees of freedom to a model allows it to explain the data better should surprise no one.

The original application of the APT focused on portfolios gleaned from factor analysis because this made the most mean-variance sense. If one is minimizing portfolio volatility and cross-sectional volatility comes from several factors, all significant factors need to be addressed. Yet the first factor, by orders of magnitude the largest, is highly correlated with the equal-weighted stock index (Jones (2001)), and given this has been often used as a
proxy for the market (under the assumption that the equal-weighted index might be a better measure of the true market, because it gives weight to small companies that proxy many unlisted companies), and the CAPM does not work much better with that market proxy, it is unsurprising that this approach has not generated significantly better results. The statistical APT approach has been quietly abandoned in recognition of its lack of explanatory power and intuition for practitioners. The now dominant extension of the CAPM draws on the APT as inspiration. The current CAPM/APT empirical work uses a three (book/market, size, market) or four (add momentum) factor model which makes sense if you believe that expected return and risk are positively related, and so a positive return for a characteristic must imply some kind of risk. However, the main factors, size and book/market, do not have strong covariances with macro factors, and in the case of book/market the factor loadings have the wrong sign (Lakonishok, Shleifer and Vishny (1994)). The size effect is not risky in any obvious sense, other than that it is a somewhat independent factor, though there are many such independent volatile groupings, such as industries, and these do not generate a premium.\(^9\) Daniel and Titman (1997) document that stocks with similar size or book/market characteristics, but not factor loadings, have similar returns, implying it is the characteristic, not their APT riskiness, that generates the return premiums to these two factors. Conversely, portfolios based on having similar factor loadings as the size and book/market portfolios, but different characteristics, have dissimilar returns to the size or book/market portfolios. Similarly, Houge and Loughran (2006) find that mutual funds with the highest loadings on the value-factor (French’s HML) reported no return premium over the same 1975-02 period, even though the HML factor generated a 6.2% average annual return over that same period. *The current factor approach was created to explain the very characteristics* 

---

\(^9\) In an interview with Eugene Fama. Q: “What is the nature of the size risk?” A: “You know, that's an
that are used as factors. That they have little external validation (e.g., via the exercises of Houge and Loughran), or the fact that as sources of variance they are not uniquely large, or strongly correlated with proxies for ‘the market’, suggests this approach is more a description of returns, rather than risk within the APT approach.

It has been suggested (Fama (1993)) that size and book/market represent distress risk, which is potentially trickier to measure than a statistical factor because it has a ‘peso-problem’ tendency to occasionally jump to near zero, or is linked to difficult-to-measure human capital. Several researchers have used models (or agency ratings) shown to statistically predict defaults to serve as a metric of financial distress risk, and then evaluated the equity performance of these models.\(^\text{10}\) In general they find that when they measure distress directly, as opposed to inferring it from size or book/value, these distressed stocks have delivered anomalously low returns, patently inconsistent with the conjecture that the value and size effects are compensation for the risk of financial distress. Campbell, Hilscher, and Szilagyi (2006) find that the distress factor can hardly explain the size and book/market factors; in fact it merely creates another anomaly because the returns are significantly in the wrong direction. Distressed firms have much higher volatility, market betas, and loadings on value and small-cap risk factors than stocks with a low risk of failure. These patterns hold in all size quintiles but are particularly strong in smaller stocks.

As one of the characteristics of distressed firms is higher volatility, it should come as little surprise that idiosyncratic volatility itself would create an anomaly. Ang, Hodrick, Xing and Zhang (2005, 2006) document that idiosyncratic variance is actually negatively correlated with total return, using data from seven of the worlds largest equity markets.\(^\text{11}\) This is

---

\(^{10}\) Dichev (1997), Agarwal and Taffler (2002), Griffin and Lemon (2002), and Vassalou and Xing (2003)

\(^{11}\) Canada, France, Germany, Italy, Japan, the U.S. and the U.K.
especially troubling to a risk/reward theory because idiosyncratic variance should pick up mismeasured factor loadings (Lehmann (1990)), and mismeasured factor loadings should help explain the poor performance of factor models such as the APT. For example, Ross and Roll (1994) point out that it is possible—though not probable—that inefficient estimates of the market portfolio are uncorrelated with returns, yet then residuals should still show a positive correlation with returns. That is, if explicit risk loadings are not empirically correlated with cross-sectional returns, the residual which will capture mismeasured risk loadings is theoretically positively correlated with returns. In practice, the correlation for residual risk goes the wrong way.

