4.5 The Markowitz inconsistency

Assume that someone tells you that the probability of an event is exactly zero. You ask him where he got this from. "Baal told me" is the answer. In such case, the person is coherent, but would be deemed unrealistic by non-Baalists. But if on the other hand, the person tells you "I estimated it to be zero," we have a problem. The person is both unrealistic and inconsistent. Something estimated needs to have an estimation error. So probability cannot be zero if it is estimated, its lower bound is linked to the estimation error; the higher the estimation error, the higher the probability, up to a point. As with Laplace's argument of total ignorance, an infinite estimation error pushes the probability toward $\frac{1}{2}$. We will return to the implication of the mistake; take for now that anything estimating a parameter and then putting it into an equation is different from estimating the equation across parameters. And Markowitz was inconsistent by starting his "seminal" paper with "Assume you know E and V" (that is, the expectation and the variance). At the end of the paper he accepts that they need to be estimated, and what is worse, with a combination of statistical techniques and the "judgment of practical men." Well, if these parameters need to be estimated, with an error, then the derivations need to be written differently and, of course, we would have no such model. Economic models are extremely fragile to assumptions, in the sense that a slight alteration in these assumptions can lead to extremely consequential differences in the results.

4.6 Psychological pseudo-biases under second layer of uncertainty.

Often psychologists and behavioral economists find "irrational behavior" (or call it under something more polite like "biased") as agents do not appear to follow a normative model and violate their models of rationality. But almost all these correspond to missing a second layer of uncertainty by a tinky-toy first-order model that doesn't get nonlinearities − it is the researcher who is making a mistake, not the real-world agent. Recall that the expansion from "small world" to "larger world" can be simulated by perturbation of parameters, or "stochasticization", that is making something that appears deterministic a random variable itself. Benartzi and Thaler [3], for instance, find an explanation that agents are victims of a disease labelled "myopic loss aversion" in not investing enough in equities, not realizing that these agents may have a more complex, fat-tailed model. Under fat tails, no such puzzle exists, and if it does, it is certainly not from such myopia.

This approach invites "paternalism" in "nudging" the preferences of agents in a manner to fit professors-without-skin-in-the-game-using-wrong-models.

The problem also applies to GMOs and how "risk experts" find them acceptable; researchers pathologize those who do not partake of the baby models (thin tailed). The point, an extension of the Pinker problem, is discussed in Chapter x.

Let us use our approach in detecting convexity to three specific problems: 1) the myopic loss aversion that we just discussed, 2) time preferences, 3) probability matching.

Figure 4.3: The effect of $H_{a,p}(t)$ "utility" or prospect theory of under second order effect on variance. Here $\sigma = 1, \mu = 1$ and t variable.

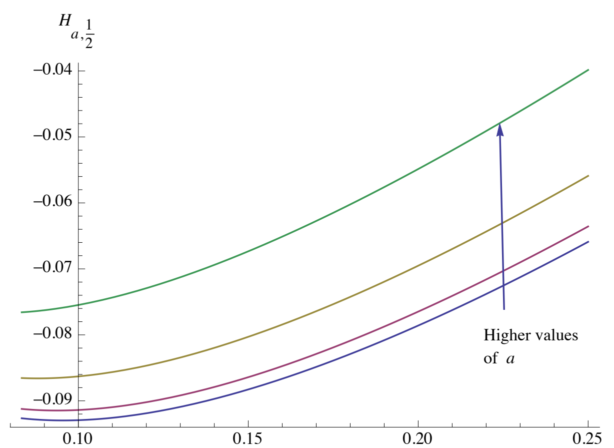
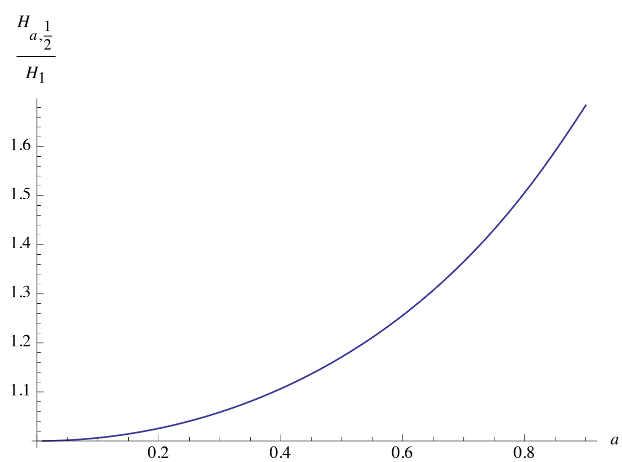


Figure 4.4: The ratio $\frac{H_{a,\frac{1}{2}}(t)}{H_0}$ or the degradation of "utility" under second order effects.



