Lecture Note 14: Uncertainty, Expected Utility Theory and the Market for Risk

David Autor, Massachusetts Institute of Technology

14.03/14.003, Microeconomic Theory and Public Policy, Fall 2010

1 Risk Aversion and Insurance: Introduction

- A significant hole in our theory of consumer choice developed in 14.03/14.003 to date is that we have only modeled choices that are devoid of uncertainty. That's convenient, but not particularly plausible.
  - Prices change
  - Income fluctuates
  - Bad stuff happens
- Most decisions are forward-looking, and these decisions depend on our beliefs about what is the optimal plan for present and future. Inevitably, such choices are made in a context of uncertainty. There is a risk (in fact, a likelihood) that not all scenarios we hoped for will be borne out. In making plans, we should take these contingencies and probabilities into account. If we want a realistic model of choice, we need to model how uncertainty affects choice and well-being.
- This model should help to explain:
  - How people choose among 'bundles' that have uncertain payoffs, e.g., whether to fly on an airplane, whom to marry.
  - Insurance: Why do people want to buy it.
  - How (and why) the market for risk operates.

1.1 A few motivating examples

1. People don't seem to want to play actuarially fair games. Such a game is one in which the cost of entry is equal to the expected payoff:

$$E(X) = P_{win} \cdot [\text{Payoff}|\text{Win}] + P_{lose} \cdot [\text{Payoff}|\text{Lose}] .$$

   - Most people would not enter into a $1,000 dollar heads/tails fair coin flip.

2. People won't necessarily play actuarially *favorable* games:

   - You are offered a gamble. We'll flip a coin. If it's heads, I'll give you $10 million dollars. If it's tails, you owe me $9 million.
     Its expected monetary value is :
     $$\frac{1}{2} \cdot 10 - \frac{1}{2} \cdot 9 = \$0.5 \text{ million}$$
     Want to play?

3. People won't pay large amounts of money to play gambles with huge upside potential. Example "St. Petersburg Paradox."

   - Flip a coin. I'll pay you in dollars 2^n, where n is the number of tosses until you get a head:
     $$X_1 = \$2, X_2 = \$4, X_3 = \$8, ...X_n = 2^n.$$
   - What is the expected value of this game?
     $$E(X) = \frac{1}{2}2 + \frac{1}{4}4 + \frac{1}{8}8 + ...\frac{1}{2^n}2^n = \infty.$$
   - How much *would* you be willing to pay to play this game? [People generally do not appear willing to pay more than a few dollars to play this game.]
   - What is the variance of this gamble? $V(X) = \infty$.

- The fact that a gamble with infinite expected monetary value has (apparently) limited 'utility value' suggests something pervasive and important about human behavior: *As a general rule, uncertain prospects are worth less in utility terms than certain ones, even when expected tangible payoffs are the same.*

- We need to be able to say how people make choices when:

  - consumers value outcomes (as we have modeled all along)
  - consumers also have feelings/preferences about the riskiness of those outcomes

2 Five Simple Statistical Notions

Definition 1 *Probability distribution*

Define states of the world $1, 2...n$ *with probability of occurrence* $\pi_1, \pi_2...\pi_n$.
A valid probability distribution satisfies:

$$\sum_{i=1}^{n} \pi_i = 1, \text{ or } \int_{-\infty}^{\infty} f(x)\partial x = 1 \text{ and } f(x) \geq 0 \ \forall \ x.$$

In this equation $f(x)$ *is the 'probability density function' (PDF) of the continuous random variable* x, *meaning that* $f(x)$ *is essentially the probability of randomly drawing the given value* x *(so,* $f(x)$ *is just like the* π_i *in the discreet case). [Note that the probability of drawing any specific value from a continuous distribution is zero since there are an infinite number of possibilities. Depending on the distribution, however, some ranges of values will be much more probable than others.]*

Definition 2 *Expected value or "expectation" (something we've used informally all semester): Say each state* i *has payoff* x_i. *Then*

$$E(x) = \sum_{i=1}^{n} \pi_i x_i \text{ or } E(x) = \int_{-\infty}^{\infty} x f(x)\partial x.$$

Example: Expected value of a fair dice roll is $E(x) = \sum_{i=1}^{6} \pi_i i = \frac{1}{6} \cdot 21 = \frac{7}{2}$.

Definition 3 *Variance (dispersion)*
Gambles with the same expected value may have different dispersion.
We'll measure dispersion with variance.

$$V(x) = \sum_{i=1}^{n} \pi_i (x_i - E(x))^2 \text{ or } V(x) = \int_{-\infty}^{\infty} (x - E(x))^2 f(x)\partial x.$$

In dice example, $V(x) = \sum_{i=1}^{n} \pi_i \left(i - \frac{7}{2}\right)^2 = 2.92$.

Definition 4 *Independence.*
A case in which the probabilities of two (or multiple) outcomes do not depend upon one another. If events A *and* B *are independent, then* $\Pr(A \text{ and } B) = \Pr(A) \cdot \Pr(B)$, *and similarly,* $E[A \cdot B] = E[A] \cdot E[B]$ *.The probability of flipping two sequential heads with a fair coin is* $\Pr(H \text{ and } H) = \Pr(H) \cdot \Pr(H) = 0.25$. *These probabilities are independent. The probabilities of hearing thunder and seeing lighting are not independent.*

Note (without proof): The variance of n identical *independent* gambles is $\frac{1}{n}$ times the variance of one of the gambles.

Definition 5 *Law of Large Numbers*
In repeated, independent trials with the same probability p of success in each trial, the chance that the percentage of successes differs from the probability p by more than a fixed positive amount $e > 0$ converges to zero as number of trials n goes to infinity for every positive e.

Dispersion and risk are closely related notions. Holding constant the expectation of X, more dispersion means that the outcome is "riskier" – it has both more upside and more downside potential.

Note (without proof): The variance of n identical *independent* gambles is $\frac{1}{n}$ times the variance of one of the gambles.

Consider four gambles:

1. \$0.50 for sure. $V(L_1) = 0$.

