Utility Theory and Risk Aversion

John Y. Campbell

Ec2723

September 2010

Outline

- Utility Theory
  - ▶ Basics
  - ▶ Jensen's Inequality

- Risk Aversion
  - ▶ Measurement and Comparison
  - ▶ The Arrow-Pratt Approximation
  - ▶ Declining Absolute Risk Aversion

- Tractable Utility Functions
- Rabin Critique
- Comparing Risks

Ordinal and Cardinal Utility

There is a basic contrast between:

- *Ordinal utility* $Y(x)$ is invariant to monotonic transformations, so $Y(x)$ is equivalent to $\Theta(Y(x))$ for any strictly increasing Θ.

- *Cardinal utility* $\Psi(x)$ is invariant to positive affine (aka linear) transformations, so $\Psi(x)$ is equivalent to $a + b\Psi(x)$ for any $b > 0$.

In finance we rely heavily on von Neumann-Morgenstern utility theory which says that choice over lotteries, satisfying certain axioms, implies maximization of the expectation of a cardinal utility function, defined over outcomes.

Reducing the Dimension of Utility

Finance theory generally works with low-dimensional arguments of the utility function. Proceeding from greater to lesser generality:

- Multiple goods, dates, and states (ordinal utility)
- Multiple goods and dates, taking expectation over states (cardinal utility)
- Multiple dates only (one-good simplification)
- Time-separable utility, adding up over dates ($U(C_1) + \delta U(C_2) + ...$)
- Single-date utility $U(C_1)$, which is equivalent to utility defined over wealth $U(W_1)$.

Jensen's Inequality

Consider a random variable $\widetilde{z}$ and a function f.

Definition: f is *concave* iff for all $\lambda \in [0, 1]$ and values a, b,

$$\lambda f(a) + (1 - \lambda)f(b) \leq f(\lambda a + (1 - \lambda)b).$$

If f is twice differentiable, then concavity implies that $f'' \leq 0$.

Jensen's Inequality: $\mathrm{E}f(\widetilde{z}) \leq f(\mathrm{E}\widetilde{z})$ for all possible $\widetilde{z}$ iff f is concave.

Jensen's Inequality and Risk Aversion

Jensen's Inequality: $\mathrm{E}f(\widetilde{z}) \leq f(\mathrm{E}\widetilde{z})$ for all possible $\widetilde{z}$ iff f is concave.

Definition: an agent is *risk averse* if he dislikes all zero-mean risk at all levels of wealth. That is, for all w_0 and risk $\widetilde{x}$ with $\mathrm{E}\widetilde{x} = 0$,

$$\mathrm{E}u(w_0 + \widetilde{x}) \leq u(w_0).$$

This is equivalent to

$$\mathrm{E}u(\widetilde{z}) \leq u(\mathrm{E}\widetilde{z}),$$

where $\widetilde{z} = w_0 + \widetilde{x}$.

Thus risk aversion is equivalent to concavity of the utility function.

Measuring Risk Aversion

A natural measure of risk aversion is f'', scaled to avoid dependence on the units of measurement for utility. Absolute risk aversion A is defined by

$$A = \frac{-f''}{f'}.$$

Note that in general this is a function of the initial level of wealth.

Comparing Risk Aversion

Let u_1 and u_2 have the same initial wealth. u_1 is more risk-averse than u_2 if u_1 dislikes all lotteries that u_2 dislikes, regardless of the common initial wealth level.

Define $\phi(x) = u_1(u_2^{-1}(x))$. What are the properties of this function?

1. $u_1(z) = \phi(u_2(z))$, so $\phi(.)$ turns u_2 into u_1.

2. $u_1'(z) = \phi'(u_2(z))u_2'(z)$, so $\phi' = u_1'/u_2' > 0$.

3. $u_1''(z) = \phi'(u_2(z))u_2''(z) + \phi''(u_2(z))u_2'(z)^2$, so

$$\phi'' = \frac{u_1'' - \phi' u_2''}{u_2'^2} = \frac{u_1'}{u_2'^2}(A_2 - A_1).$$

Comparing Risk Aversion

4. Consider a risk $\widetilde{x}$ that is disliked by u_2, that is a risk s.t.
$\mathrm{E}u_2(w_0 + \widetilde{x}) \leq u(w_0)$. We have

$$\mathrm{E}u_1(w_0 + \widetilde{x}) = \mathrm{E}\phi(u_2(w_0 + \widetilde{x})) \leq \phi(\mathrm{E}u_2(w_0 + \widetilde{x})) \leq \phi(u_2(w_0)) = u_1(w_0)$$

for all $\widetilde{x}$ iff ϕ is concave or equivalently $\phi'' \leq 0$. But then from property 3
we must have $A_1 \geq A_2$.

Comparing Risk Aversion

5. The *risk premium* π is the amount one is willing to pay to avoid a pure (zero-mean) risk. It solves

$$\mathrm{E}u(w_0 + \widetilde{x}) = u(w_0 - \pi).$$

Defining $z = w_0 - \pi$ and $\widetilde{y} = \pi + \widetilde{x}$, this can be rewritten as

$$\mathrm{E}u(z + \widetilde{y}) = u(z).$$

Now define π_2 as the risk premium for agent 2, and define z_2 and $\widetilde{y}_2$ accordingly. We have

$$\mathrm{E}u_2(z_2 + \widetilde{y}_2) = u_2(z_2).$$

If u_1 is more risk-averse than u_2, then

$$\mathrm{E}u_1(z_2 + \widetilde{y}_2) \leq u_1(z_2),$$

which implies $\pi_1 \geq \pi_2$.

