Jensen’s Inequality, Parameter Uncertainty, and Multi-period Investment

Mark Grinblatt
UCLA Anderson School of Management

Juhani T. Linnainmaa
University of Chicago, Booth School of Business

Classical approaches to estimation and decisions requiring estimation often are at odds. When values critical to the decision are convex or concave functions of unknown parameters, the statistician’s estimation error adjustments are the opposite of what is appropriate for the decision. We illustrate the conflict by studying multi-period investment problems. The proper application of Jensen’s inequality to the decision turns finance intuition on its head: Multi-period investments with negative risk premia can be profitable, risk-averse investors can have infinite demand for risky securities, settings exist in which risk-averse investors should not diversify, and demand for mutual funds with negative alphas may be rational. (JEL C11, G11)

Jensen’s inequality—the statement that \( E[V(\tilde{u})] \) is below \( V(E[\tilde{u}]) \) if \( V() \) is concave (or above \( V(E[\tilde{u}]) \) if convex)—is one of the most ubiquitous mathematical tools of financial economics. Investors prefer riskless to otherwise identical risky investments because a concave function of a random variable’s expected value exceeds the expectation of the function. Geometric means are always below arithmetic means because multi-period returns are convex functions of single-period returns. And because options are convex payoffs of underlying future values, corporate managers can transfer wealth from debt to equity holders by taking on risky projects.

All of these applications of Jensen’s inequality are well known. Less familiar are the ramifications of Jensen’s inequality for estimation. Virtually every practical problem studied in finance requires some nonlinear estimation. Modern investment theory and asset pricing take place in dynamic settings, which inherently involve compounding. Modern corporate finance, whether...
focused on capital structure, investment scale, or optimal project selection, uses concave production functions, convex options, or convex payoff functions associated with bankruptcy and taxes to analyze most of its fundamental issues.

We show that Jensen’s inequality almost always generates conflicts between the estimates of a financial decision maker, referred to as an “investor,” and estimates implemented with a classical approach, offered by an agent we refer to as a “statistician.” The Jensen’s inequality conflict is layered upon the well-known existing distinction between Bayesian and classical estimation. If the investor and statistician use the same information and hold concordant point estimates of a parameter, their estimates of any linear function of the parameter will also agree. However, their estimates of a concave or convex function of the parameter will account for the nonlinearity by adjusting the function’s estimate in opposite and conflicting directions.

A two-period example illustrates the conflict. Consider returns \( r_1 \) and \( r_2 \) generated by

\[
r_t = \mu + \sigma z_t. \tag{1}
\]

In Equation (1), the mean of the return is a constant \( \mu \), the standard deviation is \( \sigma \), and the correlation between \( z_1 \) and \( z_2 \) is zero. Under this setup, a dollar is expected to grow to

\[
\text{Future Value} = E[(1 + r_1)(1 + r_2)] = (1 + \mu)^2 \tag{2}
\]

after two periods. The sample counterpart to this equation,

\[
\text{Future Value} = E\left[(1 + m)^2\right] - \text{var}(m), \tag{3}
\]

where an unbiased estimate of \( \mu \), denoted \( m \), replaces \( \mu \), reflects a well-known result: Compounding the estimated arithmetic mean return upwardly biases the long-term return estimate.

Sample estimates of Equation (3) are commonly used to make investment decisions. Adjusting the \( E[(1 + m)^2] \) forecast for its bias by subtracting an unbiased estimate of \( \text{var}(m) \) in Equation (3) generates an unbiased forecast of the investment’s future value two periods hence. For risk-neutral investors, an investment criterion based on the sign of the difference between an unbiased estimate of the two-period expected future value and the future value obtained with risk-free investment, e.g.,

\[
[(1 + m)^2 - \text{var}(m)] - (1 + r_f)^2, \tag{4}
\]

seems eminently plausible. Alternatively, research recommends that we replace \( m \) with a lower number to eliminate the \( \text{var}(m) \) term in Equation (4). Introductory finance textbooks have also espoused the Blume (1974), Fama
Jensen’s Inequality, Parameter Uncertainty, and Multi-period Investment (1996), and Jacquier, Kane, and Marcus (2003) recommendation—take weighted averages of estimated arithmetic and (always lower) geometric means.¹

Surprisingly, the investment criterion in Equation (4) is wrong. Rather than reducing the forecast of future value, \((1 + m)^2\), we should increase it! To explain why, we derive Equation (3) from (2):

\[
(1 + \mu)^2 = \mathbb{E}\left[ (1 + m) - (m - \mu) \right]^2 \\
= \mathbb{E}\left[ (1 + m)^2 \right] + \text{var}(m - \mu) - 2\text{cov}(m, m - \mu). \tag{5}
\]

Up to this point, conditioning notation has been deliberately imprecise but is critical for understanding our point. The sign of the sum of the last two terms in Equation (5) depends on what we condition on. Conditioning on the true value of \(\mu\), Equation (5)’s future value is

\[
\mathbb{E}\left[ (1 + m)^2 \mid \mu \right] + \text{var}(m - \mu \mid \mu) - 2\text{cov}(m, m - \mu \mid \mu) \\
= \mathbb{E}\left[ (1 + m)^2 \mid \mu \right] - \text{var}(m \mid \mu), \tag{6}
\]

which is Equation (3). It becomes apparent that our analysis here is based on a classical statistician’s view of the multi-period expected future value: The true value of the parameter \(\mu\) is fixed, whereas the estimator \(m\) is a random variable. However, the decision to invest hinges on the investor’s perspective. Instead of conditioning on \(\mu\), which the investor does not know, we must condition on \(m\), which he does know.² The investor believes \(\mu\) has a distribution (centered around his estimate \(m\) of \(\mu\)’s mean). Here, the expected future value is

\[
\mathbb{E}\left[ (1 + m)^2 \mid m \right] + \text{var}(m - \mu \mid m) - 2\text{cov}(m, m - \mu \mid m) \\
= (1 + m)^2 + \text{var}(\mu \mid m). \tag{7}
\]

Equation (7) has a more Bayesian flavor than Equation (6) but does not have to be the outcome of a Bayesian updating process. Indeed, a Bayesian and a classical statistician will have different estimates in general, depending, among other things, on the Bayesian’s prior for \(\mu\). The critical point here is that conditioning on what the investor knows leads to a very different investment criterion based on the sign of

\[
(1 + m)^2 + \text{var}(\mu \mid m) - (1 + rf)^2. \tag{8}
\]

Equations (6) and (7) and their associated investment criteria illustrate two conflicting perspectives. In this example, both the classical statistician’s and

¹ See, for example, Grinblatt and Titman (2002, 419–21) and Bodie, Kane, and Marcus (2009, 147).

² Zellner and Chetty (1965) and Klein and Bawa (1977) first made this point.
the investor’s estimate of $\mu$ is $m$. Yet one believes $(1 + m)^2$ underestimates the multi-period future value and the other believes it overestimates the value. The investor concludes that adjusting the arithmetic mean toward the (always lower) geometric mean, as the literature suggests, generates even poorer investment decisions than not making any adjustment at all. If anything, the investor would like to artificially increase the arithmetic mean estimate. As a consequence, a risk-neutral investor would adopt multi-period projects that are estimated to have modest negative risk premia.

Why do we reach seemingly different conclusions, depending on whether we view truth as fixed with the estimate stochastic or the estimate as fixed with truth being one of many possible outcomes? After two periods, misestimates of the one-period population mean generate a squared (and therefore positive) random variable. Expectations of the squared number increase either the multi-period forecast (the statistician’s view) or the mean of the multi-period distribution being forecasted (the investor’s view). Even when the single-period return forecast of the investor and the statistician are identical, and therefore generate the same estimation error, it is perfectly sensible for the statistician to reduce his multi-period forecast and for the investor to increase his multi-period forecast.

We illustrate the point more generally in Figure 1, which heuristically portrays uncertainty by focusing on two outcomes for a parameter estimate (Panel A) or draw of nature (Panel B). Each of the two outcomes is labeled on the $x$-axis. The investment value $V$, a convex function, is shown on the $y$-axis. Panel A presents the statistician’s view of the deviation of the estimate from truth while the almost identical graph in Panel B portrays the investor’s view. The estimate $V(m_{\text{high}})$ overshoots $V(\mu)$ from the statistician’s perspective in Panel A. From the investor’s perspective, this corresponds to $V(\mu_{\text{low}})$ undershooting $V(m)$ in Panel B. Conversely, the estimates of $V()$ at $m_{\text{low}}$ and $\mu_{\text{high}}$ correspond to each other. In both panels, the average of the high and low outcomes of $V$ exceeds the value of the convex function in the middle. From the statistician’s perspective, the average estimate of $V$ exceeds the target. For the investor, the average of the $V$s at nature’s high and low values, $E[V(\mu)]$, is the target. This average exceeds the value of the convex function at $m$.

Remedies for the effect of convexity in Panels A and B would have to move the investment’s estimated value in opposite directions. The statistician, not knowing whether he is at Panel A’s high or low value of $m$, would lower his estimate of $V$ below $V(m)$. By contrast, the investor, not knowing if nature is at Panel B’s high or low value of $\mu$, would forecast the investment’s value $V$ to be above $V(m)$.

Figure 1 illustrates that the conflict between these points of view may apply to more than multi-period investment problems. Indeed, we formally prove that the conflict exists under quite general assumptions. However, multi-period investment, which we use as our quintessential example, offers some rather
remarkable insights into the proper perspective—the investor’s—while providing an easy platform for calibrating the distinction between the statistician’s and investor’s views.

What motivates the statistician and the investor to view the same problem differently? The statistician is motivated by the need to appear neither too high

---

**Figure 1**
**Parameter uncertainty and Jensen’s inequality: Statistician’s view versus investor’s view**
This figure categorizes uncertainty by two outcomes for a parameter estimate (Panel A) or draw of nature (Panel B) and a convex function of those parameters, which is the investment value $V$. Panel A presents the statistician’s view of the deviation of the estimate from truth, while Panel B portrays the investor’s view. The overshooting of the estimate in Panel A for the statistician, denoted with an uppercase “S,” corresponds to the undershooting of nature’s draw in Panel B for the investor, denoted by a lowercase “i.” Conversely, the undershooting of the estimate in Panel A for the statistician, denoted with a lowercase “s,” corresponds to the overshooting of nature’s draw in Panel B for the investor, denoted by an uppercase “I.”
nor too low across many independent forecasts. However, concerns about one’s long-term reputation are irrelevant for the investor. The investor cares about the distribution of the one true outcome for $\mu$ because the consequences of his decision depend on that outcome. He necessarily views $\mu$ as being drawn from a distribution, and his expected value for $V(\mu)$ is estimated from that conditional distribution.