2. Heads you receive \$1.00, tails you receive 0.
   $V(L_2) = 1 \times [0.5 \times (1 - .5)^2 + 0.50 \times (0 - .5)^2] = 0.25$

3. 4 independent flips of a coin, you receive \$0.25 on each head.
   $V(L_3) = 0.25 \times \frac{1}{4}[0.5 \times (1 - .5)^2 + 0.50 \times (0 - .5)^2]) = 0.062\,5$

4. 100 independent flips of a coin, you receive \$0.01 on each head.
   $V(L_4) = 0.01 \times \frac{1}{100}[0.5 \times (1 - .5)^2 + 0.5 \times (0 - .5)^2] = 0.000025$

All four of these "lotteries" have same expected value (50 cents), but they have different levels of risk.

3 Risk preference and expected utility theory[1]

[This section derives the Expected Utility Theorem. I will not teach this material, and will not hold you responsible for the technical details.]

[1] This section draws on Mas-Colell, Andreu, Michael D. Winston and Jerry R. Green, *Microeconomic Theory*, New York: Oxford University Press, 1995, chapter 6. For those of you considering Ph.D. study in economics, MWG is the only single text that covers almost the entire corpus of modern microeconomic theory. It is the Oxford English Dictionary of modern economic theory: most economists keep it on hand for reference; few read it for pleasure.

3.1 Description of risky alternatives

- Let us imagine that a decision maker faces a choice among a number of risky alternatives. Each risky alternative may result in one of a number of possible *outcomes*, but which outcome occurs is uncertain at the time of choice.

- Let an outcome be a monetary payoff or consumption bundle.

- Assume that the number of possible outcomes is finite, and index these outcomes by $n = 1, ..., N$.

- Assume further that the probabilities associated with each outcome are *objectively known.* Example: risky alternatives might be monetary payoffs from the spin of a roulette wheel.

- The basic building block of our theory is the concept of a *lottery.*

Definition 6 *A simple lottery L is a list $L = (p_1, ...p_N)$ with $p_n \geq 0$ for all n and $\Sigma_n p_n = 1$, where p_n is interpreted as the probability of outcome n occurring.*

- In a simple lottery, the outcomes that may result are certain.

- A more general variant of a lottery, known as a *compound lottery,* allows the outcomes of a lottery to themselves be simple lotteries.

Definition 7 *Given K simple lotteries $L_k = \left(p_1^k, ...p_N^k\right)$, $k = 1..., K$, and probabilities $\alpha_k \geq 0$ with $\Sigma_k \alpha_k = 1$, the compound lottery $(L_1, ...L_K; \alpha_1, ...\alpha_K)$, is the risky alternative that yields the simple lottery L_k with probability α_k for $k = 1, ..., K$.*

- For any compound lottery $(L_1, ...L_K; \alpha_1, ...\alpha_K)$, we can calculate a corresponding *reduced lottery* as the simple lottery $L = (p_1, ...p_N)$ that generates the same ultimate distribution over outcomes. So, the probability of outcome n in the reduced lottery is:

$$p_n = \alpha_1 p_n^1 + \alpha_2 p_n^2 + ... + \alpha_k p_n^k.$$

That is, we simply add up the probabilities, p_n^k, of each outcome n in all lotteries k, multiplying each p_n^k by the probability α_k of facing each lottery k.

3.2 Preferences over lotteries

- We now study the decision maker's *preferences over lotteries.*
- The basic premise of the model that follows is what philosophers would call a 'consequentialist' premise: for any risky alternative, the decision maker cares only about the *reduced lottery* over final outcomes. The decision maker effectively is indifferent to the (possibly many) compound lotteries underlying these reduced lotteries.
- This compound lottery premise (which could also be called an axiom) states the 'frame' or order of lotteries is unimportant. So consider a two stage lottery is follows:
  - Stage 1: Flip a coin heads, tails.
  - Stage 2:
    If it's heads, flip again. Heads yields \$1.00, tails yields \$0.75.
    If it's tails, roll a dice with payoffs $\$0.10, \$0.20, ...\$0.60$ corresponding to outcomes $1 - 6$.
- Now consider a single state lottery, where:
  - We spin a pointer on a wheel with 8 areas, 2 areas of 90^0 representing \$1.00, and \$0.75, and 6 areas of 30^0 each, representing $\$0.10, \$0.20, ...\$0.60$ each.
  - This single stage lottery has the same payouts at the same odds as the $2-$stage lottery.
  - The 'compound lottery' axiom says the consumer is indifferent between these two.
  - Counterexamples? [This is not an innocuous set of assumptions.]
  - [Is this realistic? Hard to develop intuition on this point, but much research shows that this assumption is often violated.]
- Continuing with the theory... Now, take the set of alternatives the decision maker faces, denoted by $\pounds$ to be the set of all simple lotteries over possible outcomes N.
- We assume the consumer has a rational preference relation $\succsim$ on $\pounds$, a *complete* and *transitive* relation allowing comparison among any pair of simple lotteries (I highlight the terms *complete* and *transitive* to remind you that they have specific meaning from axiomatic utility theory, given at the beginning of the semester). [This could also be called an axiom – or even two axioms!]

- **Axiom 1. Continuity**. *Small changes in probabilities do not change the nature of the ordering of two lotteries. This can be made concrete here (I won't use formal notation b/c it's a mess). If a "bowl of miso soup" is preferable to a "cup of Kenyan coffee," then a mixture of the outcome "bowl of miso soup" and a sufficiently small but positive probability of "death by sushi knife" is still preferred to "cup of Kenyan coffee."*

- Continuity rules out "lexicographic" preferences for alternatives, such as "safety first." Safety first is a lexicographic preference rule because it does not *trade-off* between safety and competing alternatives (fun) but rather simply requires safety to be held at a fixed value for any positive utility to be attained.

- The second key building block of our theory about preferences over lotteries is the so-called *Independence Axiom.*

- **Axiom 2. Independence.** *The preference relation* $\succsim$ *on the space of simple lotteries* $\pounds$ *satisfies the independence axiom if for all* $L, L', L'' \in \pounds$ *and* $\alpha \in (0, 1)$*, we have*

$$L \succsim L' \text{ if and only if } \alpha L + (1 - \alpha) L'' \succsim \alpha L' + (1 - \alpha) L''.$$

- In words, when we mix each of two lotteries with a third one, then the preference ordering of the two resulting mixtures does not depend on (is *independent of*) the particular third lottery used.