In a related finding, differences of opinion were found to be negatively correlated with stock returns. Such differences in opinion could proxy for parameter uncertainty, a perhaps better estimate of risk (Hansen and Sargent (2001)). Using analysts’ earning forecasts as a proxy for differences of opinion among investors, Diether, Malloy and Scherbina (2002) find the quintile of stocks with the greatest opinion dispersions underperformed a portfolio of otherwise similar stocks. Diether et al note that higher estimate dispersion is positively related to Beta, volatility, earnings variability. It is difficult to reconcile their findings with the idea that risk is being positively priced.

Haugen (1995) argued that empirically firms with generally low risk qualities such as lower beta, lower leverage, greater profitability, larger size, higher dividend yield, and have the highest returns. Subsequent research on distressed, volatile and uncertain companies has corroborated that finding. If the returns are lower for companies with, seemingly, higher risk, their ‘abnormal’ return in a factor model is an even greater anomaly.

Looking outside public equity returns, evidence for a straightforward positive risk-return relationship has been hard to come by. Moskowitz and Vissing-Jørgensen (2002) find
that S-Corps and Franchises have returns approximately equal to the equity market. These investments are generally large portions of a household’s wealth, are highly volatile, have a large chance of failing, are correlated with the public market indices, and correlated with the owner's human capital. If one assumes that it is impossible to hedge this risk away, because significant ownership is necessary to generate proper incentives for small business owners, these types of investments will be very risky, and therefore should generate a significant return premium to passive equity benchmarks. Heaton and Lucas (2001) calibrate that a conservative relative risk aversion coefficient of 2 generates a 10% risk premium to the market portfolio. Moskowitz and Vissing-Jørgensen find no return premium relative to the market portfolio.

Within the corporate bond universe risk is poorly rewarded. Data on junk bonds prior to 1987 is sparse, thus for the longest set of data we have on a broad sample of the highest risk debt, from 4/1987-7/2006, the average annual rate of return on the Merrill Lynch High Master II Index was only 0.77% higher than for the BBB-AA Index rated debt of similar maturity.\(^{12}\) The high yield index had a higher annualized volatility (6.7% vs. 4.8%), higher Beta with the S&P500 (0.24 vs. 0.08), and much higher transaction costs (bid-ask spreads were often 5% or more of the face value for High Yield Bonds over that period, as opposed to less than 1% for Investment grade bonds. As of August 2006 the 10-year U.S. B-rated corporates is about 270 basis points higher in yield than the comparable BBB-rated Corporate. Moody’s (2006) estimates an expected annual loss from default of 3.6% for B-rated bonds, 0.11% for BBB bonds given their expected default rates and loss-in-event of default. Going forward, the expected loss for B-rated bonds is below the current yield

\(^{12}\)Tickers H0A0 and C0C0, from Merrill (2000)
premium to the higher grade BBB-AA bonds, suggesting a negative premium to patently higher risk bonds.

Monthly returns to the 10-year bond and the US market from February 1962-July 2006 generated a beta of 0.36, and for the modern era (post 1982) a beta of 0.26, small but clearly positive.\textsuperscript{13} Further, long term bonds contain significant inflation risk because their price goes down when inflation goes up, surely a significant risk for an investor in a multifactor world (see Piazzesi and Schneider (2006)). Also, as a large investment class (US Treasury debt outstanding is greater than $3 trillion), these assets should contain some risk as they are a large portion of aggregate wealth. Thus long-term bonds appear conventionally risky, in the sense that they generate a significant wealth fluctuation that is nondiversifiable and positively correlated with the ‘true market portfolio’. A return increasing with a bond’s duration as compensation for risk would be expected in the standard risk-return theory, yet historically bond returns are not linearly increasing with duration. While there is some lift in returns from overnight to 2 years in maturity, Ilmanen (1996) finds that yield curve has produced no return premium for longer durations. A standard demonstration of the power of the risk-return ‘law’ is a comparison of the annualized volatility and returns for T-bills, Treasury bonds, and equities, which nicely shows increasing returns and volatilities.\textsuperscript{14} However, this simple demonstration conveniently omits the fact that the volatility-return trade-off that appears so clearly between the short and long end takes place for only a small portion of the yield curve.