It is well known that these two views can generate different predictions. Prior research also contains examples where classical and Bayesian statistical estimates differ for reasons other than differences in information. Sims and Uhlig (1991) observe that in time-series models with unit roots, it is not true, even asymptotically, that “…Bayesian probability statements about the unknown parameters conditional on the data are very similar to classical confidence statements about the probability of random intervals covering the true value of the parameter.” Stambaugh (1999) points out that with stochastic regressors, the two perspectives’ predictive slope estimates tend to diverge in small samples. Kandel and Stambaugh (1996) highlight that predictors can strongly influence investment even when they exhibit a weak statistical relationship with returns. Pástor and Stambaugh (2009) emphasize the differing views of Bayesians and classicists when analyzing long-horizon volatility. However, no paper in this line of research explores the conflict in the two views induced by Jensen’s inequality.

There also is an extensive literature on parameter uncertainty and dynamic portfolio choice. This literature focuses on optimizations that correctly condition on the investor’s forecast. However, the literature’s tendency to emphasize intertemporal hedging demands and investors who are more risk averse than log investors obscures the role of Jensen’s inequality when estimating a concave or convex $V(\mu)$.

Other research implicitly or explicitly recognizes that compounding of persistent risk increases expected long-term payoffs. Baron (1974, 559–61), conditioning on $m$, presents a two-period investment with two equally likely binomial trees. Because Baron’s investor maximizes the square root of wealth after two periods, he rejects the project in the absence of expansion options. Baron concludes that it is the option to expand investment after observing a good first-period outcome that makes the project an attractive investment at the outset. What he fails to recognize is that with lower risk aversion, the project is more attractive than zero-interest risk-free investing, even without expansion triggered by learning.


4 In Baron’s “good” binomial tree, the return along a single period’s segment of the tree is 95% with probability 0.8, and −100% with probability 0.2. In the “bad” binomial tree, the 95% return outcome occurs with probability 0.2. On average, investment in the project converts $1 into $1.29 after two periods, while losing $0.025 in the initial period. This trade-off is clearly attractive to risk-neutral investors.
McCardle and Winkler (1989, 1992) generalize Baron’s binomial example to longer investment horizons. They correctly point out that learning is less relevant than the compounding of persistent risk in converting bad short-term investments into good long-term ones. They consider their finding an anomaly that arises from the zero lower bound on utility for isoelastic utility functions with risk aversion below one. However, risk-neutral investors, who have no such constraint, may also find such investments to be attractive.

Despite challenges from these remarkable findings, research seems to either be mute, unaware, or confused about how to properly think about the conflict between the differing inference methodologies. For example, Blume (1974), conditioning on μ, recommends using weighted averages of arithmetic and (always lower) geometric means to estimate long-term returns. The academic literature also espouses both the statistician’s and investor’s views within the same article, as if they were not in conflict. Fama (1996, 418) points out that persistent risk, generated by a time-varying market risk premium, increases the discount rate of long-term payoffs. A page later, the paper notes that another persistent risk, market risk premium estimation error, calls for a reduction in the discount rate of long-term payoffs, achieved by weighting geometric and arithmetic averages as in Blume (1974). Jacquier, Kane, and Marcus (2005) compute both the investor’s and statistician’s views of a multi-period future value, but never emphasize that adjustments in the two settings are in opposite directions.

Pástor and Veronesi (2003) elegantly illustrate how uncertainty about the growth rate of a firm’s book value increases firm value in a multi-period finite horizon setting—an outcome that clearly stems from Jensen’s inequality. However, even this article may not fully appreciate all of the ramifications of Jensen’s inequality. We show that parameter uncertainty can make the expected utility of some investors convex in portfolio holdings, resulting in infinite utility and demand for risky assets with return parameters for which estimation is sufficiently imprecise. In such settings, the exogenously specified pricing kernels used to generate firm value may not exist without strong restrictions on preferences.

Ultimately, our article is about how finance intuition is often turned on its head when properly estimating concave and convex functions for decisions. Section 1 proves the general result that Jensen’s inequality makes statisticians and investors adjust forecasts of convex (or concave) functions in opposite ways. Section 2 examines the role of Jensen’s inequality in multi-period investing. Section 3 concludes the article.

1. The Theory

Consider a nonlinear function \( V() \) of an unknown parameter \( \mu \). The sequence of events is that nature first selects \( \mu \). An agent, either an investor (denoted by sub- or superscript “I”) or a statistician (labeled “S”), then gathers \( \mu \)-generated
information, usually historical data, and forms an information set.\(^5\) He subsequently estimates expected values for \(\mu\), denoted \(m_I\) or \(m_S\), as well as expected values for \(V(\mu)\), denoted \(E^I[V(\mu)]\) or \(E^S[V(\mu)]\) (where the latter expectation is the mean of the statistician’s confidence interval).

We focus on expectations because the investor usually forms these to decide which action to take. Examples include expected payoffs, discounted expected cash flow streams, expected utility, and risk-neutral expectations that assess fair value. Of particular interest is whether the investor perceives \(V(m_I)\) to be higher or lower than the benchmark \(E^I[V(\mu)]\). The statistician’s analogous benchmark, \(E^S[V(\mu)]\), is his unbiased estimate of \(V(\mu)\). Here, we ignore other estimation criteria like precision in order to compare the two agents’ estimates of \(E[V(\mu)]\).\(^6\)

The two agents have differing objectives. The investor makes a decision in light of nature’s lone draw of \(\mu\). Given his information, he estimates that draw to have a distribution with a mean \(m_I\). The consequences of his decision in that one case are a lottery with a distribution determined by \(\mu\)’s distribution. By contrast, the consequences of repeated estimation problems are different for the statistician. Each new estimation problem he faces has a different and independent draw of \(\mu\), a different set of data associated with that draw of \(\mu\), and a different \(m_S\). His concern for a reputation of unbiasedness calls for a different estimation philosophy—one that uses the distribution of data conditional on \(\mu\) to construct an unbiased estimate of \(V(\mu)\). Adopting this approach eliminates any tendency to over- or underestimate over multiple independent estimation problems.

If \(V(\mu)\) is linear in \(\mu\), the differing objectives of investors and statisticians do not generate estimation differences. Both the investor and the statistician estimate \(E[V(\mu)]\) as the same linear function of their estimates \(m_I\) and \(m_S\). Thus, their estimates of \(V(\mu)\) differ only when \(m_I\) and \(m_S\) differ (which can arise only from differences in the agents’ information sets, not from differing objectives). Put another way, if \(m_I\) equals \(m_S\), producing identical forecast errors, the investor’s distribution of the linear \(V\) is identical to the statistician’s distribution. However, when \(V\) is nonlinear, conditioning on \(m\) and conditioning on \(\mu\) lead to very different estimates of \(V\), even when \(m_I\) equals \(m_S\).

Denote the investor’s \(\mu\) forecast error as \(u_I = \mu - m_I\). Consistent with the discussion above, we assume the investor views \(u_I\) as perfectly correlated with nature’s draw of \(\mu\) and \(E^I[u_I | m_I] = 0\). Denote the statistician’s forecast error

---

\(^5\) The information set can combine the realized data generated by \(\mu\) with economic intuition, priors, and even irrational hunches.

\(^6\) The reader who is more familiar with statistical estimation can take comfort that quadratic loss functions are special cases of our approach. However, decision making in economics and finance generally uses expectations in lieu of loss functions. There are many ways of arriving at the same expectation without formally modeling how inference rules map data into a loss function.
as \( u_s = \mu - m_s \). The statistician views \( u_s \) as being perfectly correlated with \( m_s \) and \( E^S[\hat{u}_s | \mu] = 0 \). Our goal is to study each agent’s view of \( V(m) \) as a (flawed) estimate of \( E[V(\mu)] \). Proposition 1 shows that the investor’s and statistician’s expectations of \( V(m) \)’s forecast error are of opposite sign.

**Proposition 1.** If \( V() \) is convex,

\[
E^I[V(\mu) - V(m_I) | m_I] > 0 > E^S[V(\mu) - V(m_S) | \mu],
\]

and if \( V() \) is concave,

\[
E^I[V(\mu) - V(m_I) | m_I] < 0 < E^S[V(\mu) - V(m_S) | \mu].
\]

**Proof.** Note that \( E^I[V(\mu) - V(m_I) | m_I] = E^I[V(m_I + \hat{u}_I) - V(m_I) | m_I] \) and \( E^S[V(\mu) - V(m_S) | \mu] = E^S[V(\mu) - V(\mu - \hat{u}_S) | \mu] \), where \( E^I[\hat{u}_I | m_I] = E^S[\hat{u}_S | \mu] = 0 \). By Jensen’s inequality, \( E^I[V(m_I + \hat{u}_I) - V(m_I) | m_I] \) and \(-E^S[V(\mu) - V(\mu - \hat{u}_S) | \mu] \) are positive for convex \( V() \) and negative for concave \( V() \).

Proposition 1 tells us the sign of the statistician’s bias from estimating \( V(\mu) \) as \( V(m_S) \). However, the precise magnitude of the bias could vary with \( \mu \). In the introduction’s investment example, the statistician’s expected estimation error is independent of \( \mu \). If the source of the nonlinearity is not a second-degree polynomial (as in the introduction’s two-period return example), \( V(m) \)’s nonlinearity generates expected forecast errors that depend on \( \mu \).

Fortunately, one also avoids addressing \( \mu \)-dependent expected forecast errors when \( V() \) is an exponential function of \( \mu \) (making the estimate exponential in \( u \)) and the scaled forecast error \( u \) is multivariate normal. In this case, the percentage expected estimation error does not depend on \( \mu \) or \( m \), and there is a multiplicative adjustment to \( V(m) \) that corrects for the bias induced by Jensen’s inequality. Here, the investor’s adjustment factor, \( e^{\Delta v_I} \), solves

\[
e^{\Delta v_I} V(m) = E^I[V(m_I + \hat{u}_I) | m_I]
\]

and the statistician’s adjustment factor, \( e^{\Delta v_S} \), solves

\[
e^{\Delta v_S} E^S[V(\mu - \hat{u}_S) | \mu] = V(\mu).
\]

In this case, with \( V(u) = C_0e^{C_1u+y} \), where the normally distributed \( y \) is uncorrelated with \( \mu \)’s forecast error, the multiplicative adjustments have exponents \( \Delta v_I \) and \( \Delta v_S \) that do not depend on \( m, \mu \), or \( C_0 \) (which determines the convexity or concavity of \( V \) by its sign). Specifically, \( \Delta v_I = \frac{1}{2}C_1^2\text{var}(\hat{u}_I) \) and \( \Delta v_S = -\frac{1}{2}C_1^2\text{var}(\hat{u}_S) \). This article quantifies Jensen’s inequality’s effect on
estimates of the expected utility of terminal wealth using this tractable class of cases.\textsuperscript{7}

In short, Jensen’s inequality forces the investor and the statistician to adjust a $V(m)$ estimate of a concave or convex function $V(\mu)$ in opposite directions. This implies that the statistician’s view is rarely appropriate for decisions that hinge on nonlinear functions of unknown parameters. To better illustrate how misleading the statistician’s perspective is, we now analyze multi-period investment, a key area of finance in which convexity plays a major role. Study of this topic helps calibrate the impact of Jensen’s inequality and points to obvious flaws in decisions that would be made using the statistician’s perspective. However, the proper investor perspective also leads to a number of surprising but understandable results that the finance literature has largely ignored. These include the desirability of investment with negative risk premia or fund management with negative alphas, infinite demand for risky assets by risk-averse investors, and higher utility from not diversifying.