Comparing Risk Aversion

6. Consider a risk that may have a nonzero mean μ. It pays $\mu + \widetilde{x}$ where $\widetilde{x}$ has zero mean. The *certainty equivalent* C^e satisfies

$$\mathrm{E}u(w_o + \mu + \widetilde{x}) = u(w_0 + C^e).$$

This implies that

$$C^e(w_0, u, \mu + \widetilde{x}) = \mu - \pi(w_0 + \mu, u, \widetilde{x}).$$

Thus if u_1 is more risk-averse than u_2, then $C_1^e \leq C_2^e$.

John Y. Campbell (Ec2723)   Utility Theory and Risk Aversion   September 2010   11 / 35

Comparing Risk Aversion

In summary, the following statements are equivalent:

- u_1 is more risk-averse than u_2.
- u_1 is a concave transformation of u_2.
- $A_1 \geq A_2$.
- $\pi_1 \geq \pi_2$.
- $C_1^e \leq C_2^e$.

The Arrow-Pratt Approximation

Consider a pure risk $\widetilde{y} = k\widetilde{x}$, where k is a scale factor. Write the risk premium as a function $g(k)$: $g(k) = \pi(w_0, u, k\widetilde{x})$ satisfies

$$Eu(w_0 + k\widetilde{x}) = u(w_0 - g(k)).$$

Note that $g(0) = 0$.
Now differentiate w.r.t. k:

$$E\widetilde{x}u'(w_0 + k\widetilde{x}) = -g'(k)u'(w_0 - g(k)).$$

Since $E\widetilde{x} = 0$, this implies that $g'(0) = 0$.

The Arrow-Pratt Approximation

Differentiate w.r.t. k a second time:

$$\mathrm{E}\widetilde{x}^2 u''(w_o + k\widetilde{x}) = g'(k)^2 u''(w_0 - g(k)) - g''(k)u'(w_0 - g(k)),$$

which implies that

$$g''(0) = \frac{-u''(w_0)}{u'(w_0)}\mathrm{E}\widetilde{x}^2 = A(w_0)\mathrm{E}\widetilde{x}^2.$$

The Arrow-Pratt Approximation

Now take a Taylor approximation of $g(k)$:

$$g(k) \approx g(0) + kg'(0) + \frac{1}{2}k^2 g''(0).$$

Substituting in, we get

$$\pi \approx \frac{1}{2}A(w_0)k^2 \mathrm{E}\widetilde{x}^2 = \frac{1}{2}A(w_0)\mathrm{E}\widetilde{y}^2.$$

The risk premium is proportional to the *square* of the risk. This property of differentiable utility is known as *second-order risk aversion*. It implies that people are approximately risk-neutral with respect to small risks. We also find that

$$C^e \approx k\mu - \frac{1}{2}A(w_0)k^2 \mathrm{E}\widetilde{x}^2,$$

so a positive mean has a dominant effect for small risks.

Relative Risks

Define a multiplicative risk by $\widetilde{w} = w_0(1 + k\widetilde{x}) = w_0(1 + \widetilde{y})$. Define $\widehat{\pi}$ as the share of wealth one would pay to avoid this risk:

$$\widehat{\pi} = \frac{\pi(w_0, u, w_0 k\widetilde{x})}{w_0}.$$

Then

$$\widehat{\pi} \approx \frac{1}{2} w_0 A(w_0) k^2 \mathrm{E}\widetilde{x}^2 = \frac{1}{2} R(w_0) \mathrm{E}\widetilde{y}^2,$$

where $R(w_0) = w_0 A(w_0)$ is the *coefficient of relative risk aversion*.

Decreasing Absolute Risk Aversion

The following conditions are equivalent:

- π is decreasing in w_0.
- $A(w_0)$ is decreasing in w_0.
- $-u'$ is a concave transformation of u, so $-u'''/u'' \geq -u''/u'$ everywhere. The ratio $-u'''/u'' = P$ has been called *absolute prudence* by Kimball, who relates it to the theory of precautionary saving.

Decreasing absolute risk aversion (DARA) seems intuitively appealing. Certainly we should be uncomfortable with increasing absolute risk aversion.

Tractable Utility Functions

Almost all applied theory and empirical work in finance uses some member of the class of linear risk tolerance (LRT) or hyperbolic absolute risk aversion (HARA) utility functions. These are defined by

$$u(z) = \zeta(\eta + \frac{z}{\gamma})^{1-\gamma},$$

defined over z s.t. $\eta + z/\gamma > 0$.

For these utility functions, we get

$$T(z) = \frac{1}{A(z)} = \eta + \frac{z}{\gamma},$$

which is linear in z, and

$$A(z) = \left(\eta + \frac{z}{\gamma}\right)^{-1},$$

which is hyperbolic in z.