- Example: If bowl of miso soup is preferred to cup of Kenyan coffee, then the lottery (bowl of miso soup with 50% probability, death by sushi knife with 50% probability) is preferred to the lottery (cup of Kenyan coffee with 50% probability, death by sushi knife with 50% probability).

3.3 Expected utility theory

- We now want to define a class of utility functions over risky choices that have the "expected utility form." We will then prove that if a utility function satisfies the definitions above for *continuity* and *independence* in preferences over lotteries, then the utility function has the expected utility form.

- It's important to clarify now that "expected utility theory" does *not* replace consumer theory, which we've been developing all semester. Expected utility theory extends the model of consumer theory to choices over risky outcomes. Standard consumer theory continues to describe the utility of consumption of specific *bundles.* Expected utility theory describes how a consumer might select among risky bundles.

- **Definition 3.** *The utility function $U : \pounds \to \mathbb{R}$ has an expected utility form if there is an assignment of numbers $(u_1, ...u_N)$ to the N outcomes such that for every simple lottery $L = (p_1, ..., p_N) \in \pounds$ we have that*

$$U(L) = u_1 p_1 + ... + u_N p_N.$$

- A utility function with the expected utility form is called a Von Neumann-Morgenstern (VNM) expected utility function.

- The term *expected utility* is appropriate because with the VNM form, the utility of a lottery can be thought of as the expected value of the utilities u_n of the N outcomes.

- In other words, a utility function has the expected utility form if and only if:

$$U\left(\sum_{k=1}^{K} \alpha_k L_k\right) = \sum_{k=1}^{K} \alpha_k U(L_k)$$

  for any K lotteries $L_k \in \pounds$, $k = 1, ..., K$, and probabilities $(\alpha_1, ..., \alpha_K) \geq 0$, $\Sigma_k \alpha_k = 1$.

- Intuitively, a utility function that has the expected utility property if the utility of a lottery is simply the (probability) weighted average of the utility of each of the outcomes.

- A person with a utility function with the expected utility property flips a coin to gain or lose one dollar. The utility of that lottery is

$$U(L) = 0.5U(w+1) + 0.5U(w-1),$$

  where w is initial wealth.

- Q: Does that mean that

$$U(L) = 0.5(w+1) + 0.5(w-1) = w?$$

  No. We haven't actually defined the utility of an *outcome*, and we certainly don't want to assume that $U(w) = w$.

3.4 Proof of expected utility property

Proposition 8 (1) *(Expected utility theory) Suppose that the rational preference relation $\succsim$ on the space of lotteries $\pounds$ satisfies the continuity and independence axioms. Then $\succsim$ admits a utility representation of the expected utility form. That is, we can assign a number u_n to* *each outcome* $n = 1, ..., N$ *in such a manner that for any two lotteries* $L = (p_1, ...P_N)$ *and* $L' = (p'_1, ...P'_N)$, *we have* $L \succsim L'$ *if and only if*

$$\sum_{n=1}^{N} u_n p_n \geq \sum_{n=1}^{N} u_n p'_n$$

Proof: Expected Utility Property (in five steps)

Assume that there are best and worst lotteries in $£$, $\bar{L}$ and $\underline{L}$.

1. If $L \succ L'$ and $\alpha \in (0, 1)$, then $L \succ \alpha L + (1 - \alpha) L' \succ L'$. This follows immediately from the independence axiom.

2. Let $\alpha, \beta \in [0, 1]$. Then $\beta \bar{L} + (1 - \beta)\underline{L} \succ \alpha \bar{L} + (1 - \alpha)\underline{L}$ if and only if $\beta > \alpha$. This follows from the prior step.

3. For any $L \in £$, there is a unique α_L such that $\left[\alpha_L \bar{L} + (1 - \alpha_L) \underline{L}\right] \sim L$. Existence follows from continuity. Uniqueness follows from the prior step.

4. The function $U : £ \to \mathbb{R}$ that assigns $U(L) = \alpha_L$ for all $L \in £$ represents the preference relation $\succsim$.
   Observe by Step 3 that, for any two lotteries $L, L' \in £$, we have

   $$L \succsim L' \text{ if and only if } \left[\alpha_L \bar{L} + (1 - \alpha_L) \underline{L}\right] \succsim \left[\alpha_{L'} \bar{L} + (1 - \alpha_{L'}) \underline{L}\right].$$

   Thus $L \succsim L'$ if and only if $\alpha_L \geq \alpha_{L'}$.

5. The utility function $U(\cdot)$ that assigns $U(L) = \alpha_l$ for all $L \in £$ is linear and therefore has the expected utility form.
   We want to show that for any $L, L' \in £$, **and** $\beta [0, 1]$, **we have** $U(\beta L + (1 - \beta) L') = \beta U(L) + (1 - \beta) U'(L)$.
   By step (3) above, we have

   $$\begin{aligned} L &\sim U(L) \bar{L} + (1 - U(L)) \underline{L} = \alpha_L \bar{L} + (1 - \alpha_L) \underline{L} \\ L' &\sim U(L') \bar{L} + (1 - U(L')) \underline{L} = \alpha'_L \bar{L} + (1 - \alpha'_L) \underline{L}. \end{aligned}$$

   By the Independence Axiom,

   $$\beta L + (1 - \beta) L' \sim \beta \left[U(L) \bar{L} + (1 - U(L)) \underline{L}\right] + (1 - \beta) \left[U(L') \bar{L} + (1 - U(L')) \underline{L}\right].$$

   Rearranging terms, we have

   $$\begin{aligned} \beta L + (1 - \beta) L' &\sim [\beta U(L) + (1 - \beta) U(L')] \bar{L} + \left[\beta (1 - U(L)) + (1 - \beta) \left(1 - U(L)'\right)\right] \underline{L} \\ &= [\beta U(L) + (1 - \beta) U(L')] \bar{L} + [1 - \beta U(L) + (\beta - 1) U(L')] \underline{L}. \end{aligned}$$

By step (4), this expression can be written as

$$
\begin{aligned}
& [\beta \alpha_L + (1-\beta)\,\alpha_{L'}]\,\bar{L} + [1 - \beta\alpha_L + (\beta - 1)\,\alpha'_L]\,\underline{L} \\
= \;& \beta\left(\alpha_L \bar{L} + (1-\alpha_L)\,\underline{L}\right) + (1-\beta)\left(\alpha_{L'}\bar{L} + (1-\alpha'_L)\,\underline{L}\right) \\
= \;& \beta U(L) + (1-\beta)\,U(L')\,.
\end{aligned}
$$

This establishes that a utility function that satisfies continuity and the Independence Axiom, has the expected utility property: $U(\beta L + (1-\beta)\,L') = \beta U(L) + (1-\beta)\,U(L')$

[End of optional self-study section.]