2. Applications to Multi-period Investing

This section develops a more realistic model of returns and preferences to quantify the impact of Jensen’s inequality on multi-period investment. It shows that an enormous gap exists between the statistician’s and the investor’s assessments of long-term projects.

An “apples-to-apples” comparison that isolates the effect of $V(\cdot)$’s nonlinearity on estimates of $V(\cdot)$ using the two approaches requires the statistician and the investor to (i) have the same $\mu$ estimate, that is, $m_I = m_S = m$; and (ii) perceive identical unconditional distributions of the error from estimating $\mu$ as $m$, so that $\text{var}(\tilde{u}_I) = \text{var}(\tilde{u}_S) = \theta^2$. These requirements are unusual for those familiar with the contrast between the Bayesian and frequentist approaches to statistical estimation. The confidence interval of the frequentist usually coincides with the posterior distribution of the Bayesian only when the Bayesian has an uninformative prior.\textsuperscript{8} Although we respect a long literature that points out this difference in the two approaches, our goal is to isolate an additional difference that generates conflicting expectations of concave or convex functions. To properly calibrate the impact of the difference induced by Jensen’s

\textsuperscript{7} In these cases, $C_0 = \frac{1}{1-\gamma}$, $C_1 = (1 - \gamma)xT$, $\gamma = \text{risk aversion}$, $x = \text{risky asset investment}$, and $T = \text{investment horizon}$. In more general cases, not knowing $\mu$, the statistician could weight the forecast error attached to each realized $\mu$, $E[ V(\mu) - V(m_S) ]$, by some subjective probability to obtain an adjustment to $V(m)$ that meets his objective of being unbiased across many independent forecasts. Consider an example where $V(\mu) = \mu^4$ and $\tilde{\mu} \sim N(0, 1)$. In this case, the statistician’s expected $\mu$-contingent forecast error is $E[ \mu^4 - (\mu - \tilde{\mu})^4 | \mu ] = -6\mu^2\theta^2 - 3\theta^4$. If the statistician specifies a density $f(\mu)$, he can achieve an unbiased estimate of $V(\mu)$ by adding $\int (6\mu^2\theta^2 + 3\theta^4)f(y)dy$ to $V(m)$. Determining the best distribution for $\mu$ is beyond the scope of this article. However, $\mu$’s conditional distribution given $m$ represents one possibility. In this case, the adjustment captured in the integral above adds $6m^2\theta^2 + 3\theta^4 = 6m^2\theta^2 + 9\theta^4$ to $V(m)$.

\textsuperscript{8} See Sims and Uhlig (1991) for a notable exception.
inequality, we will analyze a variety of finance problems where each agent is assumed to have the same estimate of $\mu$ but estimation objectives that align with either the frequentist or the Bayesian. Moreover, we do not impose rules for constructing $\mu$’s estimate. As long as the rules are consistent across the two types of agents, each can form inferences using principles from behavioral economics rather than from Bayes’ rule. For this reason, we refer to the agents as investors and (classical) statisticians rather than as Bayesians and frequentists.

2.1 Model for Calibration
To quantify the impact of Jensen’s inequality on the decision to invest in a risky multi-period project, consider an investor with power utility over $T$-year terminal wealth

$$E[U(\tilde{W}_T)] = E\left[\frac{\tilde{W}_T^{1-\gamma}}{1-\gamma}\right],$$

where $\gamma$ is the risk-aversion parameter. For $\gamma = 1$, $\ln(\tilde{W}_T)$ replaces the expression above.

The project, which the investor either adopts or rejects, has value $P$ with dynamics given by the standard geometric diffusion process

$$\frac{d\tilde{P}_t}{P_t} = (m + \tilde{u})dt + \sigma d\tilde{z}_t, \quad (13)$$

where $u = \mu - m$. Because the investor does not know $\mu$, his forecast error $u$ is a persistent random variable assumed to be normally distributed with constant variance $\theta^2$. The constant $\sigma$ parametrizes the variance from non-persistent uncertainty, $d\tilde{z}_t$. Thus, $\sigma d\tilde{z}_t$ is the deviation of each instantaneous return from $\mu dt$.

2.1.1 The Investor’s View. The project adoption decision is isomorphic to a portfolio optimization problem. The investor earns utility from investing either in the risky project or a risk-free asset. Without loss of generality, assume initial wealth $W_0 = 1$. The fraction of wealth spent on the risky project $x$ can be either 0 or 1. Letting $\tilde{z}$ denote the stochastic integral $\frac{1}{\sqrt{T}} \int_0^T d\tilde{z}_t$, which is scaled to have unit variance, Itô’s lemma implies that terminal wealth is

$$\tilde{W}_T = e^{\left[r_f + x(m + \tilde{u} - r_f - \frac{1}{2} x \sigma^2)\right]T + x \sigma \tilde{z}_t \sqrt{T}}, \quad (14)$$

9 We have verified that our findings are robust to models of utility from consumption streams rather than terminal wealth. Our insights are also robust to uncertainty about the horizon for investing. As long as the investor has utility that is a probability weighting of utilities at multiple horizons, the analysis remains largely unchanged.
where \((\tilde{u}, \tilde{z})\) are uncorrelated bivariate Normal\((0,1)\) random variables. Expected utility is a scaling of the moment-generating function of a normally distributed random variable,

\[
E \left[ \frac{\tilde{W}_{T}^{1-\gamma}}{1-\gamma} \right] = \frac{1}{1-\gamma} e^{(1-\gamma)T[r_f + (m-r_f - \frac{1}{2} x \sigma^2)x] + \frac{1}{2} (1-\gamma)^2 (T^2 \theta^2 + T \sigma^2) x^2}
\]

\[
= \frac{1}{1-\gamma} e^{(1-\gamma)T[r_f + (m-r_f)x + \frac{1}{2} ((1-\gamma)T \theta^2 - \gamma \sigma^2) x^2]}, \tag{15}
\]

where the \(T^2\) multiplying \(\theta^2\) in the exponent arises from the persistence in \(m\)'s forecast error. Expected wealth at \(T\) is Equation (15) with \(\gamma = 0\),

\[
E(\tilde{W}_T) = e^{r_f T} e^{(m-r_f)x + \frac{1}{2} T^2 \theta^2 x^2}. \tag{16}
\]

Because of the \(T^2\) term in Equations (15) and (16), expected utility and wealth from investing \((x = 1)\) exceeds that from not investing \((x = 0)\) for sufficiently large \(T\), even when the instantaneous risk premium \(m - r_f\) is negative. Indeed, as \(T\) grows, the expected wealth from investing is unbounded. The quadratic effect of \(T\), which arises because we compound parameter uncertainty, implies an investment can simultaneously have a negative risk premium over a short investment horizon and a positive risk premium when measuring returns over a longer horizon. As the horizon lengthens, the expected utility of investors who are less risk averse than log investors eventually becomes convex in parameter uncertainty.

Parameter uncertainty’s effect on utility stands in marked contrast to the effect on utility from the \(d\tilde{z}\)'s, the deviations of returns from nature’s draw of the mean return. The latter source of risk reduces expected utility, irrespective of \(\gamma\), because the \(d\tilde{z}\)'s are uncorrelated over time, whereas \(\tilde{u}\) is a persistent random variable. If each date’s \(\tilde{u}\) had been an independent draw from a distribution, rather than being identical, mistakes in forecasting \(\mu\) would never compound and thus could not offset the adverse utility effect of increased risk, even for investors with low risk aversion.

Could the quadratic horizon effect lead mildly risk-averse investors to rationally invest in risky investments with negative risk premia per period, even when risk premia are measured and estimated over periods of finite length? To investigate the issue, compute the sum of one plus the expected return from investing over a finite period of length \(\Delta t\). This sum is Equation (16), replacing \(T\) with \(\Delta t\) and setting \(x = 1\),

\[
1 + E(\tilde{r}_j) = e^{(m + \frac{1}{2} \theta^2 \Delta t) \Delta t}. \tag{17}
\]

We henceforth drop the superscript \(I\) and \(S\) from expectations, forecasts, and forecast errors, as it is clear which agent’s perspective we are taking.

Although \(x = x^2\) here, we write the equation in this form to facilitate later study of scalable investment.
Compounding this return for $\frac{T}{\Delta t}$ periods generates a value that depends on period length $\Delta t$,

$$(1 + E(\tilde{r}_j))^{\frac{T}{\Delta t}} = e^{mT + \frac{1}{2}\theta^2 T \Delta t}. \quad (18)$$

The utility from the cumulated expected gross return in Equation (18) can be factored out of the expected utility of an investor who adopts the project. The latter expected utility, Equation (15) evaluated at $x = 1$, is factored as follows:

$$E\left[\frac{W^{1-\gamma}_T}{1-\gamma}\right] = \frac{1}{1-\gamma} e^{(1-\gamma)T[m + \frac{1}{2}((1-\gamma)T \theta^2 - \gamma \sigma^2)]}$$

$$= e^{\frac{1}{2}(1-\gamma)T((1-\gamma)T - \Delta t)\theta^2 - \gamma \sigma^2)} \left[\frac{(1 + E(\tilde{r}_j))^{\frac{T}{\Delta t}(1-\gamma)}}{1-\gamma}\right]. \quad (19)$$

Thus, expected utility is decreasing in $\sigma$, but other parameters’ effects may depend on the sign of $1 - \gamma$.

Equation (19) illustrates how means and risk interact in a multi-period setting to influence expected utility. Expected utility is the product of an exponential term on the left and utility from a riskless compounded return (in brackets) on the right. A riskless investment has both $\theta$ and $\sigma$ equal to zero. Hence, a hypothetical riskless investment, earning the same expected per-period return as the risky project, would generate expected utility only from the bracketed term on the right. The sign of the exponent of the term on the left determines whether the risky project’s uncertainty, due to both $\theta$ and $\sigma$, enhances or diminishes expected utility, as compared to this hypothetical riskless benchmark. The key insight is summarized below.

**Proposition 2.** A risk-averse investor may prefer a multi-period risky investment to a multi-period riskless investment when both are expected to earn the same return per period or when the riskless investment is expected to earn a slightly larger return.