Confounding data appear in a variety of obvious and not so obvious venues. Malkiel (1995) finds that a fund manager’s beta is uncorrelated with his fund’s average return. Commodities seem to offer a risk premium, but only if one looks at the indices with a heavy

\textsuperscript{13} Market data from Ken French, Bond Data from the Federal Reserve’s H.15.
energy weighting (Gorton and Rouwenhorst (2005)), as opposed to those that favor agriculture (Erb and Harvey (2005)). Lotteries with the lowest odds and largest jackpots (higher variance) generate more revenues, are more in demand, than higher odds, lower jackpots (Garrett and Sobel (2004)). DeVany (2003) documents that from 1986-99 Hollywood produced more than 1,000 R-rated movies, while producing a mere 60 G-rated movies. The G-rated movies generated lower volatility and higher returns to movie makers. In sports books there is the favorite-long shot bias, where favorites with low payoffs (low volatility) have higher expected returns than the higher volatility long shots, (Cain, Law, and Peel (2000)).

The more you look at the data, the more you have to think that risk, as measured by nondiversifiable volatility, is not positively and linearly related to risk, even as a tendency. If anything, the data suggest the opposite, especially for idiosyncratic risk (which as mentioned above, is a potential metric for ‘dumb money’ demand). A good theory should be consistent with this stylized fact, and consistent with intuition about disliking ‘risk’.

B. *Is the Risk Premium Zero?*

One of the main empirical implications of the theory presented in this paper is that the expected return on the market should be the risk free rate. That may seem improbable given that the ‘equity premium puzzle’ is that this premium is too large for existing models. Mehra and Prescott (1985) find that for any value of the preference parameters that generate an expected real return to Treasury bills of less than 4% (which is usually the case), the risk premium is less than 0.35%, which given the standard error of these estimates, is empirically indistinguishable from 0%. However much hope has been not so much that the equity

---

14 Malkiel (2003)
premium estimates will fall—though that would be helpful—but that a new modification within the standard framework (e.g., Epstein-Zin preferences) would cause theoretical risk premiums to rise. Certainly modifications to this model, such as shorting frictions or heterogeneous status groups that people assign themselves to, may allow this approach to create a risk premium greater than zero as well.

Nonetheless, it is worth noting that the current estimates of the risk premium are falling over time. The original Mehra-Prescott (1985) equity risk-premium estimate was 6.2%. The 1996 edition of the popular Brealey and Meyers finance textbook presented an 8.4% risk premium, and an influential Ibbotson (2000) estimate based on the U.S. from 1926-97 was 8.9%. More recently, Welch (2001) surveyed academic financial economists and found the average equity premium was 5.5%, down from 7.0% in 1998. In 2002 the AIMR sponsored a forum on the equity risk premium with several well-known researchers in this issue, and those offering a forward-looking premium estimated the following: Ibbotson (4.0%), Cornell (2.5%), Siegel (2.0), Williamson (2.0%), Campbell (1.5%), and Arnott (0%). A forward looking mean of 3.0% is a reasonable inference from this group of specialists. This is in line with dividend and earnings growth models that generate estimates of around 3.5% (see Fama-French (2002) and Blanchard (1993)), and consistent with a new historical estimate from Dimson, Marsh and Staunton (2006) that corrects for many technical issues.

After falling from 6.2%-8.4%, to 3.5% over a decade, reaching 0% is not farfetched, and there are several independent avenues that may lead this adjustment. For example,

---

15 Interestingly, these same finance economists claimed they tended to decrease their estimates after above average returns in the market, when in this instance they did the opposite.
16 Of course, there were qualifications making these estimates typically academic, qualified in various ways. See AIMR (2002).
18 This reminds one of the remark attributed to Albert Einstein: If the facts don’t fit the theory, change the facts.
most estimates of the equity subtracted a short-term bond return from a long-term equity return, but one could argue the long bond return is more appropriate because equities, like bonds, are long maturity instruments. If there is a substantial cash-like quality to bills that generates benefits not found in bonds or equities, bonds are probably better for computing the risk premium. Several estimates of the difference between short and long term bonds put the estimated mean return difference at around 1.0% (see Dimson et al (2006), Siegel (2002)), whereas Ibbotson (2006) estimates a 1.8% differential based on 1927-2005 data.\textsuperscript{19}


A final additional cost that may be unappreciated is transaction costs, which include commissions, bid-ask spread, and trade impact. There is a tradeoff whereby small traders with low trade impact generally have high commissions, cross the bid-ask spread to transact, and do not have sufficient experience or scale trading to do so in an optimal manner (algorithmic trading optimizes the size, frequency and limit order specifications, currently done thru computer-based execution of equity orders via direct market-access channels, Domowitz and Yegerman (2005)), while large institutions generate small commissions and trade based on strategies optimized through extensive trial-and-error, but in turn have a larger trade-impact due to their larger size. As brokers optimize over these three costs in a relationship, they must be considered simultaneously to get at the true trading costs.\textsuperscript{20}

\textsuperscript{19} Tax-exempt debt returns during the Gold Standard (1880-1934) were close to 4% in the US (Seigel, Homey and Sylla (1991)). Subsequent to that there periods during 1940-55 that regulated debt markets in terms of what sort of rates were applicable, and what kind of entities could hold various types of debt.