Important Special Cases

- *Quadratic* utility has $\gamma = -1$. This implies increasing absolute risk aversion and the existence of a bliss point at which $u' = 0$. These are important disadvantages, although quadratic utility is tractable in models with additive risk.

- *Exponential or constant absolute risk averse* (CARA) utility is the limit as $\gamma \rightarrow -\infty$. To obtain constant absolute risk aversion A, we need

$$-u''(z) = Au'(z).$$

Solving this differential equation, we get

$$u(z) = \frac{-\exp(-Az)}{A},$$

where $A = 1/\eta$. This form of utility is tractable with normally distributed risks because then utility is lognormally distributed.

Important Special Cases

- *Power or constant relative risk averse* (CRRA) utility has $\eta = 0$ and $\gamma > 0$. Relative risk aversion $R = \gamma$. For $\gamma \neq 1$,

$$u(z) = \frac{z^{1-\gamma}}{1-\gamma}.$$

For $\gamma = 1$,

$$u(z) = \log(z).$$

Power utility is appealing because it implies stationary risk premia and interest rates even in the presence of long-run economic growth. Also it is tractable in the presence of multiplicative lognormally distributed risks.

- A negative η represents a *subsistence level*. Rubinstein has argued for this model, but economic growth renders any fixed subsistence level irrelevant in the long run. Models of habit formation have time-varying subsistence levels which can grow with the economy.

Rabin Critique

Matthew Rabin (2000) has argued that standard utility theory cannot explain observed aversion to small gambles without implying ridiculous aversion to large gambles. This follows from the fact that differentiable utility has second-order risk aversion.

To understand Rabin's critique, consider a gamble that wins \$11 with probability $1/2$, and loses \$10 with probability $1/2$. With diminishing marginal utility, the utility of the win is at least $11u'(w_0 + 11)$. The utility cost of the loss is at most $10u'(w_0 - 10)$. Thus if a person turns down this gamble, we must have $10u'(w_0 - 10) > 11u'(w_0 + 11)$ which implies

$$\frac{u'(w_0 + 11)}{u'(w_0 - 10)} < \frac{10}{11}.$$

Rabin Critique

$$\frac{u'(w_0 + 11)}{u'(w_0 - 10)} < \frac{10}{11}.$$

Now suppose the person turns down the same gamble at an initial wealth level of $w_0 + 21$. Then

$$\frac{u'(w_0 + 21 + 11)}{u'(w_0 + 21 - 10)} = \frac{u'(w_0 + 32)}{u'(w_0 + 11)} < \frac{10}{11}.$$

Combining these two inequalities,

$$\frac{u'(w_0 + 32)}{u'(w_0 - 10)} < \left(\frac{10}{11}\right)^2 = \frac{100}{121}.$$

If this iteration can be repeated, it implies extremely small marginal utility at high wealth levels, which would induce people to turn down apparently extremely attractive gambles.

Risk Aversion and Expected-Utility Theory: A Calibration Theorem

Matthew Rabin

TABLE I

IF AVERSE TO 50-50 LOSE \$100 / GAIN g BETS FOR ALL WEALTH LEVELS,
WILL TURN DOWN 50-50 LOSE L / GAIN G BETS; G'S ENTERED IN TABLE.

		g		
L	\$101	\$105	\$110	\$125
\$400	400	420	550	1,250
\$600	600	730	990	∞
\$800	800	1,050	2,090	∞
\$1,000	1,010	1,570	∞	∞
\$2,000	2,320	∞	∞	∞
\$4,000	5,750	∞	∞	∞
\$6,000	11,810	∞	∞	∞
\$8,000	34,940	∞	∞	∞
\$10,000	∞	∞	∞	∞
\$20,000	∞	∞	∞	∞

TABLE II

TABLE I REPLICATED, FOR INITIAL WEALTH LEVEL $290,000, WHEN l/g BEHAVIOR IS ONLY KNOWN TO HOLD FOR $w \leq$ $300,000.

L	$101	$105	g $110	$125
$400	400	420	550	1,250
$600	600	730	990	36,000,000,000
$800	800	1,050	2,090	90,000,000,000
$1,000	1,010	1,570	718,190	160,000,000,000
$2,000	2,320	69,930	12,210,880	850,000,000,000
$4,000	5,750	635,670	60,528,930	9,400,000,000,000
$6,000	11,510	1,557,360	180,000,000	89,000,000,000,000
$8,000	19,290	3,058,540	510,000,000	830,000,000,000,000
$10,000	27,780	5,503,790	1,300,000,000	7,700,000,000,000,000
$20,000	85,750	71,799,110	160,000,000,000	540,000,000,000,000,000,000

TABLE III

If a Person Has CARA Utility Function and Is Averse to 50/50 Lose l/Gain g Bets for all Wealth Levels, then (i) She Has Coefficient of Absolute Risk Aversion No Smaller than ρ and (ii) Invests X in the Stock Market when Stock Yields Are Normally Distributed with Mean Real Return 6.4% and Standard Deviation 20%, and Bonds Yield a Riskless Return of 0.5%.

l/g	ρ	X
$100/$101	.0000990	$14,899
$100/$105	.0004760	$3,099
$100/$110	.0009084	$1,639
$100/$125	.0019917	$741
$100/$150	.0032886	$449
$1,000/$1,050	.0000476	$30,987
$1,000/$1,100	.0000908	$16,389
$1,000/$1,200	.0001662	$8,886
$1,000/$1,500	.0003288	$4,497
$1,000/$2,000	.0004812	$3,067
$10,000/$11,000	.0000090	$163,889
$10,000/$12,000	.0000166	$88,855
$10,000/$15,000	.0000328	$44,970
$10,000/$20,000	.0000481	$30,665

Responses to Rabin Critique

1. As we increase wealth, a person will continue to turn down a given absolute gamble indefinitely only if absolute risk aversion is constant or increasing. Rabin's most extreme results assume this, and can be understood as a critique of constant or increasing absolute risk aversion.