To understand Proposition 2, notice that the exponential term, the first factor in Equation (19), exceeds one when $(T - \Delta t - \gamma T)\theta^2 - \gamma \sigma^2$ is positive and is below one when it is negative. The boundary where this exponential term equals one is

$$\gamma^b = \frac{(T - \Delta t)\theta^2}{T \theta^2 + \sigma^2}. \quad (20)$$

Hence, when $\gamma < \gamma^b$, expected utility from $T$-year risky investment exceeds expected utility from risk-free investment that earns the same one-period expected return, $E(\tilde{r}_j)$, with certainty. Equation (20)’s risk-aversion parameter of indifference, $\gamma^b$, is increasing both in $T$ and $\frac{\theta}{\sigma}$ (the ratio of the uncertainty
Table 1
Certainty-equivalent Returns and Risk-aversion Parameters of Indifference (The Investor’s View)

Panel A: Certainty-equivalent Returns

<table>
<thead>
<tr>
<th>γ = 0</th>
<th>θ</th>
<th>Horizon (in Quarters)</th>
<th>1</th>
<th>4</th>
<th>12</th>
<th>40</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.00</td>
<td>4.00%</td>
<td>4.00%</td>
<td>4.00%</td>
<td>4.00%</td>
<td></td>
<td></td>
</tr>
<tr>
<td>0.04</td>
<td>4.00%</td>
<td>4.02%</td>
<td>4.06%</td>
<td>4.20%</td>
<td></td>
<td></td>
</tr>
<tr>
<td>0.08</td>
<td>4.00%</td>
<td>4.06%</td>
<td>4.23%</td>
<td>4.81%</td>
<td></td>
<td></td>
</tr>
<tr>
<td>0.12</td>
<td>4.00%</td>
<td>4.14%</td>
<td>4.52%</td>
<td>5.84%</td>
<td></td>
<td></td>
</tr>
<tr>
<td>0.16</td>
<td>4.00%</td>
<td>4.25%</td>
<td>4.92%</td>
<td>7.30%</td>
<td></td>
<td></td>
</tr>
<tr>
<td>0.20</td>
<td>4.00%</td>
<td>4.39%</td>
<td>5.44%</td>
<td>9.20%</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>γ = 1/4</th>
<th>θ</th>
<th>Horizon (in Quarters)</th>
<th>1</th>
<th>4</th>
<th>12</th>
<th>40</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.00</td>
<td>2.97%</td>
<td>2.97%</td>
<td>2.97%</td>
<td>2.97%</td>
<td></td>
<td></td>
</tr>
<tr>
<td>0.04</td>
<td>2.97%</td>
<td>2.97%</td>
<td>2.99%</td>
<td>3.06%</td>
<td></td>
<td></td>
</tr>
<tr>
<td>0.08</td>
<td>2.99%</td>
<td>2.99%</td>
<td>3.07%</td>
<td>3.36%</td>
<td></td>
<td></td>
</tr>
<tr>
<td>0.12</td>
<td>3.01%</td>
<td>3.20%</td>
<td>3.85%</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>0.16</td>
<td>3.05%</td>
<td>3.38%</td>
<td>4.54%</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>0.20</td>
<td>3.09%</td>
<td>3.61%</td>
<td>5.44%</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>γ = 2</th>
<th>θ</th>
<th>Horizon (in Quarters)</th>
<th>1</th>
<th>4</th>
<th>12</th>
<th>40</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.00</td>
<td>−0.08%</td>
<td>−0.08%</td>
<td>−0.08%</td>
<td>−0.08%</td>
<td></td>
<td></td>
</tr>
<tr>
<td>0.04</td>
<td>−0.09%</td>
<td>−0.10%</td>
<td>−0.14%</td>
<td>−0.28%</td>
<td></td>
<td></td>
</tr>
<tr>
<td>0.08</td>
<td>−0.12%</td>
<td>−0.18%</td>
<td>−0.34%</td>
<td>−0.89%</td>
<td></td>
<td></td>
</tr>
<tr>
<td>0.12</td>
<td>−0.17%</td>
<td>−0.30%</td>
<td>−0.66%</td>
<td>−1.90%</td>
<td></td>
<td></td>
</tr>
<tr>
<td>0.16</td>
<td>−0.24%</td>
<td>−0.48%</td>
<td>−1.11%</td>
<td>−3.30%</td>
<td></td>
<td></td>
</tr>
<tr>
<td>0.20</td>
<td>−0.33%</td>
<td>−0.70%</td>
<td>−1.69%</td>
<td>−5.07%</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>γ = 5</th>
<th>θ</th>
<th>Horizon (in Quarters)</th>
<th>1</th>
<th>4</th>
<th>12</th>
<th>40</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.00</td>
<td>−5.90%</td>
<td>−5.90%</td>
<td>−5.90%</td>
<td>−5.90%</td>
<td></td>
<td></td>
</tr>
<tr>
<td>0.04</td>
<td>−5.92%</td>
<td>−5.98%</td>
<td>−6.13%</td>
<td>−6.65%</td>
<td></td>
<td></td>
</tr>
<tr>
<td>0.08</td>
<td>−5.99%</td>
<td>−6.22%</td>
<td>−6.81%</td>
<td>−8.88%</td>
<td></td>
<td></td>
</tr>
<tr>
<td>0.12</td>
<td>−6.11%</td>
<td>−6.61%</td>
<td>−7.95%</td>
<td>−12.47%</td>
<td></td>
<td></td>
</tr>
<tr>
<td>0.16</td>
<td>−6.27%</td>
<td>−7.17%</td>
<td>−9.51%</td>
<td>−17.27%</td>
<td></td>
<td></td>
</tr>
<tr>
<td>0.20</td>
<td>−6.48%</td>
<td>−7.88%</td>
<td>−11.49%</td>
<td>−23.05%</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Panel B: Risk-aversion Parameters of Indifference

<table>
<thead>
<tr>
<th>Horizon (in Quarters)</th>
<th>θ</th>
<th>1</th>
<th>4</th>
<th>12</th>
<th>40</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
<td></td>
</tr>
<tr>
<td>0.000</td>
<td>0.000</td>
<td>0.007</td>
<td>0.027</td>
<td>0.089</td>
<td></td>
</tr>
<tr>
<td>0.000</td>
<td>0.000</td>
<td>0.029</td>
<td>0.098</td>
<td>0.279</td>
<td></td>
</tr>
<tr>
<td>0.000</td>
<td>0.000</td>
<td>0.062</td>
<td>0.195</td>
<td>0.462</td>
<td></td>
</tr>
<tr>
<td>0.000</td>
<td>0.103</td>
<td>0.297</td>
<td>0.600</td>
<td></td>
<td></td>
</tr>
<tr>
<td>0.000</td>
<td>0.150</td>
<td>0.393</td>
<td>0.696</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

This table calibrates the effect of multi-period investing on investor utility. The risky investment has a forecasted mean of 4% per quarter and an instantaneous return standard deviation, given μ, of 0.4 per year. For parameter uncertainty (θ) that is 0, 0.04, 0.08, 0.12, and 0.2 per year and horizons of 1, 4, 12, and 40 quarters, Panel A reports the utility equivalent risk-free return per quarter. Certainty equivalents are computed as $CE^T = (1 + 0.04)e^{\frac{1}{4}T(\text{Δ}r - \frac{1}{2}\sigma^2)} - 1$, with $\sigma = 0.4$, $\Delta r = 0.25$, and $T \in \{0.25, 1, 3, 10\}$. Panel A provides these calculations for risk-aversion parameters (γ) of 0, 1/4, 2, and 5. Panel B reports $γ^b = \frac{(T - Δr)σ^2}{Tσ^2 + 2}$, the risk-aversion parameter that makes one indifferent between a riskless investment and a risky project forecasted to earn the same quarterly return.

about μ to the uncertainty about the return conditional on μ). The critical risk-aversion parameter ranges from 0 (at T = Δr) to 1 (for infinite T or θ). Thus, as T or uncertainty about μ gets sufficiently large, those less risk averse than log utility investors strictly prefer risky multi-period projects to otherwise identical risk-free projects.

Table 1 quantifies how multi-period investing influences utility. The table assumes a period length of 3 months, and a risky investment with a forecasted mean m of 4% per quarter and a standard deviation, given μ, of 0.4 per year. For parameter uncertainty (θ) that is 0, 0.04, 0.08, 0.12, 0.16, and 0.2 per year and horizons of 1, 4, 12, and 40 quarters, Panel A reports the utility equivalent risk-free return per quarter, using risk-aversion parameters (γ) of 0, 1/4, 2, and 5. Panel B provides the risk-aversion level, γ^b, that makes one indifferent
between a riskless investment and a risky project that is expected to earn the same return per quarter.

The horizon “1 Quarter” columns of Table 1, Panel A, merely indicate that parameter uncertainty lowers expected utility because it increases risk. However, comparing the numbers in these columns to those on their right indicates that the compounding of parameter uncertainty is equivalent to increasing the perceived quarterly expected return for the two lowest risk-aversion parameters and decreasing it for the two larger ones. The effect is approximately linear in the investment horizon $T$.

The adjustments in Panel A are strikingly large. Consider the case $\theta = 0.08$, a level of parameter uncertainty that accounts for a mere 1% of the variance of the logged quarterly return. Alternatively, $\theta = 0.08$ is the standard error of $\mu$’s estimate after observing 100 quarterly returns. This value of $\theta$ implies that the 4% per quarter expected return on a project (a 16% APR) corresponds to a risk-neutral investor’s certainty equivalent of 4.81% per quarter (a 19.3% APR) at a 10-year horizon. For the same level of parameter uncertainty, the certainty equivalent shifts in the opposite direction and by a similar magnitude for $\gamma = 2$: from $-0.12\%$ per quarter at the 1-quarter horizon to $-0.89\%$ per quarter at the 40-quarter horizon.

Panel B indicates that when $\theta = 0.08$ and $T = 10$ years, every investor with risk aversion below 0.279 prefers the risky project to a risk-free investment earning the same 4% per quarter. When risk aversion exceeds 0.279, the reverse is true. Given how risky the project is, the degree to which compounded parameter uncertainty compensates for risk is rather remarkable. Numbers for $\gamma^b$, such as 0.279, have an interesting “$R$-squared” interpretation. They approximately represent the fraction of the variance of date $T$ log wealth accounted for by parameter uncertainty. Hence, even though a $\theta$ of 0.08 accounts for a mere 1% of each quarter’s return variance from the $\sigma = 0.4$ project, compounding of the persistent risk from parameter uncertainty magnifies its importance. After 10 years, more than 25% of the cumulated return variance derives from $\theta$.