4 Summary of Expected Utility property

- **The key equation**: Preferences that satisfy VNM Expected Utility theory have the property that:

  $$U(\beta L + (1-\beta)\,L') = \beta U(L) + (1-\beta)\,U(L'),$$

  where L and L' are bundles with $L \neq L'$ and $\beta \in (0,1)$. This equation says that for a person with VNM preferences, the utility of consuming two bundles L and L' with probabilities β and $(1-\beta)$, respectively, is equal to β times the utility of consuming bundle β plus $(1-\beta)$ times the utility of consuming bundle L'. Thus, the utility function is *linear* in probabilities though *not necessarily linear* in preferences over the bundles. That is, VNM does *not* imply that $U(2L) = 2 \times U(L)$. [As we'll see below, that equation would *only* hold for risk neutral preferences.]

- A person who has VNM EU preferences over lotteries will act as if she is maximizing *expected utility*—a weighted average of utilities of each state, where weights equal probabilities.

- If this model is correct, then we don't need to know exactly how people feel about risk *per se* to make strong predictions about how they will optimize over risky choices.

- [If the model is not entirely correct—which it surely is not—it may still provide a useful description of the world and/or a normative guide to how one should analytically structure choices over risky alternatives.]

- To use this model, two ingredients are needed:

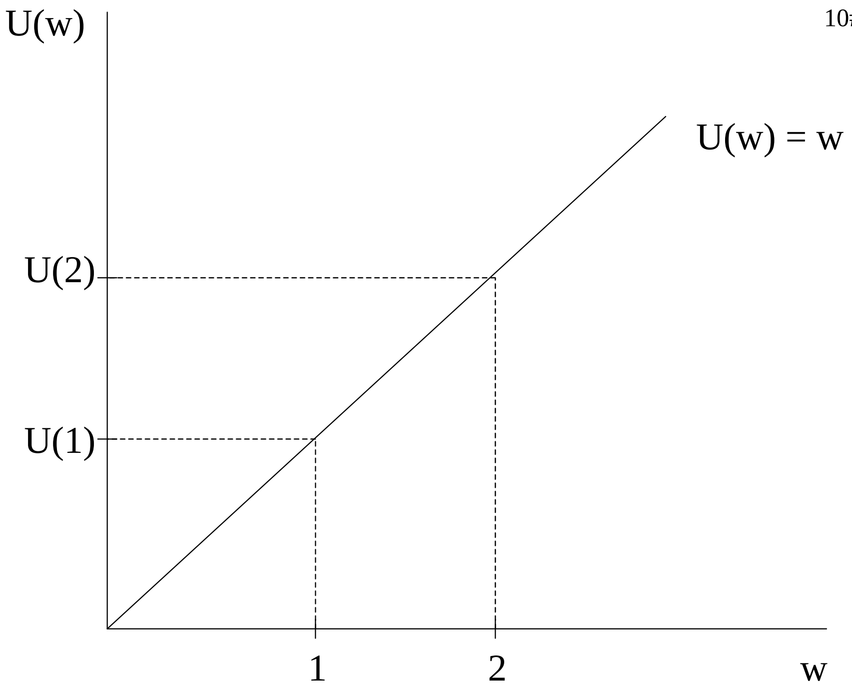
1. First, a utility function that assigns bundles an ordinal utility ranking. Note that such functions are defined up to an affine (i.e., positive linear) transformation. This means they are required to have more structure (i.e., are more restrictive) than standard consumer utility functions, which are only defined up to a monotone transformation.
2. Second, the VNM assumptions. These make strong predictions about the maximizing choices consumers will take when facing risky choices (i.e., probabilistic outcomes) over bundles, which are of course ranked by this utility function.

5 Expected Utility Theory and Risk Aversion

- We started off to explain risk aversion. What we have done to far is lay out expected utility theory, which is a set of (relatively restrictive) axioms about how consumers make choices among risky bundles.
- Where does risk aversion come in?
- Consider the following three utility functions characterizing three different expected utility maximizers:
-