2. Observed aversion to small risks probably results from some other aspect of human psychology besides declining marginal utility of wealth. But this does not necessarily mean that we should abandon utility theory for studying large risks in financial markets.

Responses to Rabin Critique

3. What does explain the observed aversion to small risks? One simple story would be that people suffer fixed disutility from any loss, regardless of size. More attractive are nonstandard preference models with "first-order risk aversion", in which the premium paid to avoid a small loss is proportional to the loss, not the squared loss. Such models have a kink in the utility function at the initial level of wealth. Kahneman and Tversky's (1979) prospect theory has this feature.

Responses to Rabin Critique

4. Barberis, Huang, and Thaler (2006) point out that even first-order risk aversion cannot generate substantial aversion to small delayed gambles. During the time between the decision to take a gamble and the resolution of uncertainty, the agent will be exposed to other risks and will merge these with the gamble under consideration. If the gamble is uncorrelated with the other risks, it is diversifying. In effect the agent will have second-order risk aversion with respect to delayed gambles. To deal with this problem, Barberis et al. argue that people treat gambles in isolation, that is, they use "narrow framing".

5. Chetty and Szeidl (2007) show that "consumption commitments" (fixed costs to adjust a portion of consumption) raise risk aversion over small gambles, relative to risk aversion over large gambles where extreme outcomes would justify paying the cost to adjust all consumption.

CONSUMPTION COMMITMENTS AND RISK PREFERENCES*

RAJ CHETTY AND ADAM SZEIDL

Many households devote a large fraction of their budgets to "consumption commitments"—goods that involve transaction costs and are infrequently adjusted. This paper characterizes risk preferences in an expected utility model with commitments. We show that commitments affect risk preferences in two ways: (1) they amplify risk aversion with respect to moderate-stake shocks, and (2) they create a motive to take large-payoff gambles. The model thus helps resolve two basic puzzles in expected utility theory: the discrepancy between moderate-stake and large-stake risk aversion and lottery playing by insurance buyers. We discuss applications of the model such as the optimal design of social insurance and tax policies, added worker effects in labor supply, and portfolio choice. Using event studies of unemployment shocks, we document evidence consistent with the consumption adjustment patterns implied by the model.

TABLE I
Expenditure Shares and Adjustment Frequencies by Consumption Category

Consumption category	Share of total expenditures	Fraction of households actively reducing consumption
Shelter (%)	22.2	8.7
Cars (excluding gas + maint) (%)	14.7	10.6
Apparel (%)	5.1	0.3
Furniture/appliances (%)	4.4	0.5
Health insurance (%)	3.0	32.9
Food and alcohol (%)	18.1	42.8
Utilities (%)	8.2	45.8
Other transportation (%)	7.3	49.2
Entertainment (%)	6.1	48.7
Out-of-pocket health (%)	3.0	47.8
Education (%)	2.0	45.2
Housing operations (%)	1.9	44.3
Personal care (%)	1.0	41.0
Tobacco (%)	0.9	36.6
Reading materials (%)	0.6	45.7
Miscellaneous (%)	1.5	39.3

Note: First column in table shows aggregate expenditure shares for consumption categories in the CEX, following methodology described in the Data Appendix. Second column reports fraction of households who actively reduce consumption (beyond depreciation) of each category from first quarter to last quarter in CEX. For apparel and furniture, households that reduce consumption are defined as those with negative net expenditures. For all other categories, households that reduce expenditure are those with negative nominal growth ($g_{cit} < 0$). See text and Data Appendix for the definition of g_{cit}. Categories above dotted line are classified as "commitments" by frequency-of-adjustment definition.