The relationship between $R$-squared and $\gamma^b$ becomes exact with continuous compounding. At the continuous-time limit, the certainty equivalent’s closed-form solution,

$$m^* = m + \frac{1}{2} \left[ (1 - \gamma) T \theta^2 - \gamma \sigma^2 \right],$$

is exactly linear in the investment horizon $T$. The risk-aversion parameter that makes the certainty equivalent identical to $m$ in Equation (21) is the $\gamma$ obtained from Equation (20) with $\Delta t = 0$. This “break-even” risk aversion is the ratio of two variances: the variance of date $T$ log wealth from parameter uncertainty divided by the total variance (from all sources) of date $T$ log wealth. Moreover, because $m^*$, the certainty equivalent, represents the critical instantaneous

---

12 Reduce the total risk and the risky project becomes even more appealing. For example, if $\sigma$ drops in half to 0.2 per year, $\gamma^b$ would jump up to 0.696.
risk-free rate that determines project acceptance, Equation (21) quantifies the project adoption rule with a particularly simple mathematical expression, outlined below.

**Proposition 3.** Whenever \( m > r_f - \frac{1}{2} \left[ (1 - \gamma) T \theta^2 - \gamma \sigma^2 \right] \), the risky project should be accepted. If the inequality is reversed, the risky project should be rejected.

### 2.1.2 The Statistician’s View.

The analysis above computes expected wealth and utility using the investor’s expectation; that is, the computation conditions on \( m \). The statistician’s view of expected wealth and utility is very different. Recall Section 1’s discussion of multiplicative adjustment for exponential functions of normal variables. This discussion suggested that one could examine the investor’s bias adjustment in the exponent of a utility (or wealth) forecast. The statistician, instead of adding this bias adjustment in the exponent, subtracts it. Consistent with this discussion, the statistician’s unbiased estimate of expected wealth is

\[
E(\tilde{W}_T) = e^{r_f T} e^{(m-r_f) T x - \frac{1}{2} x^2 \theta^2 T^2},
\]  

and his unbiased estimate of expected utility is

\[
E \left[ \frac{\tilde{W}_T^{1-\gamma}}{1 - \gamma} \right] = \frac{1}{1 - \gamma} e^{(1-\gamma) T \left[ r_f + x (m-r_f - \frac{1}{2} x \gamma \sigma^2) \right] - \frac{1}{2} x^2 \theta^2 T^2 (1-\gamma)^2}.
\]  

Equation (23) is not the investor’s expected utility of wealth (or even his expected wealth for \( \gamma = 0 \)). As noted earlier, Jensen’s inequality requires the investor to make a positive rather than a negative adjustment for the uncertainty about \( \mu \). The statistician’s adjustment, subtracting \( \frac{1}{2} \theta^2 T^2 (1 - \gamma)^2 \) in the exponent (for \( x = 1 \)), is designed to offset the upward bias in expected utility from forecasting \( \mu \) as \( \mu - \tilde{u} \). The adjustment achieves an enviable track record but does not assess the true distribution of rewards and penalties from a decision.

No more obvious evidence for this exists than the paradoxical result that the statistician’s unbiased expected utility estimate can increase as risk aversion increases. As \( \gamma \) increases, Equation (23) becomes less negative in the region where \( (1 - \gamma)^2 \) dominates the other terms in the equation. The “paradox” stems from the statistician’s recognition that over- and undershooting \( \mu \) by a large amount tends to reduce forecasted expected utility. To offset the downward bias that arises from such parameter uncertainty, the statistician raises his utility estimate. This statistical correction can exceed the utility reduction that is the normal outcome of higher \( \gamma \) in the absence of parameter uncertainty. Even with \( m - r_f < 0 \) (a negative risk premium), the bias adjustment forecasts larger expected utility from risky project adoption than from risk-free investment if
$\gamma$ is large enough. However, being more risk averse makes a risky project less, not more, attractive. Thus, the statistician’s utility estimate, which makes the risky project seem relatively more attractive as risk aversion increases beyond some value, should not be treated as a guide for investment decisions.

Table 2 illustrates what happens if we make this mistake. Panel A computes a hypothetical certainty equivalent—the risk-free return that generates the same forecasted expected utility as the statistician’s unbiased forecast. The grouping in the top-left corner computes this figure for risk-neutral investors using the same calibration as in Table 1. This adjustment is essentially Blume’s (1974), quantified with a more general parameter uncertainty model than that in Jacquier, Kane, and Marcus (2003). The three other Table 2 groups analyze the statistician’s expected utility estimate at three positive values of $\gamma$, using Table 1’s assumptions.

The numbers clearly differ from the corresponding numbers in Table 1. Recall from Table 1 that the investor found a 40-quarter project with a true expected return of 4% per quarter to be equivalent to a 4.81% per quarter riskless project. The statistician’s forecasted long-term payoff for such a project has a certainty equivalent of 3.19% per quarter. No one can dispute that there is an economically meaningful difference between a 19.3% APR estimate (as seen in Table 1) and a 12.8% APR estimate (as seen in Table 2).

Table 2 also complements our earlier discussion in showing that at long horizons, increases in risk aversion perversely seem to increase the hypothetical certainty equivalent when sufficient parameter uncertainty exists. This is evident not just from the 10-year horizon in Panel A with the two largest $\theta$s, but from Panel B as well. For example, Panel B’s entry for $\theta = 0.16$ and $T = 10$ shows that for risk aversion above 2.600, the statistician estimates larger expected utility with a risky project than with risk-free investment earning the same forecasted 4% per quarter. When risk aversion is below 2.600, the reverse is true. At this level of parameter uncertainty and higher, we obtain the perverse result that a 10-year risky project is desirable only to the more risk averse of two investors. There could be no stronger evidence for the inappropriateness of the statistician’s view than Table 2. For this reason, the remainder of the article concentrates on the investor’s view.

### 2.2 Project Scalability

The above analysis restricts investment to be either one dollar in the riskless asset or one dollar in the risky project. Here, we study the issue of splitting the dollar between the two investments. Because utility is not defined for negative wealth, we require the fraction of wealth $x$ invested in the risky asset to be between 0 and 1 inclusive.

---

13 They use diffuse priors and Bayesian updating of $\mu$ from historical data to generate $\tilde{m}$’s distribution.

14 A $\theta$ of 0.16 corresponds to the standard error of $\mu$’s estimate after observing 25 quarters of returns.
Table 2
Certainty-equivalent Returns (The Statistician’s View)

<table>
<thead>
<tr>
<th>Panel A: Certainty-equivalent Returns</th>
<th>Panel B: Risk-aversion Parameters of Indifference</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\gamma = 0$</td>
<td>$\gamma = \frac{1}{T}$</td>
</tr>
<tr>
<td>$\theta$</td>
<td>Horizon (in Quarters)</td>
</tr>
<tr>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>0</td>
<td>4.00%</td>
</tr>
<tr>
<td>0.04</td>
<td>4.00%</td>
</tr>
<tr>
<td>0.08</td>
<td>4.00%</td>
</tr>
<tr>
<td>0.12</td>
<td>4.00%</td>
</tr>
<tr>
<td>0.16</td>
<td>4.00%</td>
</tr>
<tr>
<td>0.20</td>
<td>4.00%</td>
</tr>
</tbody>
</table>

Note: N/A indicates that no risk-aversion parameter of indifference exists.

This table calibrates the statistician’s estimate of multi-period investing on utility. The risky investment has a forecasted mean of 4% per quarter and an instantaneous return standard deviation, given $\mu$, of 0.4 per year. For parameter uncertainty ($\theta$) that is 0, 0.04, 0.08, 0.12, and 0.2 per year and horizons of 1, 4, 12, and 40 quarters, Panel A reports the utility equivalent risk-free return per quarter. Certainty equivalents are computed as

$$CE = (1 + 0.04)e^{- \frac{1}{2} \Delta \left( \frac{(1+\gamma)^T - \Delta \gamma^2}{2 + \gamma^2}\right)} - 1,$$

where $\gamma$ is 0, $\frac{1}{T}$, 2, and 5. Panel B reports $\gamma^b = \frac{(T - \Delta \gamma)^2}{T \theta^2 - \sigma^2}$, the risk-aversion parameter that makes the statistician’s estimate of expected utility identical between a riskless investment and a risky project forecasted to earn the same quarterly return.

Terminal wealth is given by the random variable

$$\tilde{W}_T = e^{r_f T} \left[ (1 - x) + xe^{(m - r_f - \frac{1}{2} \sigma^2 + \bar{u})T + \sigma \bar{z} \sqrt{T}} \right].$$  \hspace{1cm} (24)

Its derivative with respect to $x$ is

$$\frac{d(W_T)}{dx} = \tilde{x},$$  \hspace{1cm} (25)

where $\tilde{x}$ is the $T$-year excess return on the risky asset,

$$\tilde{x} = e^{(m - r_f - \frac{1}{2} \sigma^2 + \bar{u})T + \sigma \bar{z} \sqrt{T}} - e^{r_f T}.$$  \hspace{1cm} (26)
Given that buying and holding to \( T \) is isomorphic to the static one-period portfolio decision, the derivative of expected utility with respect to \( x \) is the familiar equation

\[
\frac{d E[U'(\tilde{W}_T)]}{dx} = E[U'(\tilde{W}_T)\tilde{\pi}] = E[\tilde{W}_T^{-\gamma} \tilde{\pi}], \tag{27}
\]

and the second derivative,

\[
\frac{d^2 E[U'(\tilde{W}_T)]}{dx^2} = -\gamma E[\tilde{W}_T^{-\gamma-1} \tilde{\pi}^2], \tag{28}
\]

is negative, implying expected utility is concave in \( x \). At \( x = 1 \), the first derivative,

\[
\left. \frac{d E[U'(\tilde{W}_T)]}{dx} \right|_{x=1} = E \left[ e^{-\gamma \left( (m-\frac{1}{2}\sigma^2+\tilde{u})T+\sigma \tilde{z}\sqrt{T} \right) - e^{\gamma T}} \right] \\
= E \left[ e^{(1-\gamma) \left( (m-\frac{1}{2}\sigma^2+\tilde{u})T+\sigma \tilde{z}\sqrt{T} \right) - e^{\gamma \left( (m-\frac{1}{2}\sigma^2+\tilde{u})T+\sigma \tilde{z}\sqrt{T} \right)}} \right], \tag{29}
\]

is the difference in the expected values of two lognormally distributed variables. Using the moment-generating function for such variables, we obtain

\[
\left. \frac{d E[U'(\tilde{W}_T)]}{dx} \right|_{x=1} = e^{(1-\gamma)T \left[ m + \frac{1}{2} ((1-\gamma)\theta^2 T - \gamma \sigma^2) \right] - e^T \left[ r_f - \gamma m + \frac{1}{2} \gamma (\gamma \theta^2 T + (1+\gamma)\sigma^2) \right]}, \tag{30}
\]

which is positive whenever the first exponent exceeds the second. This occurs when

\[
m > r_f - \frac{1}{2} \theta^2 T + \gamma (\sigma^2 + \theta^2 T). \tag{31}
\]

Inequality (31) is more stringent than Proposition 3’s condition,

\[
m > r_f - \frac{1}{2} \left( (1 - \gamma)T \theta^2 - \gamma \sigma^2 \right). \tag{32}
\]

When the latter condition is met but inequality (31) is violated, the optimal scale of the project, \( x \), is between zero and one. Note that inequality (31) is
satisfied when $\gamma < \gamma^*$, where

$$\gamma^* = \frac{m - r_f + \frac{1}{2} \theta^2 T}{\sigma^2 + \theta^2 T}.$$  \tag{33}

We summarize this result as follows.