\textsuperscript{20} As the AMEX Broker/Dealer equity index has risen at over twice the annualized rate of the S&P500 since it began tracking in 1993 (XBD Index), and the price of a NYSE seat has also risen at a faster rate than the S&P500 since 1993, brokers have figured out a way to squeeze money out of their trading flow even as both
top end, Barber and Odean (2001) finds the top quintile of brokerage accounts generate 6% annual drag in expenses through excess trading (and that excludes full service brokers, or funds that charge 5% loads). Another way of triangulating this expense is the fact that the average mutual fund underperforms by 2.0% according to Malkiel (2004), which assuming no aggregate skill, is a reflection of total costs. Finally there are funds with low turnover and management fees than annualize around a 0.2% expense (e.g., Vanguard 500). It is not clear whether the marginal investor affecting equilibrium over the past 100 years is patient and savvy (0.2%), average (2%), or impatient and naïve (6%).

C. Further implications of this approach

The main implication of this model is that assets have the same expected return. However, the value of a model is its ability to explain not merely an outstanding problem (the empirical failure of beta), but seemingly diverse phenomena, and I will discuss some of them here. First, this model consistent with prior research on utility functions that look at status mentioned above, and therefore helps validate those approaches to those various problems. For example, Harbaugh and Kornienko (2001) derive a model based on assuming that people derive utility from their local status, just as assumed here. They use their model to explain why people are risk averse in gains and risk loving in losses, as an alternative to assuming a kinked utility function as in Prospect Theory, a deeper explanation if one can accept their premise that people are maximizing status.

In a world of status orientation, deviations are necessarily due to perceived superior absolute returns because in general there is no return for risk of any sort. In the CAPM average commissions and bid-ask spreads have fallen by about half over this period (see Jones (2002)), a period that included the reappearance of odd-eights quotes for NASDAQ in 1994, and introduction of decimalization in 2002.
world there should be some if not many situations where a $\beta = 0.5$ stock will be expected to outperform on a risk-adjusted basis, but not an absolute basis, yet it is extremely rare for a brokerage to tout a ‘buy’ or ‘strong buy’ recommendation that has a lower-than-average market return (see Bradshaw (2002), who finds that ‘buy’ and ‘strong buy’ recommendations from brokerages have an average one-year expected return of 25% and 34%, respectively). The fact that these recommendations are absent suggests that people only take risk, defined here as disproportionate positions relative to the consensus, if they expect an absolute greater-than-average return.

Goetzmann and Kumar (2005) document extreme underdiversification among investors using more than 40,000 equity investment accounts. More than 25% of investor portfolios contain only 1 stock, more than 50% of them contain fewer than 3 stocks. This is massively irrational diversification in the CAPM context, but as the authors note in the paper, a significant group of investors in our sample believe that they can earn superior performance. (Barber and Odean (2000) study investment clubs and find they show a distinct preference towards more volatile stocks). This is exactly what we should see if people make risky investments only when they perceive they have excess-return stocks, as misguided as this may be on average. The statistical inefficiency of these picks makes sense to these undiversified investors because in general one can rationalize overconfident insights for only a limited number of situations. Further, given the well-known objective difficulty in forecasting the risk premium (Goyal and Welch (2005), Campbell and Shiller (2001)), or market factors (Cremers and Petajisto (2006)), we should expect more risk taking in the direction of idiosyncratic risks as opposed to nondiversifiable risks, because these risks includes a greater number of parochial situations that are less competitive, and more uncertain (as opposed to risky as defined by Knight (1919)).
As every risky position implies an equal market weighting of its opposite (section II.D), it is not obvious whether disproportionate ownership is related to prescience or ignorance. One candidate for a characteristic that would tend to attract more ‘dumb money’ would be volatility. Volatility has attraction for risk takers because a generic overconfident active investor would see more value in more volatile stocks simply because if one believes they have an investment edge, for any reason, more volatile stocks present greater rewards (e.g., if they can pick stocks that go up, pick the one that will go up 40%, not the one who’s upside is 20%). Since attention is a limited resource for any investor, volatile stocks provide greater value to someone who must choose which stocks to analyze. It is also likely that volatile companies are more likely to be considered by overconfident investors looking to apply their prescience—big movers make the news—which also lowers cost of information for the time-constrained investor looking for unique situations (Barber and Odean (2005)).