FIGURE II
Event Study of Consumption Around Unemployment Shocks
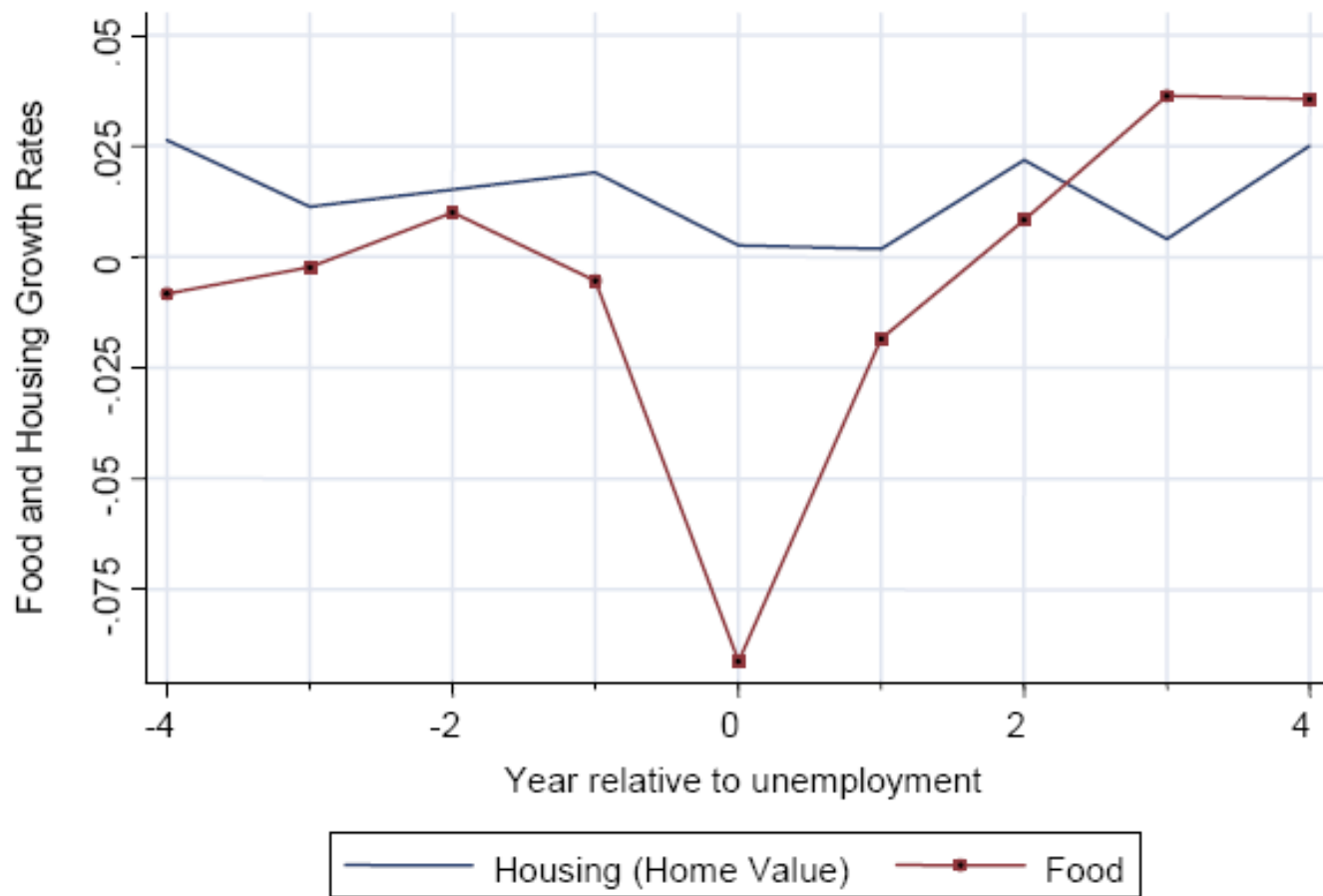
a. Renters