**Proposition 4.** When $\gamma < \gamma^*$, investors want to maximize the scale of the risky project (i.e., $x = 1$). Thus, if $\theta$ or $T$ are sufficiently large, investors who are less risk averse than $\gamma < \frac{1}{2}$ maximize project scale irrespective of how negative the per-period risk premium is.

To prevent negative wealth outcomes, Proposition 4 limits risky project maximum investment to $x = 1$. In contrast to risk-averse isoelastic utility investors, risk-neutral investors can weigh negative wealth outcomes against positive wealth outcomes. In this case ($\gamma = 0$), inequality (31) reduces to $m > r_f - \frac{1}{2} \theta^2 T$. This implies the following:

**Proposition 5.** If $m > r_f - \frac{1}{2} \theta^2 T$ and the amount invested is unconstrained, risk-neutral investors would like to hold an infinite amount of the risky project.

Proposition 5’s inequality is equivalent to the statement that the expectation of Equation (26)’s long-horizon risk premium is positive. Obviously, risk-neutral investors desire infinitely leveraged positions in any asset with a positive long-horizon risk premium, even if the per-period risk premium, $m - r_f$, is negative. The infinite demand for large $T$ or $\theta$ is obvious from Equation (18)’s expression for expected utility. The risky project’s date $T$ payoff in this expression increases with $\theta^2 T^2$, whereas growth in the risk-free asset’s value is proportional to $T$.

In continuous time, the risk-averse investor can hit a zero wealth boundary prior to $T$ without crossing it. Hence, by treating zero wealth as an absorbing state, one can study the demand for a multi-period buy-and-hold asset and allow leveraged positions in the asset. We analyze this problem with numerical methods.

There are model inputs, close to those that make the investor prefer $x = 1$ to lower values of $x$, for which investor utility consistently increases as $x$ substantially increases above 1. The tested values of $x$ are large enough to convince most skeptics that, for certain model inputs, infinite leverage is the solution to the buy-and-hold portfolio problem. For $\gamma > 1$, leveraged buy-and-hold strategies generate infinite negative utility for the finite fraction of paths that hit the zero wealth boundary, making them suboptimal. Hitting the zero wealth boundary causes no similar problem when the investor is less risk averse than the log investor.
If we allow for continuous rebalancing with an unconstrained initial portfolio weight $x$ that is maintained until $T$, wealth (Equation (14)) never reaches 0, even for leveraged positions. For this strategy, the optimal weight has a closed-form solution. Note that for $x = 0$ or $x = 1$, the rebalancing and buy-and-hold strategies are identical. However, we study the rebalancing strategy to understand the optimal $x$ without constraints on the initial $x$. Later, we show that these insights generalize to dynamic portfolio choice with learning.

Equation (15) is expected utility from a continuous rebalancing strategy with constant weight $x$ invested in the risky asset and the remainder in the risk-free asset. Irrespective of how negative the instantaneous risk premium is, the maximum expected utility is unbounded when the term multiplying $x^2$ in Equation (15)’s exponent is positive,

$$(1 - \gamma)T\theta^2 - \gamma \sigma^2 > 0. \quad (34)$$

This is the now familiar condition,

$$\gamma < \frac{T\theta^2}{T\theta^2 + \sigma^2}. \quad (35)$$

The critical boundary in Equation (35) is the same as in Equation (20), with $\Delta t$ set to zero. As $T$ gets large, condition (35) on $\gamma$ converges to the class of utility functions that are less risk averse than log utility. If $\gamma$ is above the critical threshold in (35), the optimal demand from the first-order condition of Equation (15) is

$$x^* = \frac{m - rf}{\gamma \sigma^2 - (1 - \gamma)T\theta^2}. \quad (36)$$

It is wrong to think that Equation (36) gives the optimal portfolio weight for $\gamma$s equal to or lower than the critical threshold. Equation (36) proposes a finite $x^*$ of the opposite sign as the instantaneous risk premium when the denominator is negative. However, when risk aversion is sufficiently low, expected utility (Equation (15)) is infinite for both infinite positive and negative $x$—irrespective of the sign or magnitude of the instantaneous risk premium $m - rf$. (In these cases, the weight in Equation (36) satisfies a first-order condition that generates a local minimum.)

We also used numerical methods to compare the constant rebalancing strategy’s expected utility to buy-and-hold utility for the same initial $x$. These comparisons indicate that, for all values of risk aversion, expected utility is lower for the rebalancing strategy when $x$ is between 0 and 1 and $\theta/\sigma$ is not too

---

15 The optimization problem here has previously been studied by Jacquier, Kane, and Marcus (2005). Our treatment of the issue differs slightly in its use of geometric Brownian motion diffusion. Unlike their model, we employ a drift term that does not depend on $\sigma^2$. We also study a wider variety of issues with the model, extend it to multiple assets (in a later subsection), and allow priors to be more general.
small, but higher for $x > 1$. This is obvious for $\gamma > 1$ because the leveraged buy-and-hold strategy hits the zero wealth boundary for a finite fraction of paths.

Figure 2 verifies the points made above. For an instantaneous risk premium, $m - r_f$, of 0, it plots expected utility from buy-and-hold and rebalancing strategies as a function of $x$ at values of $x$ ranging from 0 to 8. Panel A assumes $\gamma = 0.75$, whereas Panel B assumes $\gamma = 0.35$. These choices straddle the critical $\gamma$'s of indifference. Given the other parameters, $\sigma = 0.1$, $\theta = 0.05$, and $T = 16$, the critical $\gamma$'s are 0.8 for the continuous rebalancing strategy and 0.4 for the buy-and-hold strategy. In Panel A, the rebalancing strategy's expected utility is a convex, monotonically increasing function of $x$, whereas the buy-and-hold strategy's expected utility decreases after reaching a maximum below $x = 1$. In Panel B, expected utility increases to infinity for both strategies. Interestingly, the buy-and-hold strategy has two points at which the slope is zero, but as $x$ gets large, it ultimately becomes a concave increasing function of $x$.

### 2.3 Multiple Risky Assets and Diversification

Uncertainty about nature’s draw of $\mu$ increases the risky project’s expected payoff after multiple periods. Could this mean that diversification, which reduces uncertainty, has the potential to harm investors? Consider a manager allocating capital of $1 to a risk-free asset and $N$ risky projects, indexed by $n$. Project $n$’s value follows the diffusion

$$\frac{d\tilde{P}_{t,n}}{P_{t,n}} = (m + \tilde{u}_n)dt + \sigma d\tilde{z}_{t,n}. \quad (37)$$

For any two distinct dates, $t$ and $t'$, or any two distinct projects, $n$ and $n'$, we assume

$$\text{cov}(d\tilde{z}_{t,n}, d\tilde{z}_{t',n'}) = \text{cov}(\tilde{u}_n, \tilde{u}_{n'}) = 0. \quad (38)$$

(Also, the correlation structure of any given project’s random variables is the same as in the single-project model described earlier.) In short, projects are independent but otherwise identical. When $\theta = 0$, the optimal risky portfolio would continuously rebalance some equal weighting of the risky projects. It would also mix that equal-weighted portfolio of risky projects with the risk-free asset, where the investor’s risk aversion determines the mix.

It is tempting to believe that one should weight the risky projects equally because equal weighting diversifies both $\tilde{z}$ and $\tilde{u}$ risk. This logic forgets that while risk from $\tilde{u}$ adversely affects utility by increasing instantaneous return

---

16 As $\theta/\sigma$ approaches zero, the rebalancing strategy becomes optimal.
Figure 2

Expected utility for continuous rebalancing and buy-and-hold strategies

This figure plots expected utility for a buy-and-hold strategy against a constant rebalancing strategy for multiple values of $x$ ranging from 0 to 8. Both panels fix the following parameters: $\sigma = 10\%$, $\theta = 5\%$, $m = r_f$, and $T = 16$. Given these parameter values, the critical $\gamma^*$ of indifference is 0.8 for the continuous rebalancing strategy and 0.4 for the buy-and-hold strategy. In Panel A, the risk-aversion coefficient is $\gamma = 0.75$, which is below the critical value of $\gamma^* = 0.8$ for the continuous rebalancing strategy; in Panel B, the risk-aversion coefficient is $\gamma = 0.35$, which is below the critical value of $\gamma^* = 0.4$ for the buy-and-hold strategy.
variance, portfolios that reduce that variance also decrease the expected multi-period payoff of the portfolio. For sufficiently low risk aversion, the mean effect dominates the risk effect, and the investor would prefer a single project to an equal weighting of the $N$ projects. We summarize this result below.

**Proposition 6.** Assume a continuously rebalanced portfolio maintaining constant weights on $N$ independent but otherwise identical risky assets and a risk-free asset. If the risk-aversion parameter $\gamma < \frac{T \theta^2}{T \theta^2 + \sigma^2}$ and if $x$ cannot be negative, the optimal portfolio mixes (any) one of the risky projects with the risk-free asset.

**Proof.** The portfolio problem selects nonnegative $x_1, x_2, \ldots, x_N$ to maximize

\[
\mathbb{E}\left[\frac{\tilde{W}_T^{1-\gamma}}{1-\gamma}\right] = \frac{1}{1-\gamma} e^{(1-\gamma)T \left[ r_f + (m-r_f) \sum x_n + \frac{1}{2} ((1-\gamma)T \theta^2 - \gamma \sigma^2) \sum x_n^2 \right]},
\]

which is equivalent to maximizing

\[
(m - r_f) \sum x_n + \frac{1}{2} (1 - \gamma)T \theta^2 - \gamma \sigma^2 \sum x_n^2.
\]

For $(1 - \gamma)T \theta^2 - \gamma \sigma^2 > 0$, this is achieved by making any $x_n$ infinite. If the risky asset portfolio weights sum to a finite number (positive for $m > r_f$ and negative otherwise), one needs to simply maximize $\sum x_n^2$—requiring all but one of the $x_n$s to equal zero. By contrast, if $(1 - \gamma)T \theta^2 - \gamma \sigma^2 < 0$, then minimizing $\sum x_n^2$ maximizes expected utility. With $\sum x_n$ fixed, investing equal amounts in each of the risky assets achieves this goal.