$$\mathbf{u}_1(\mathbf{w}) = \mathbf{w}$$

10#1a

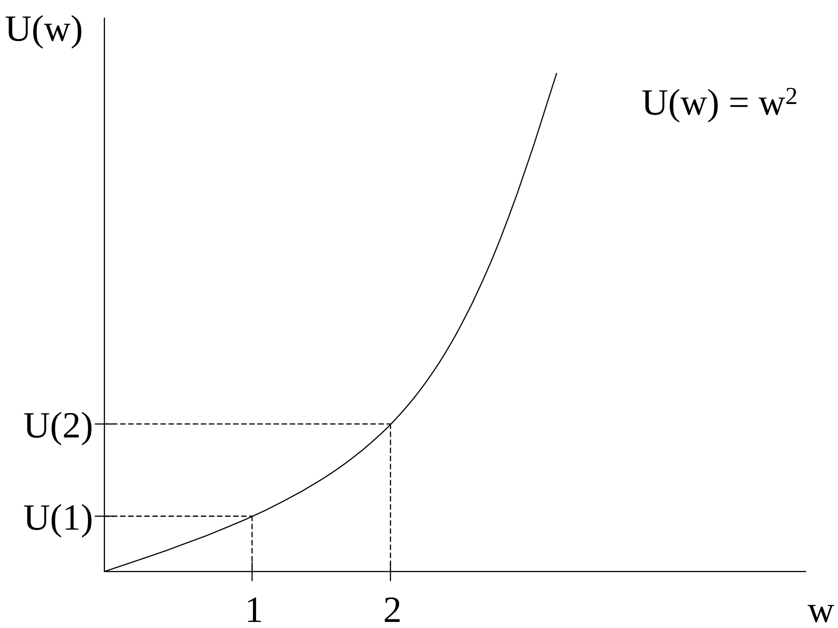


–

•

$$\mathbf{u}_2(\mathbf{w}) = \mathbf{w}^2$$

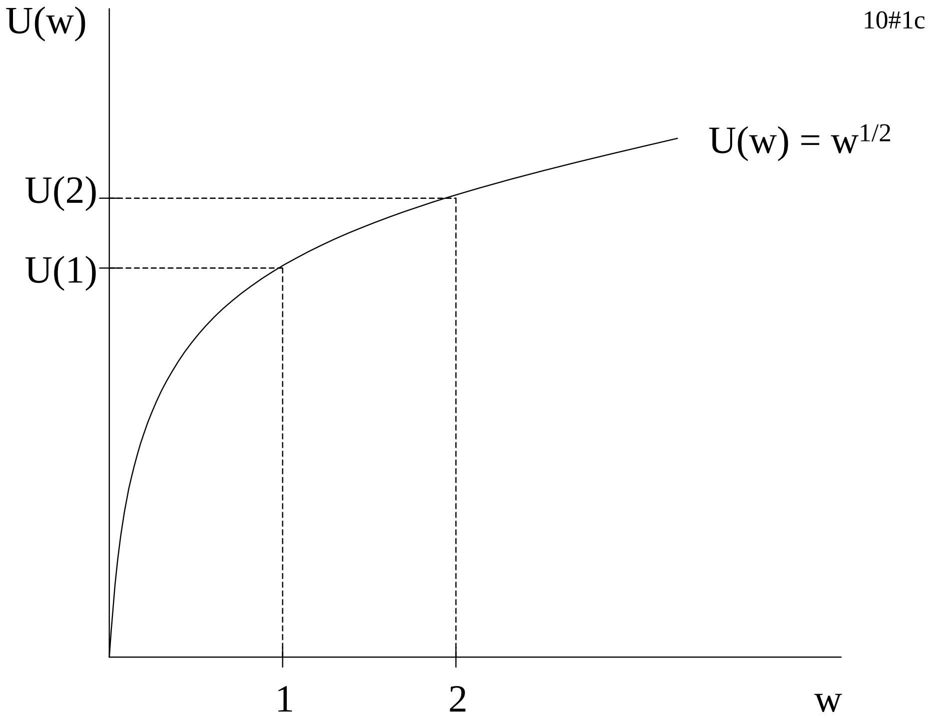
10#1b



–

•

$$\mathbf{u}_3(\mathbf{w}) = \sqrt{w}$$



–

- Consumer a lottery where the consumer faces 50/50 odds of either receiving two dollars or zero dollars. The expected monetary value of this lottery is \$1.
- How do these three consumers differ in risk preference?
- First notice that $u_1(1) = u_2(1) = u_3(1) = 1$. That is, they all value *one dollar with certainty* equally.
- Now consider the *Certainty Equivalent* for a lottery L that is a 50/50 gamble over \$2 versus \$0. The certainty equivalent is the amount of cash that the consumer be willing to accept with certainty in lieu of facing lottery L.
  - Step 1: What is the expected utility value?
    1. $u_1(L) = .5 \cdot u_1(0) + .5 \cdot u_1(2) = 0 + .5 \cdot 2 = 1$
    2. $u_2(L) = .5 \cdot u_1(0) + .5 \cdot u_1(2) = 0 + .5 \cdot 2^2 = 2$
    3. $u_3(L) = .5 \cdot u_1(0) + .5 \cdot u_1(2) = 0 + .5 \cdot 2^{.5} = .71$
  - Step 2: What is the "Certainty Equivalent" of lottery L for these three utility functions—that is, the cash value that the consumer would take in lieu of facing these lotteries?

1. $CE_1(L) = U_1^{-1}(1) = \$1.00$
2. $CE_2(L) = U_2^{-1}(2) = 2^{.5} = \$1.41$
3. $CE_3(L) = U_3^{-1}(0.71) = 0.71^2 = \$0.51$

- Depending on the utility function, a person would pay $\$1, \$1.41$, or $\$0.51$ dollars to participate in this lottery.

- Although the expected monetary value $E(V)$ of the lottery is $1.00, the three utility functions value it differently:

  1. The person with U_1 is *risk neutral*: $CE = \$1.00 = E(Value) \Rightarrow$ Risk neutral
  2. The person with U_2 is *risk loving*: $CE = \$1.41 > E(Value) \Rightarrow$ Risk loving
  3. The person with U_3 is *risk averse*: $CE = \$0.50 < E(Value) \Rightarrow$ Risk averse

- *What gives rise to these inequalities is the shape of the utility function. Risk preference comes from the concavity/convexity of the utility function*:

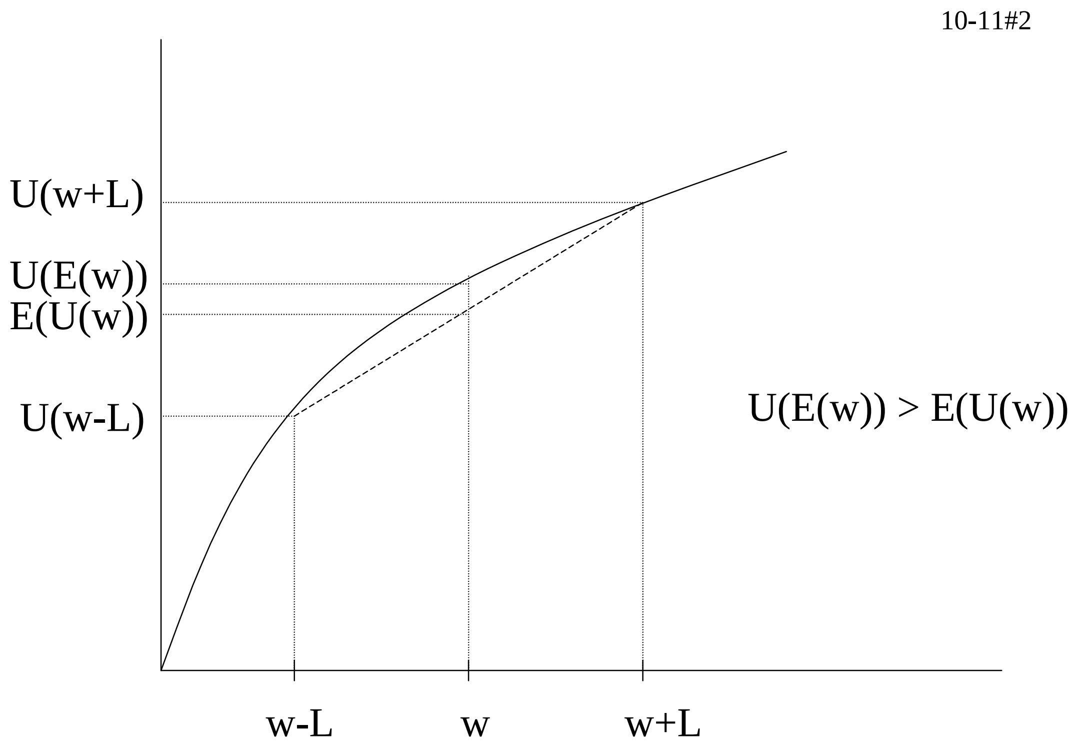
- Expected utility of wealth: $E(U(w)) = \sum_{i=1}^{N} p_i U(w_i)$

- Utility of expected wealth: $U(E(w)) = U\left(\sum_{i=1}^{N} p_i w_i\right)$

- Jensen's inequality:

  - $E(U(w)) = U(E(w)) \Rightarrow$ Risk neutral
  - $E(U(w)) > U(E(w)) \Rightarrow$ Risk loving
  - $E(U(w)) < U(E(w)) \Rightarrow$ Risk averse

- So, the core insight of expected utility theory is this: *For a risk averse consumer, the expected utility of wealth is less than the utility of expected wealth (given non-zero risk).*



- The reason this is so:
  If wealth has diminishing marginal utility (as is true if $U(w) = w^{1/2}$), losses cost more utility than equivalent monetary gains provide.

- Consequently, a risk averse consumer is better off to receive a given amount of wealth *with certainty* than the same amount of wealth *on average* but with variance around this quantity.

5.1 Application: Risk aversion and insurance

- Consider insurance that is *actuarially fair,* meaning that the premium is equal to expected claims: Premium $= p \cdot A$ where p is the expected probability of a claim, and A is the amount that the insurance company will pay in the event of an accident.

- How much insurance will a risk averse person buy?

- Consider a person with an initial endowment consisting of three things:.A level of wealth w_0; a probability of an accident of p; and the amount of the loss, L (in dollars) should a loss occur:

$$\begin{aligned}
\Pr(1-p) &: U(\cdot) = U(w_o), \\
\Pr(p) &: U(\cdot) = U(w_o - L)
\end{aligned}$$

- If insured, the endowment is (incorporating the premium pA, the claim paid A if a claim is made,and the loss L):

$$\begin{aligned}
\Pr(1-p) &: U(\cdot) = U(w_o - pA), \\
\Pr(p) &: U(\cdot) = U(w_o - pA + A - L)
\end{aligned}$$

- Expected utility if uninsured is:

$$E(U|I=0) = (1-p)U(w_0) + pU(w_o - L).$$

- Expected utility if insured is:

$$E(U|I=1) = (1-p)U(w_0 - pA) + pU(w_o - L + A - pA). \tag{1}$$

- How much insurance would this person wish to buy (assuming they can buy up to their total wealth, $w_0 - pL$, at actuarially fair prices)? To solve for the optimal amount of insurance that the consumer should purchase, maximize their utility with respect to the insurance policy:

$$\begin{aligned}
\max_A E(U) &= (1-p)U(w_0 - pA) + pU(w_o - L + A - pA) \\
\frac{\partial E(U)}{\partial A} &= -p(1-p)U'(w_0 - pA) + p(1-p)U'(w_o - L + A - pA) = 0. \\
&\Rightarrow U'(w_0 - pA) = U'(w_o - L + A - pA), \\
&\Rightarrow A = L, \text{ which implies that wealth is } w_0 - L \text{ in both states of the world (insurance c}
\end{aligned}$$

- A risk averse person will optimally buy *full insurance* if the insurance is actuarially fair.
- Is the person better off for buying this insurance? Absolutely. You can verify that expected utility rises with the purchase of insurance *although expected wealth is unchanged.*
- You could solve for *how much* the consumer would be willing to pay for a given insurance policy. Since insurance increases the consumer's welfare, s/he will be willing to pay some positive price *in excess of the actuarially fair premium* to defray risk.

- What is the intuition for why consumers want full insurance?
  - *The consumer is seeking to equate the marginal utility of wealth across states.*
  - Why? For a risk averse consumer, the utility of average wealth is greater than the average utility of wealth.
  - The consumer therefore wants to distribute wealth evenly across states of the world, rather than concentrate wealth in one state.
  - The consumer will attempt to maintain wealth at the same level in all states of the world, assuming she can costlessly transfer wealth between states of the world (which is what actuarially fair insurance allows the consumer to do).
- This is exactly analogous to convex indifference curves over consumption bundles.
  - Diminishing marginal rate of substitution across goods (which comes from diminishing marginal utility of consumption) causes consumer's to want to diversify across goods rather than specialize in single goods.
  - Similarly, diminishing marginal utility of wealth causes consumers to wish to diversify wealth across possible states of the world rather than concentrate it in one state.
- Q: How would answer to the insurance problem change if the consumer were *risk loving?*
- A: They would want to be at a corner solution where all risk is transferred to the least probable state of the world, again holding constant expected wealth.
- The more risk the merrier. Would buy "uninsurance."
- **OPTIONAL:**
  - For example, imagine the consumer faced probability p of some event occurring that induces loss L.
  - Imagine the policy pays $A = \frac{w_0}{p}$ in the event of a loss and costs pA.

$$
\begin{aligned}
W(\text{No Loss}) &= w_0 - p\left(\frac{w_0}{p}\right) = 0, \\
W(\text{Loss}) &= w_0 - L - p\left(\frac{w_0}{p}\right) + \frac{w_0}{p} = \frac{w_0}{p} - L. \\
E(U) &= (1-p)U(0) + pU\left(\frac{w_0}{p} - L\right).
\end{aligned}
$$

  - For a risk loving consumer, putting all of their eggs into the least likely basket maximizes expected utility.