b. Homeowners

FIGURE IIIa
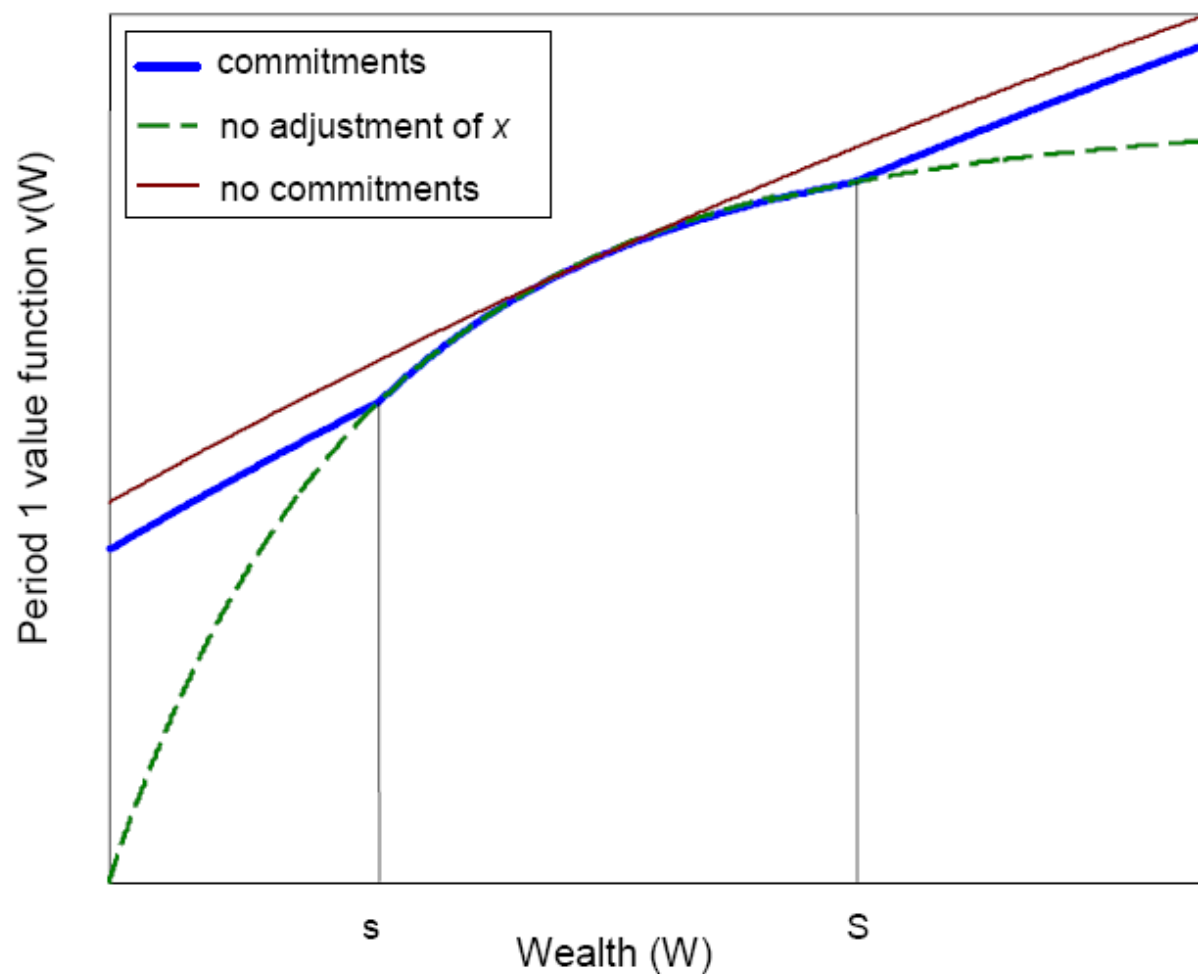Effect of Commitments on Value Function
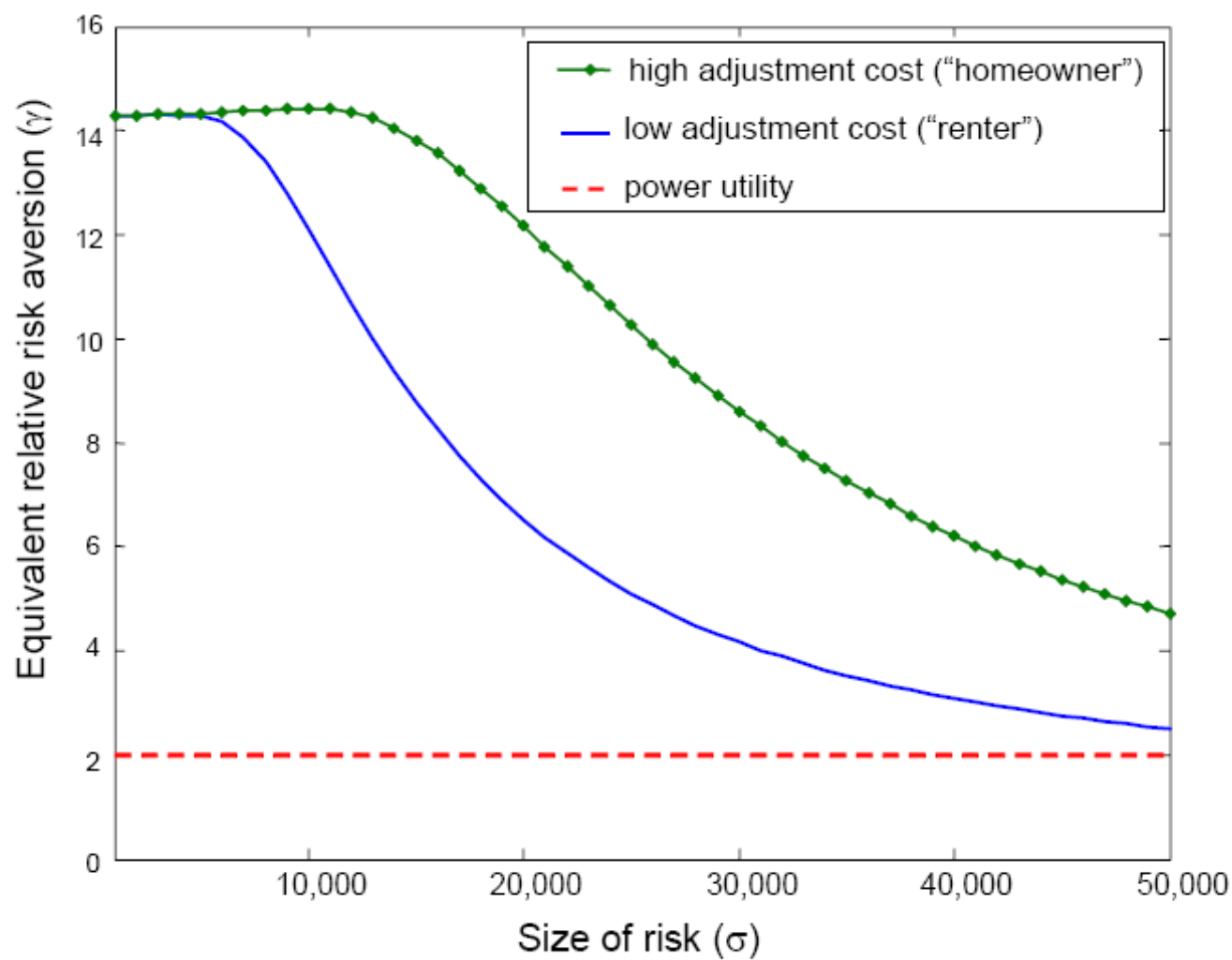
FIGURE IV
Risk Aversion Over Small versus Large Risks