The intuition for Proposition 6 is clear when the investor is risk neutral. Here, $e^{mT + \frac{1}{2} \theta^2 T^2}$ is date $T$’s expected wealth from $\$1$ invested in a single project, whereas the smaller value $e^{mT + \frac{1}{2} \tilde{\theta}^2 T^2}$ is expected wealth from weighting the $N$ projects equally, reflecting the diversified risky portfolio’s lower parameter uncertainty. By extension, when the projects have independent $\tilde{\theta}$ risk but different $\theta$s, and $m$ does not vary across the risky projects, investing the fixed $\sum x_n$ in the project with the largest $\theta$ is optimal with sufficiently low risk aversion.

One could obtain closed-form solutions to more general portfolio selection models with risks that are correlated across assets and risk premium forecasts that differ across the $N$ projects. However, the simpler model gets the point across; that is, putting all your eggs in one basket may pay! By contrast, the buy-and-hold investor who lowers parameter uncertainty risk by diversifying does not reduce the portfolio’s multi-period expected return. For example, if all assets are forecasted to have the same expected return, then all vectors of risky asset portfolio weights with the same sum generate identical expected wealth.

24
at $T$. Since risk-neutral investors are indifferent to diversification with the buy-and-hold strategy, but not with a rebalancing strategy, risk-averse investors pursuing a buy-and-hold strategy should want to diversify. This intuition is indeed correct, as the following proposition demonstrates.

**Proposition 7.** Assume a buy-and-hold portfolio of $N$ independent but otherwise identical risky assets and a risk-free asset. If the risky assets have positive instantaneous risk premia, the optimal portfolio has equal positive weight on each of the risky assets.

**Proof.** Generalizing Equation (27), the $N$ first-order conditions are

$$
\frac{\partial E[U(\tilde{W}_T)\tilde{\pi}_n]}{\partial x_n} = E[U'(\tilde{W}_T)\tilde{\pi}_n] = E[\tilde{W}_T^{-\gamma}\tilde{\pi}_n] = 0,
$$

which is only satisfied for some scaling of equal $x_n$s. Generalizing Equation (28) to multiple assets, the Hessian indicates that the solution satisfies the second-order condition.

2.4 Corporate Finance Interpretations

The continuous rebalancing strategy partially sells the better-performing assets and buys the poorly performing assets in order to maintain constant weights on each. Parameter uncertainty implies that past performance is correlated with expected future returns; thus, rebalancing risky asset weights relative to one another represents poor timing that diminishes expected wealth. Keeping $\sum x_n$ fixed, the poor timing’s expected wealth effect is greatest when maximally diversified. By contrast, the “put all your eggs in one basket” strategy avoids this poor timing—there is no other risky asset to move into or out of in response to relative performance.

Even though the buy-and-hold investor prefers to diversify, heterogeneous parameter risk still generates an incentive to tilt weights in favor of assets with larger parameter risk, $\theta$. Over time, positive buy-and-hold portfolio weights automatically revise in favor of assets that performed relatively well in the past. If the investor had complete dynamic control of the portfolio weights, revisions would be greatest in the assets with the greatest parameter uncertainty. Posteriors, and thus weight revisions, are more sensitive to signals about assets with the least informative priors. Tilting the buy-and-hold portfolio in favor of assets with greater parameter uncertainty better mimics this superior revision strategy. Such tilts make portfolio revisions larger for the assets with greater parameter uncertainty. At the margin, this tilt has to be profitable for $N$ otherwise identical assets. Compared to equal weighting, the marginal increase in variance for a small tilt is an order of magnitude less than the marginal increase in expected return.
Timing of buys and sells also explains the rebalancing strategy’s appetite for leverage. In the single risky asset case, a leveraged rebalanced portfolio buys more of the risky asset when the asset has experienced good past performance. This is like the homeowner taking out a second mortgage to renovate his house because the home price increase raised his equity in the property. If changes in home values are persistent, the renovation financed with the second mortgage is expected to increase future wealth.

The above intuition has a real options flavor. Various strategies mimic desirable investment policies the investor would implement if he had complete control over project scale at each instant in time. For example, multi-period investment in a single risky project, whether based on buy-and-hold or rebalanced weights, is in some sense a vehicle to increase or decrease investment at the right time. When past returns are high, both expected future returns and next period’s total dollar investment in the risky project are larger. Multi-period investment can thus be viewed as an automatic mechanism for increasing total dollar investment in a risky project when returns are expected to be larger and vice versa. It is the positive covariance between capital invested and future returns that makes expected wealth increase with $T$.

This finding bears some similarity to the project-scale options in McDonald and Siegel (1985, 1986). To illustrate the similarity, it is useful to show that if the risky project’s total investment does not increase or decrease in response to past returns, parameter uncertainty’s only effect is to increase variance; its multi-period mean effects disappear in the absence of compounding.

To study this issue, assume that all returns from the risky asset, rather than being reinvested, are placed in a riskless asset assumed to earn a return of zero (without loss of generality) and thus do not compound. Date $t$’s “simple” risky-asset return is

$$\frac{d\tilde{P}_t}{P_0} = (m + \tilde{u})dt + \sigma d\tilde{z}_t.$$  \hspace{1cm} (42)

With initial wealth of $1, investing $x$ in the risky project generates date $T$ terminal wealth

$$\tilde{W}_T = 1 + x \left( (m + \tilde{u})T + \sigma \int d\tilde{z}_t \right).$$  \hspace{1cm} (43)

The derivative of the expected utility of this wealth with respect to $x$ is

$$\frac{dE[U(\tilde{W}_T)]}{dx} = E \left[ \left( 1 + x \left( (m + \tilde{u})T + \sigma \int d\tilde{z}_t \right) \right)^{-\gamma} \left( (m + \tilde{u})T + \sigma \int d\tilde{z}_t \right) \right].$$  \hspace{1cm} (44)
By Stein’s Lemma,\textsuperscript{17} Equation (44) can be rewritten as

\[
\frac{dE[U(W_T)]}{dx} = E \left[ \left( 1 + x \left( (m + \tilde{u})T + \sigma \int d\tilde{z}_t \right) \right)^{-\gamma} \right] mT \tag{45}
\]

\[
- \gamma E \left[ \left( 1 + x \left( (m + \tilde{u})T + \sigma \int d\tilde{z}_t \right) \right)^{-\gamma - 1} \right]
\]

\[
\times \left( \theta^2 T + \sigma^2 \right) xT. \tag{46}
\]

If the risky asset’s forecasted instantaneous risk premium (here, \( m \) because \( r_f = 0 \)) is not positive, this derivative is negative for all positive \( x \). Thus, in the absence of compounding, risky assets with negative risk premia are not held. This finding is not surprising, as maximizing the expected utility of wealth is isomorphic to a one-period optimization problem here.

\[17\] Stein’s Lemma states that, if \( \tilde{W}_T \) and \( \tilde{x} \) are bivariate normal, 

\[
E[g(\tilde{W}_T) \tilde{x}] = E[g(\tilde{W}_T)]E[\tilde{x}] + E[g'(\tilde{W}_T)] \text{cov}(\tilde{W}_T, \tilde{x}),
\]

with \( \tilde{x} \equiv (m + \tilde{u})T + \sigma \int d\tilde{z}_t \).

\[18\] Grinblatt and Linnainmaa (2011) develop this idea further.
earlier model, which provides a useful benchmark for Brennan’s model, as the “constrained rebalancing model.”

Brennan’s model, by contrast, allows the investor to continuously alter the risky asset’s portfolio weight as he learns about the asset’s conditional mean return through Bayesian updating. The investor’s prior belief about \( \mu \) is normal with mean \( m_0 \) and variance \( v_0 \). The asset’s price path determines the path of the posterior \( m_t \). The conditional variance declines deterministically in \( t \) as the investor continuously resolves uncertainty about \( \mu \). Because new observations are of the order of \( dt \), the posterior mean and conditional variance change at rates of the order of \( dt \).

Recall that the risky asset’s optimal weight in the constrained rebalancing model,

\[
    x^* = \frac{m - r_f}{\gamma \sigma^2 - (1 - \gamma)T\theta^2},
\]

is valid only for positive denominators. When the denominator is negative, both demand and utility are infinite. This insight suggests that the solutions found for continuous-time learning models, such as those derived in Brennan (1998), do not apply for \( \gamma \leq T\theta^2/\sigma^2 + \sigma^2 \).

Note that Brennan’s model and the constrained rebalancing model remain fundamentally distinct even though they generate the same initial portfolio weight on the risky asset. For example, when prices increase, the mean forecast \( m \) increases. In this case, Brennan’s investor will increase his weight on the risky asset, whereas the constrained rebalancing investor is prevented from doing so. Thus, the dynamic evolution of the risky asset weight in Brennan’s model differs from the evolution in the constrained rebalancing model. Nevertheless, their starting points are the same and the insights generated earlier for the constrained rebalancing model also apply to Brennan’s model (and to continuous-time models with the same setup as Brennan’s model, but with no learning). This justifies our extensive treatment of the constrained rebalancing model. Rather than being an artificial model, it is a gateway to insights about models with more realistic portfolio choice assumptions.

### 2.6 Investing at the Margin: When Does Active Management Pay?

We now study the case of two risky assets and one risk-free asset. One risky asset lacks parameter uncertainty and is the optimal risky investment in the

19 For \( \gamma \) this low, the dynamic programming problem’s first-order conditions yield suboptimal portfolios because the second-order conditions are not satisfied.

20 We take no credit for this solution. This solution was provided to us by the late Yihong Xia many years ago and it was derived independently by Rogers (2001).
absence of parameter uncertainty. A second asset has an unknown mean, but its forecast precludes positive investment in the asset in the absence of parameter uncertainty. We allow scalability in asking the question: Under what conditions does it pay to add the asset with parameter uncertainty to one’s overall investment portfolio? The answer here does not depend on the investor’s risk aversion.