4.6.1 Myopic loss aversion

Take the prospect theory valuation w function for x changes in wealth.

$$w_{\lambda,\alpha}(x) = x^\alpha \, \mathbb{1}_{x\geq 0} - \lambda(-x^\alpha) \, \mathbb{1}_{x<0}$$

Where $\phi_{\mu t,\sigma\sqrt{t}}(x)$ is the Normal Distribution density with corresponding mean and standard deviation (scaled by t)

The expected "utility" (in the prospect sense):

$$H_0(t) = \int_{-\infty}^{\infty} w_{\lambda,\alpha}(x)\phi_{\mu t,\sigma\sqrt{t}}(x)\,\mathrm{d}x \tag{4.5}$$

$$\begin{aligned}
&= \frac{1}{\sqrt{\pi}} 2^{\frac{\alpha}{2}-2} \left(\frac{1}{\sigma^2 t}\right)^{-\frac{\alpha}{2}} \\
&\left(\Gamma\left(\tfrac{\alpha+1}{2}\right)\left(\sigma^\alpha t^{\alpha/2}\left(\tfrac{1}{\sigma^2 t}\right)^{\alpha/2} - \lambda\sigma\sqrt{t}\sqrt{\tfrac{1}{\sigma^2 t}}\right) {}_1F_1\left(-\tfrac{\alpha}{2};\tfrac{1}{2};-\tfrac{t\mu^2}{2\sigma^2}\right)\right. \\
&\qquad\qquad\qquad\qquad + \tfrac{1}{\sqrt{2}\sigma}\mu\Gamma\left(\tfrac{\alpha}{2}+1\right) \\
&\left.\left(\sigma^{\alpha+1} t^{\frac{\alpha}{2}+1}\left(\tfrac{1}{\sigma^2 t}\right)^{\frac{\alpha+1}{2}} + \sigma^\alpha t^{\frac{\alpha+1}{2}}\left(\tfrac{1}{\sigma^2 t}\right)^{\alpha/2} + 2\lambda\sigma t\sqrt{\tfrac{1}{\sigma^2 t}}\right) {}_1F_1\left(\tfrac{1-\alpha}{2};\tfrac{3}{2};-\tfrac{t\mu^2}{2\sigma^2}\right)\right)
\end{aligned} \tag{4.6}$$

We can see from 4.6 that the more frequent sampling of the performance translates into worse utility. So what Benartzi and Thaler did was try to find the sampling period "myopia" that translates into the sampling frequency that causes the "premium" —the error being that they missed second order effects.

Now under variations of σ with stochatic effects, heuristically captured, the story changes: what if there is a very small probability that the variance gets multiplied by a large number, with the total variance remaining the same? The key here is that we are not even changing the variance at all: we are only shifting the distribution to the tails. We are here generously assuming that by the law of large numbers it was established that the "equity premium puzzle" was true and that stocks *really* outperformed bonds.

So we switch between two states, $(1+a)\,\sigma^2$ w.p. p and $(1-a)$ w.p. $(1-p)$.

Rewriting 4.5

$$H_{a,p}(t) = \int_{-\infty}^{\infty} w_{\lambda,\alpha}(x)\left(p\,\phi_{\mu\,t,\sqrt{1+a}\,\sigma\sqrt{t}}(x) + (1-p)\,\phi_{\mu\,t,\sqrt{1-a}\,\sigma\sqrt{t}}(x)\right)\mathrm{d}x \tag{4.7}$$

Result. Conclusively, as can be seen in figures 4.3 and 4.4, second order effects cancel the statements made from "myopic" loss aversion. This doesn't mean that myopia doesn't have effects, rather that it cannot explain the "equity premium", not from the outside (i.e. the distribution might have different returns", but from the inside, owing to the structure of the Kahneman-Tversky value function $v(x)$.