TABLE II
Risk Preferences over Small and Large Stakes: Calibration Results

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	$g = 1025$		$g = 1050$		$g = 1100$		$g = 1250$	
Loss (L)	CRRA	Commit	CRRA	Commit	CRRA	Commit	CRRA	Commit
100	100.2	100.2	100.5	100.5	101	101	102	102
200	201	201	204	202	204	204	208	208
500	506	506	524	512	524	524	555	555
1000	*1025*	*1025*	*1050*	*1050*	*1100*	*1100*	*1250*	*1250*
2000	2103	2103	2211	2211	2446	2446	3373	3370
5000	5695	5695	6573	6570	9388	9359	∞	14879
10000	13243	13235	19615	19387	∞	23388	∞	23841
15000	23805	23732	98027	29741	∞	30149	∞	30290
20000	39975	34003	∞	36660	∞	37005	∞	37239
25000	69671	43003	∞	43860	∞	44388	∞	44746
30000	175622	50354	∞	51591	∞	52360	∞	52884
40000	∞	66569	∞	68897	∞	70372	∞	71393
50000	∞	85166	∞	89184	∞	91789	∞	93623
Curvature over food (γ_f)		3.7		7.1		13.6		29.9
ERRA for L = 50000 gamble		2.47		2.63		2.72		2.78

A special case. The following specification of utility yields a simple expression that is useful in calibrating the effect of commitments on risk aversion:

$$(6) \qquad u(f, x) = \frac{f^{1-\gamma_f}}{1 - \gamma_f} + \mu \cdot \frac{x^{1-\gamma_x}}{1 - \gamma_x}.$$

From (3), the ratio of curvatures over wealth with and without adjustment costs at $\overline{W}$ is

$$(7) \qquad \frac{\gamma^c}{\gamma^n} = 1 + \frac{x}{f} \frac{\gamma_f}{\gamma_x}.$$

Equation (7) shows that the commitment share of the budget is a key factor in determining how much commitments magnify risk aversion over small shocks. When commitments constitute a higher share of expenditures, shocks are concentrated on a smaller set of goods, and risk aversion is higher. When $\gamma_f > \gamma_x$, the consumer is particularly risk averse over adjustable goods, increasing the amplification effect.

Comparing Risks

We would like to compare the riskiness of different distributions. Three possible notions of increasing risk:

1) Something that all concave utility functions dislike.
2) More weight in the tails of the distribution.
3) Added noise.

The classic analysis of Rothschild and Stiglitz shows that these are all equivalent. They are *not* equivalent to higher variance.

Comparing Risks

Consider random variables $\widetilde{X}$ and $\widetilde{Y}$ which have the same expectation.

1) $\widetilde{X}$ is weakly less risky than $\widetilde{Y}$ if no individual with a concave utility function prefers $\widetilde{Y}$ to $\widetilde{X}$:

$$E\left[u(\widetilde{X})\right] \geq E\left[u(\widetilde{Y})\right]$$

for all concave u (.). $\widetilde{X}$ is less risky (without qualification) if there is *some* concave $u(.)$ which strictly prefers $\widetilde{X}$.

Note that this is a partial ordering. It is not the case that for any $\widetilde{X}$ and $\widetilde{Y}$, either $\widetilde{X}$ is weakly less risky than $\widetilde{Y}$ or $\widetilde{Y}$ is weakly less risky than $\widetilde{X}$. We can get a complete ordering if we restrict attention to a smaller class of utility functions than the concave, such as the quadratic.

Comparing Risks

2) $\widetilde{X}$ is less risky than $\widetilde{Y}$ if the density function of $\widetilde{Y}$ can be obtained from that of $\widetilde{X}$ by applying a mean-preserving spread (MPS). An MPS $s(x)$ is defined by

$$s\left(x\right) = \left(\begin{array}{c} \alpha \text{ for } c < x < c + t \\ -\alpha \text{ for } c' < x < c' + t \\ -\beta \text{ for } d < x < d + t \\ \beta \text{ for } d' < x < d' + t \\ 0 \text{ elsewhere} \end{array}\right)$$

where $\alpha, \beta, t > 0$; $c + t < c' < d - t < d + t < d'$; and

$\alpha\left(c' - c\right) = \beta(d' - d)$, that is, "the more mass you move, the less far you can move it."

Comparing Risks

An MPS is something you add to a density function $f(x)$. If $g(x) = f(x) + s(x)$, then (i) $g(x)$ is also a density function, and (ii) it has the same mean as $f(x)$.

(i) is obvious because $\int s(x)\, dx =$ area under $s(x) = 0$.

(ii) follows from the fact that the "mean" of $s(x)$, $\int x s(x) = 0$, which follows from $\alpha(c' - c) = \beta(d' - d)$.

In what sense is an MPS a spread? It's obvious that if the mean of $f(x)$ is between $c\prime + t$ and d, then $g(x)$ has more weight in the tails. It's not so obvious when the mean of $f(x)$ is off to the left or the right. Nevertheless, we can show that $\widetilde{Y}$ with density g is riskier than $\widetilde{X}$ with density f in the sense of 1) above. In this sense the term "spread" is appropriate.