5.2 Operation of insurance: State contingent commodities

- To see how risk preference generates demand for insurance, it is useful to think of insurance as a 'state contingent commodity,' a good that you buy now *but only consume* if a specific state of the world arises.

- Insurance is a state contingent commodity: when you buy insurance, you are buying a claim on \$1.00. This insurance is purchased before the state of the world is known. You can only make the claim for the payout if the relevant state arises.

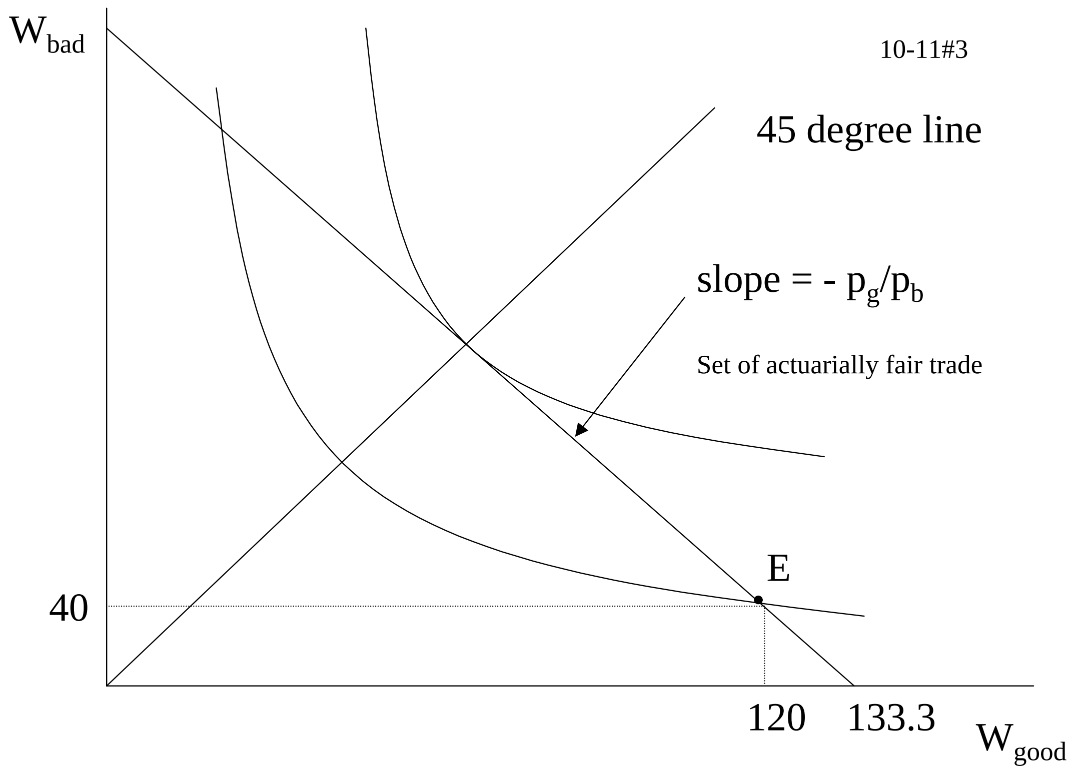
- Previously, we've drawn indifference maps across goods X, Y. Now we will draw indifference maps across states of the world: *good, bad.*

- Consumers can use their endowment (equivalent to budget set) to shift wealth across states of the world via insurance, just like budget set can be used to shift consumption across goods X, Y.

- Example: Two states of world, good and bad, with $w_0 = 120$, $p = 0.25$, $L = 80$.

$$
\begin{aligned}
w_g &= 120 \\
w_b &= 120 - 80 \\
\Pr(g) &= (1 - p) = 0.75 \\
\Pr(b) &= p = 0.25 \\
E(w) &= 0.75(120) + .25(40) = 100 \\
E(u(w)) &< u(E(w)) \text{ if consumer is risk averse.}
\end{aligned}
$$

- See FIGURE.



- Let's say that this consumer can buy actuarially fair insurance. What will it sell for?
- If you want \$1.00 in Good state, this will sell of \$0.75 *prior to the state being revealed.*
- If you want \$1.00 in Bad state, this will sell for \$0.25 *prior to the state being revealed.*
- Why these prices? Because these are the expected probabilities of making the claim. So, a risk neutral consumer (say a central bank) could sell you insurance against bad states at a price of \$0.25 on the dollar and insurance again good states (assuming you wanted to buy it) at a price of \$0.75 on the dollar.
- The price ratio is therefore
$$\frac{P_g}{P_b} = \frac{p}{(1-p)} = 3.$$
- The set of fair trades among these states can be viewed as a 'budget set' and the slope of which is $-\frac{P}{(1-P)}$, and which passes through the initial endowment.
- Now we need indifference curves.
- Recall that the utility of this lottery (the endowment) is:
$$u(L) = pu(w_g) + (1-p)u(w_b).$$

- Along an indifference curve

$$\begin{aligned} dU &= 0 = pu'(w_g)dw_g + (1-p)u'(w_b)dw_b, \\ \frac{dw_b}{dw_g} &= -\frac{pu'(w_g)}{(1-p)u'(w_b)} < 0. \end{aligned}$$