Comment. We used the (1+a) heuristic largely for illustrative reasons; we could use a full distribution for σ^2 with similar results. For instance the gamma distribution with density $f(v) = \frac{v^{\gamma-1} e^{-\frac{\alpha v}{V}} \left(\frac{V}{\alpha}\right)^{-\gamma}}{\Gamma(\gamma)}$ with expectation V matching the variance used in the "equity premium" theory.

Rewriting 4.7 under that form,

$$\int_{-\infty}^{\infty} \int_{0}^{\infty} w_{\lambda,\alpha}(x) \phi_{\mu\, t, \sqrt{v\, t}}(x)\, f(v)\, \mathrm{d}v\, \mathrm{d}x$$

Which has a closed form solution (though a bit lengthy for here).

4.6.2 Time preference under model error

This author once watched with a great deal of horror one Laibson [37] at a conference in Columbia University present the idea that having one massage today to two tomorrow, but reversing in a year from nowm is irrational and we need to remedy it with some policy. (For a review of time discounting and intertemporal preferences, see [27], as economists temps to impart what seems to be a varying "discount rate" in a simplified model).

Intuitively, what if I introduce the probability that the person offering the massage is full of balloney? It would clearly make me both prefer immediacy at almost any cost and conditionally on his being around at a future date, reverse the preference. This is what we will model next.

First, time discounting has to have a geometric form, so preference doesn't become negative: linear discounting of the form Ct, where C is a constant ant t is time into the future is ruled out: we need something like C^t or, to extract the rate, $(1+k)^t$ which can be mathematically further simplified into an exponential, by taking it to the continuous time limit. Exponential discounting has the form $e^{-k\, t}$. Effectively, such a discounting method using a shallow model prevents "time inconsistency", so with $\delta < t$:

$$\lim_{t\to\infty} \frac{e^{-k\, t}}{e^{-k\, (t-\delta)}} = e^{-k\, \delta}$$

Now add another layer of stochasticity: the discount parameter, for which we use the symbol λ, is now stochastic.

So we now can only treat $H(t)$ as

$$H(t) = \int e^{-\lambda\, t} \phi(\lambda)\, \mathrm{d}\lambda$$

It is easy to prove the general case that under symmetric stochasticization of intensity $\Delta\lambda$ (that is, with probabilities $\frac{1}{2}$ around the center of the distribution) using the same technique we did in 2.4:

$$H'(t, \Delta\lambda) = \frac{1}{2}\left(e^{-(\lambda-\Delta\lambda)t} + e^{-(\lambda+\Delta\lambda)t}\right)$$

$$\frac{H'(t,\Delta\lambda)}{H'(t,0)} = \frac{1}{2}e^{\lambda t}\left(e^{(-\Delta\lambda-\lambda)t} + e^{(\Delta\lambda-\lambda)t}\right) = \cosh(\Delta\,\lambda t)$$

Where cosh is the cosine hyperbolic function $-$ which will converge to a certain value where intertemporal preferences are flat in the future.

Example: Gamma Distribution. Under the gamma distribution with support in $\mathbb{R}^+$, with parameters α and β, $\phi(\lambda) = \frac{\beta^{-\alpha}\lambda^{\alpha-1}e^{-\frac{\lambda}{\beta}}}{\Gamma(\alpha)}$

we get:

$$H(t,\alpha,\beta) = \int_0^\infty e^{-\lambda\,t}\frac{\left(\beta^{-\alpha}\lambda^{\alpha-1}e^{-\frac{\lambda}{\beta}}\right)}{\Gamma(\alpha)}\,d\lambda = \beta^{-\alpha}\left(\frac{1}{\beta}+t\right)^{-\alpha}$$

so

$$\lim_{t\to\infty}\frac{H(t,\alpha,\beta)}{H(t-\delta,\alpha,\beta)} = 1$$

Meaning that preferences become flat in the future no matter how steep they are in the present, which explains the drop in discount rate in the economics literature.

Further, fudging the distribution and normalizing it, when

$$\phi(\lambda){=}\frac{e^{-\frac{\lambda}{k}}}{k},$$

we get the *normatively obtained* (not empirical pathology) so-called hyperbolic discounting:

$$H(t) = \frac{1}{1+k\,t}$$