Comparing Risks

3) A formal definition of "added noise" is that $\widetilde{X}$ is less risky than $\widetilde{Y}$ if $\widetilde{Y}$ has the same distribution as $\widetilde{X} + \widetilde{\varepsilon}$, where $E\left[\widetilde{\varepsilon}|X\right] = 0$ for all values of X. We say that $\widetilde{\varepsilon}$ is a "fair game" with respect to X.

This condition is stronger than zero covariance, $\text{Cov}\left(\widetilde{\varepsilon}, \widetilde{X}\right) = 0$. It is weaker than independence, $\text{Cov}\left(f\left(\widetilde{\varepsilon}\right), g\left(\widetilde{X}\right)\right) = 0$ for all functions f and g. It is equivalent to $\text{Cov}\left(\widetilde{\varepsilon}, g\left(\widetilde{X}\right)\right) = 0$ for all functions g.

Comparing Risks

3) is sufficient for 1):

$$
\begin{aligned}
E\left[U(\widetilde{X}+\widetilde{\varepsilon})|X\right] &\leq U(E\left[\widetilde{X}+\widetilde{\varepsilon}|X\right]) = U(X) \\
&\Rightarrow E\left[U(\widetilde{X}+\widetilde{\varepsilon})\right] \leq E\left[U(\widetilde{X})\right] \\
&\Rightarrow E\left[U(\widetilde{Y})\right] \leq E\left[U(\widetilde{X})\right]
\end{aligned}
$$

because $\widetilde{Y}$ and $\widetilde{X}+\widetilde{\varepsilon}$ have the same distribution.

In fact, Rothschild-Stiglitz show that 1), 2) and 3) are all equivalent. This is a powerful result because one or the other condition may be most useful in a particular application.

Comparing Risks

Why are these not equivalent to $\widetilde{Y}$ having greater variance than $\widetilde{X}$? It is obvious from 3) that if $\widetilde{Y}$ is riskier than $\widetilde{X}$ then $\widetilde{Y}$ has greater variance than $\widetilde{X}$. The problem is that the reverse is not true in general. Greater variance is necessary but not sufficient for increased risk. $\widetilde{Y}$ could have greater variance than $\widetilde{X}$ but still be preferred by some concave utility functions if it has more desirable higher-moment properties. This possibility can only be eliminated if we confine attention to a limited class of distributions such as the normal distribution.

Stochastic Dominance

The following are definitions.

$\widetilde{X}$ *dominates* $\widetilde{Y}$ if $\widetilde{Y} = \widetilde{X} + \widetilde{\zeta}$, where $\widetilde{\zeta} \leq 0$.

$\widetilde{X}$ *first-order stochastically dominates* $\widetilde{Y}$ if $\widetilde{Y}$ has the distribution of $\widetilde{X} + \widetilde{\zeta}$, where $\widetilde{\zeta} \leq 0$. Equivalently, if $F(.)$ is the cdf of $\widetilde{X}$ and $G(.)$ is the cdf of $\widetilde{Y}$, then $\widetilde{X}$ first-order stochastically dominates $\widetilde{Y}$ if $F(z) \leq G(z)$ for every z. First-order stochastic dominance implies that every increasing utility function will prefer $\widetilde{X}$.

$\widetilde{X}$ *second-order stochastically dominates* $\widetilde{Y}$ if $\widetilde{Y}$ has the distribution of $\widetilde{X} + \widetilde{\zeta} + \widetilde{\varepsilon}$, where $\widetilde{\zeta} \leq 0$ and $\mathrm{E}\left[\widetilde{\varepsilon}|X + \zeta\right] = 0$. Second-order stochastic dominance implies that every increasing, concave utility function will prefer $\widetilde{X}$. Increased risk is the special case of second-order stochastic dominance where $\widetilde{\zeta} = 0$.

Application: The Principle of Diversification

Consider n lotteries with payoffs $\widetilde{x}_1$, $\widetilde{x}_2$, ..., $\widetilde{x}_n$ that are independent and identically distributed (iid). You are asked to choose weights α_1, α_2, ..., α_n subject to the constraint that $\sum_i \alpha_i = 1$. It seems obvious that the best choice is complete diversification, with weights $\alpha_i = 1/n$ for all i. The payoff is then

$$\widetilde{z} = \frac{1}{n} \sum_{i=1}^{n} \widetilde{x}_i.$$

Application: The Principle of Diversification

To prove that this is optimal, note that the payoff on any other strategy is

$$\sum_i \alpha_i \widetilde{x}_i = \widetilde{z} + \sum_i \left(\alpha_i - \frac{1}{n} \right) \widetilde{x}_i = \widetilde{z} + \widetilde{\varepsilon},$$

and

$$\mathrm{E}\left[\widetilde{\varepsilon} | z \right] = \sum_i \left(\alpha_i - \frac{1}{n} \right) \mathrm{E}\left[\widetilde{x}_i | z \right] = k \sum_i \left(\alpha_i - \frac{1}{n} \right) = 0.$$

Note that diversification increases utility because you are averaging independent risks. Adding independent risks does not have the same effect.