Mutual fund investment provides a concrete analogy for the portfolio problem outlined above. The risky asset with no parameter uncertainty is an index fund consisting of the market portfolio; the other is an actively managed fund with unknown alpha. Without loss of generality, assume that the actively managed fund has no beta risk. Thus, for initial wealth of one, investing $x_p$ in the passive index fund and $1 - x_p$ in the risk-free asset generates wealth

\[
\tilde{W}_T^{(\text{passive})} = e^{r_f T} \left[ 1 - x_p + x_p e^{\mu_p - r_f - \frac{1}{2} \sigma_p^2 + \sigma_p \tilde{z}_p \sqrt{T}} \right].
\]

(48)

Investing an additional amount $x_a$ in the actively managed fund and financing it with $x_a$ less in the risk-free asset leads to wealth of

\[
\tilde{W}_T = e^{r_f T} \left[ 1 - x_a - x_p + x_a e^{(m_a - r_f - \frac{1}{2} \sigma_a^2 + \theta_a \tilde{u}_a) T + \sigma_a \tilde{z}_a \sqrt{T}} \right.

\[+ x_p e^{\mu_p - r_f - \frac{1}{2} \sigma_p^2 + \sigma_p \tilde{z}_p \sqrt{T}} \right].
\]

(49)

Denoting the excess return from active investing as

\[
\tilde{\pi}_a = e^{(m_a - \frac{1}{2} \sigma_a^2 + \theta_a \tilde{u}_a) T + \sigma_a \tilde{z}_a \sqrt{T}} - e^{r_f T},
\]

(50)

the decision of whether to invest a small amount in the actively managed fund hinges on the sign of the derivative of expected utility at $x_a = 0$,

\[
\frac{\partial}{\partial x_a} \left. E[U(\tilde{W}_T)] \right|_{x_a=0} = E[U'(\tilde{W}_T^{(\text{passive})}) \tilde{\pi}_a] = E[\tilde{W}_T^{-(\gamma)} (\text{passive}) \tilde{\pi}_a],
\]

(51)

because $\tilde{u}_a$, $\tilde{z}_a$, and $\tilde{z}_p$ are independent. This derivative is positive when

\[
E[\tilde{\pi}_a] = e^{m_a T + \frac{1}{2} \theta_a^2 T^2} - e^{r_f T} > 0,
\]

(52)

or equivalently,

\[
m_a - r_f > -\frac{1}{2} \theta_a^2 T.
\]

(53)

---

21 If the actively managed fund has beta risk, investing in the actively managed fund would increase beta risk above the investor’s target level. However, by adjusting the investment in the passively managed fund downward by an equal amount, the investor achieves the same target beta.
Here, since $m_a - r_f$ is the estimate of the fund’s instantaneous alpha, expected utility is enhanced by buying some amount of the actively managed fund, provided that the fund’s estimated alpha per period is not too negative. Moreover, no matter how negative the per-period alpha is, for sufficient alpha uncertainty or horizon length, combining the negative-alpha active fund with the passively managed market portfolio is still profitable.

The key assumption here is that $\tilde{u}_a$ and $\tilde{z}_p$ are uncorrelated; that is, estimation risk is independent of future market returns. The correlation between $\tilde{z}_a$ and $\tilde{z}_p$ is not relevant to the result. If the latter two variables are correlated (e.g., if the fund has a positive beta), one could project $\tilde{z}_a$ onto $\tilde{z}_p$, obtaining $\tilde{z}_a = \beta_a \tilde{z}_p + \tilde{\epsilon}_a$. Holding $\beta_a$ units less of the passive portfolio for every unit added of the active fund yields the same terminal wealth distribution as the solution to the optimization problem with the zero beta active fund.

Our results also could be generalized to account for parameter uncertainty about the passively managed fund’s risk premium. Provided that the parameter uncertainty of the actively managed fund is orthogonal to the other random variables that determine wealth, the result above remains the same. We summarize the result as follows.

**Proposition 8.** Irrespective of risk aversion, an investor would like to add an actively managed portfolio with a per-period alpha above $-\frac{1}{2} \theta_a^2 T$ to his passive optimal portfolio if parameter uncertainty about the alpha of the active manager is independent of the return of his passive optimal portfolio.

Proposition 8 implies that counseling investors to avoid funds or groups of funds because they have negative alpha estimates is naive: These arguments are based on per-period alphas. The condition $m_a > r_f - \frac{1}{2} \theta_a^2$ is equivalent to saying that the fund’s long-horizon alpha is positive. Adding a positive long-horizon alpha asset to a portfolio raises the portfolio’s Sharpe ratio. Proposition 8 merely points out that alpha uncertainty makes it possible to have a negative per-period alpha and a (more relevant) positive long-horizon alpha at the same time.

Proposition 8’s conclusion applies even if one learns about the true alpha over time. Recall that learning in continuous time does not alter parameter uncertainty’s influence on ex ante portfolio choice. Moreover, even in discrete time, simulations indicate that learning’s effect on Proposition 8’s result is relatively small. As in Brennan’s (1998) model, the subsequent path of the optimal portfolio is influenced by learning even though learning does not alter the initial portfolio. Thus, if investors learn about alphas over time, we expect fund flows to chase past returns—a well-documented phenomenon.

### 3. Conclusion

The field of finance has long pushed the classical statistician’s view that unbiased estimates of long-term mean returns are the appropriate tool for
investment decisions. We now know this view is wrong. Any persistence in realized returns positively influences both the long-term mean and the long-term variance of future wealth. Compounding magnifies variance effects on utility more than mean effects for those more risk averse than log investors; mean effects amplify more than variance effects for those less risk averse than log investors. For long horizons, at risk-aversion levels below the fraction of total variance accounted for by parameter uncertainty, the mean effects actually dominate and the investor’s expected utility can become convex in the risk born. Consequently, Jensen’s inequality may make it rational for risk-neutral and mildly risk-averse investors to adopt projects with negative risk premia.

Part of the pitfall here is that one can empirically confirm the statistician’s view by observing, for example, a long time series of asset returns. The investor’s view has no empirical test for it. Nature has one draw of the true return-generating process, so the investor cannot observe parameter uncertainty from a time series of returns. Given the one true value, the data look like a random walk. The investor is unperturbed, as he expects the data to appear this way. At the same time, he understands that persistence in future returns, although unmeasurable, naturally arises from his uncertainty about the asset’s return-generating process.

Both the statistician and the investor make perfectly rational but different long-term return forecasts. This is because the targets they forecast differ from one another, due to Jensen’s inequality and differences in the way they form conditional distributions. It is easy to understand why the conflict appears in the multi-period setting but not the single-period one. When a forecast of the one-period mean is too high, the effect on the multi-period future value’s overestimate is not offset by a symmetric underestimate of the one-period mean, as it is in the one-period problem. The classical statistician’s objective, to eliminate the bias, justifies shifting the estimated arithmetic mean downward. However, from the investor’s perspective, overestimation of the one-period $\mu$ corresponds to a smaller multi-period return disappointment than the corresponding reward from a symmetric underestimate. The investor thus expects the multi-period return to be higher than that obtained from compounding the estimated one-period mean. Gambles with great uncertainty about nature’s draw of the mean are better if the outcomes compound.

Our simple insight here is obvious in hindsight but has ramifications for the most sophisticated models in portfolio theory. In particular, our finding that constant rebalancing strategies may prefer infinite leverage and generate infinite utility casts suspicions on dynamic programming solutions for investors with low risk aversion. We also find that, for some strategies, diversification does not pay for mildly risk-averse investors because the beneficial long-term mean effects from non-diversification outweigh the loss in utility from increased long-term risk.

Jensen’s inequality, which drives our multi-period investment results, also has unexpected ramifications for the study of many other issues in finance.
Persistence in return shocks can arise for a variety of reasons that are distinct from parameter uncertainty. For example, many investment models assume that nature’s mean follows an autoregressive process. In a multi-period setting, the persistence dies off, but for mildly risk-averse investors making multi-period investment decisions, such persistence enhances means. Our article presents a fresh perspective on how to frame the effects of such persistence on utility and rational decisions.

Many of our novel results hinge on low risk aversion. This begs the question, “Should we care about low risk aversion?” It would be wrong to dismiss our analysis because low risk aversion is not consistent with some empirical estimates of the intertemporal elasticity of substitution and the market risk premium. Experiments point to considerable heterogeneity in risk aversion and there is some evidence from market prices justifying assumptions of low risk aversion, such as the level of the risk-free rate. Moreover, even in the absence of parameter uncertainty, Ross (1999) has shown that only isoelastic utility functions with risk aversion below one are consistent with time diversification: the appealing concept that investors may find repeated i.i.d. multiplicative gambles with positive per-period risk premia attractive when a single gamble is not attractive. It should now be clear that if you take the first “i” out of the “i.i.d.” assumption, Ross’s sequence of gambles becomes even more attractive—to the point where they might be utility enhancing even if they have negative risk premia.

We also know that low risk aversion is an important concept when studying decisions at the margin. As the analysis of active fund management proved, the critical risk-aversion issue is whether parameter uncertainty or some other form of persistent shock to returns is correlated with consumption. At the margin, believing that everyone is risk neutral with respect to certain persistent shocks is plausible. As long as parameter uncertainty is idiosyncratic, persistence has effects on means that will be overlooked in models that focus on high risk aversion.

At the margin, it is well known that risks with zero betas should be treated as if the investor were risk neutral. Our analysis of the desirability of actively managed mutual funds is a pedagogical device to illustrate a key point that generalizes to other assets: Even if the risky investment has a positive beta, as long as its parameter uncertainty is uncorrelated with the market, the risk aversion of the investor is irrelevant at the margin.

The setting for our analysis is one in which there is uncertainty about an exogenous return process. This partial equilibrium view is more appropriate for

---

22 See, for example, Weil (1989) and Browning, Hansen, and Heckman (1999) for a discussion of the risk-free rate puzzle. Studies such as those by Metrick (1995), Barsky, Juster, Kimball, and Shapiro (1997), and Cohen and Einav (2007) measure individuals’ risk preferences through surveys and by observing individuals’ choices in TV game shows and car insurance contracts. They find significant heterogeneity in attitudes toward risk. In Barsky, Juster, Kimball, and Shapiro (1997), for example, 10% to 20% of respondents have estimated γ s below one. Metrick (1995) finds that the data from “Jeopardy!” imply near risk-neutrality.
corporate projects than for traded liquid assets. If uncertainty is about a primitive process, such as dividends, and assets are traded in liquid markets, learning can have additional effects in equilibrium.\textsuperscript{23} Addressing the consequences of learning, including the question about the existence (or impossibility) of equilibrium, would be an interesting extension of the ideas presented here. However, one key takeaway from our exercise is that financial economics is as likely to generate key insights by focusing as much attention on compounding as on learning. This lesson is one that has been lost to finance, not just for a couple of decades (e.g., see McCardle and Winkler 1989) but for a couple of centuries. In 1790, Benjamin Franklin, as part of his will, deposited what amounted to about 1100 dollars in a Boston bank.\textsuperscript{24} To teach the world about compound interest, he bound the bank to not reveal the size of the account for 200 years. Despite interest rates far below the 5% Franklin projected, the account was “unfrozen” in 1990 and the value was revealed to be approximately five million dollars!

\textbf{References}


\textsuperscript{23} Veronesi (2000), for example, shows that uncertainty about the dividend growth rate leads to a counterintuitive result: The risk premium becomes negative for a sufficiently risk-averse representative agent.

\textsuperscript{24} The actual amount was 1000 pounds, about 4400 dollars, but 75% was removed in 1890.