If overconfidence is a primarily delusional, assets attractive to the merely overconfident, such as idiosyncratic volatility, should be negative correlated with expected returns. This could explain the low returns to idiosyncratic risk discussed above.

If people see their peers taking on more conventionally risky assets, these assets become, in this model, less risky. In a dynamic setting such a movement is unstable as supply may not be able to provide assets with sufficient returns indefinitely. A utility function that looks at relative performance may create an endogenously unstable market, relevant to models of financial panics as in Minsky (1975) and excess volatility (Shiller (1981)).

This model is consistent with the disposition effect, the finding that people generally have an unwillingness to sell losers versus winners. Shefrin and Statman (1984) argue that if a stock pick has gone down one regrets the investment, and in hoping that the stock price
will rise in the next period and thereby avoid regret, hold the stock. A stock that has gone up, however, generates pride, and holding it risks turning it into a loser, so winners are cashed out. That interpretation is consistent with risk-taking in this model, because for someone taking a risk, a position outside the consensus, they see their bet as not only an investment, but a reflection on their judgment. As such, losers are, hopefully, merely winners in gestation, while winners that have been closed are forever proof of good judgment.

The benchmark for investors may depend on their relevant peer group, with jet-setting wealthy investors comparing themselves to a world-wide return, while most investors are more interested in those assets common to their countrymen. Thus this approach is consistent with the home bias in portfolio holdings (Coval and Moskowitz (1999)), as people are trying to stay even with their countrymen, not the world. A concern for local status, as opposed to global status, is a minor modification of this approach.

IV. Conclusion

The implications of the CAPM and its extensions are logical, powerful, and slightly surprising, things one wants in a theory, but if we are to apply the principal that empirical data is the test of all theory, we have given the CAPM well too many excuses. The anomalies to the CAPM are not exceptions to a general tendency: there is no general tendency. The standard utility function of the CAPM may be a good normative theory, but a preference for status generates a more accurate positive theory, and its assumption—caring only about relative status—is at least as plausible as assuming people care only about absolute wealth.
A relative-status oriented utility function generates a factor model consistent with the familiar CAPM and APT model, except the risk premiums are zero. Beta is still descriptive of relative volatility, and generates normative predictions for volatility minimization. However, there is no cross-sectional return to $\beta$, no positively-sloped Security Market Line. The portfolio optimization algorithm for a trader without any special edge is trivial: allocate assets to the standard categories of conventional wisdom, because this minimizes relative wealth volatility and maximizes returns. The idea that nondiversifiable risk is not priced because it is unnecessary is augmented, so that diversifiable risk is also not priced because it too is unnecessary in the sense of avoiding deviations from the average portfolio. One has to take risk to generate return, as the saying goes, but there is no general relationship, no general expectation of higher expected returns merely for being bold.

A normative recommendation to investors with these preferences would be to adhere to average portfolio weights in stocks, bonds, home mortgage, and perhaps equity styles (e.g., growth stocks), and to allocate, cautiously, a disproportionately higher allocation to situations where one has good reason to believe they have an edge (the Peter Lynch recommendation to ‘invest in what you know’, see Lynch (2000)). Deviance from a market-weighted consensus may define risk, but many people think the conventional wisdom is wrong on more than one dimension, and they are probably right. Yet merely deviating from conventional wisdom produces no necessary reward any more than taking a risk in any aspect of one’s life generally produces reward on average, and it seems reasonable that for every risk-taker confident because he is smart, there is a risk-taker confident because he is merely deluded. In light of the severe underdiversification of most portfolios, the higher expected returns that appear to drive them, and the common method of allocating assets to standard categories (e.g., cash, bonds, stocks), this seems more realistic description of the
investing process in practice than the CAPM assumption that people invest based on a target beta between their household wealth and the market.
References


Cremers, Martin, and Antti Petajisto, 2006, How Active is your Fund Manager? A New Measure that Predicts Performance, Working Paper.


Harbaugh, Rick, and Titiana Kornienko, 2001, Local Status and prospect theory, Claremont Colleges working paper.


Lynch, Peter and John Rothchild, 2000, One Up on Wall Street, Simon and Schuster.


Kim, Kun-Ok, 1993, What is Behind Face-Saving in Cross-Cultural Communication? Intercultural Communications Studies III:1

Knight, Frank, 1921, Risk, Uncertainty, and Profit, Hart, Schaffner & Marx.


Sobel, Joel, 2005, Interdependent Preferences and Reciprocity, *Journal of Economic Literature*, 43, 392-436