- Provided that $u()$ is concave, these indifference curves are bowed towards the origin in probability space. [It can readily be proven that indifference curves are convex to origin by taking second derivatives, but the intuition is straightforward.]
  - Flat indifference curves would indicate risk neutrality – because for risk neutral consumers, expected utility is linear in expected wealth.
  - Convex indifference curves mean that you must be compensated to bear risk.
  - i.e., if I gave you \$133.33 in good state and 0 in bad state, you are strictly worse off than getting \$100 in each state, even though your expected wealth is

$$E(w) = 0.75 \cdot 133.33 + 0.25 \cdot 0 = 100.$$

  - So, I would need to give you more than \$133.33 in the good state to compensate for this risk.
  - Bearing risk is psychically costly – must be compensated.
- Therefore there are potential utility improvements from reducing risk.
- In the figure, the movement from the lower (closer to the origin) to the upper indifference curve is the gain from shedding risk.
- Notice from the Figure that along the 45^0 line, $w_g = w_b$.
- But if $w_g = w_b$, this implies that

$$\frac{dw_b}{dw_g} = -\frac{pu'(w_g)}{(1-p)u'(w_b)} = \frac{p}{(1-p)} = \frac{P_g}{P_b}.$$

- Hence, the indifference curve will be tangent to the budget set at exactly the point where wealth is equated across states. [This is an alternative way of demonstrating the results above that a risk averse consumer will always fully insure if insurance is actuarially fair.]
- This is a very strong restriction that is imposed by the expected utility property: *The slope of the indifference curves in expected utility space must be tangent to the odds ratio.*

6 Aside: A Counterexample (from MacLean, 1986) [Optional]

Consider the following far-fetched example, which attacks the intuition of expected utility theory. A decision-maker and her group of six employees is captured by terrorists. The decision-maker is given six bullets and six guns, each a six-shooter (that is, holding six bullets), and told that her employees must play Russian Roulette. The rules are as follows:

- She distributes the six bullets among the six guns as she prefers.
- She then chooses one of two 'games:'

  1. She picks one gun at random and fires it sequentially at all six employees.

-or-

- 2. She gives one gun at random to each employee. Each employee spins the barrel of her gun (to randomize bullet locations) and fires once.

Consider the following possible variations the decision-maker could choose (there are 3.9 million variations):

1. She puts one bullet in each pistol, chooses a pistol at random and fires it sequentially at all six employees. This yields one death with certainty.
2. She puts all six bullets in one pistol and distributes the pistols among the employees. This again yields one death with certainty.
3. She puts one bullet in each pistol and distributes the six pistols. There are 7 possible outcomes in this case:
   (a) No deaths: $p = 0.335$
   (b) 1 death: $p = 0.402$
   (c) 2 deaths: $p = 0.201$
   (d) 3 deaths: $p = 0.054$
   (e) 4 deaths: $p = 0.008$
   (f) 5 deaths: $p = 0.001$
   (g) 6 deaths: $p = 1/6^6 \approx 0$

4. She puts six bullets in one pistol, selects a pistol at random, and fires it sequentially:
   (a) No deaths: $p = 0.833$
   (b) 6 deaths: $p = 0.167$
5. She puts two bullets in each of three pistols, picks one at random, and fires sequentially:
   (a) No deaths: 0.5
   (b) 2 deaths: 0.5
6. She puts two bullets in each of 3 pistols, and distributes all pistols at random:
   (a) No deaths: 0.297
   (b) 1 death: 0.444
   (c) 2 deaths: 0.222
   (d) 3 deaths: 0.037

MacLean argues that, if all employees are considered interchangeable (have identical utility value) then these six scenarios are ex ante identical from the perspective of expected utility theory—and thus, the decision-maker should be indifferent among them. Clearly, most people would not be indifferent among these possible allocations of risk.

Is MacLean's argument correct? It's certainly thought-provoking. And the argument could be correct, depending on whether these outcomes (deaths) are measured on the linear or non-linear section of the utility function. What MacLean has in mind is:

$$U(L) = \Sigma_{i=1}^{6}\left[p_i U(D_i)\right].$$

Here, the $p_i's$ sum to 1, and the utility loss from each death is identical. Hence, any combination of $p_i's$ that sum to 1 have identical expected utility value to the decision maker. This may appear realistic.

Let's say that, instead, that deaths enter into the non-linear section of the utility function. Hence

$$U(L) = p_1 U(1 \cdot D) + p_2 U(2 \cdot D) + ... + p_6 U(6 \cdot D).$$

Let's assume further that the decision-maker is risk averse, meaning that deaths have increasing marginal harm (the disutility of two deaths is more than twice the disutility of one death). What prediction does this model make about the decision maker's choice? Does that seem obviously correct?

7 The Market for Insurance

Now consider how the market for insurance operates. If everyone is risk averse (and it's pretty safe to assume that most are), how can insurance exist at all? Who would sell it?

There are actually three distinct mechanisms by which insurance can operate: risk pooling, risk spreading and risk transfer.

7.1 Risk pooling

Risk pooling is the main mechanism underlying most *private* insurance markets. Its operation depends on the Law of Large Numbers. Relying on this mechanism, it defrays risk, which is to say that it makes it disappear.

- For example, for any number of tosses n of a fair coin, the expected fraction of heads H is $E(H) = \frac{0.5n}{n} = 0.5$. But the variance around this expectation (equal to $\frac{p(1-p)}{n}$) is declining in the number of tosses:

$$
\begin{aligned}
V(1) &= 0.25 \\
V(2) &= 0.125 \\
V(10) &= 0.025 \\
V(1,000) &= 0.00025
\end{aligned}
$$

- We cannot predict the **share** of heads in one coin toss with any precision, but we can predict the **share** of heads in $10,000$ coin tosses with considerable confidence. It will be vanishingly close to 0.5.

- Therefore, by *pooling* many independent risks, insurance companies can treat uncertain outcomes as *almost known.*

- So, "risk pooling" is a mechanism for providing insurance. It *defrays* the risk across independent events by exploiting the law of large numbers – makes risk effectively disappear.

7.1.1 Example:

- Let's say that each year, there is a $1/250$ chance that my house will burn down. If it does, I lose the entire $\$250,000$ house. The expected cost of a fire in my house each year is therefore about $\$1,000$.

- Given my risk aversion, it is costly in expected utility terms for me to bear this risk (i.e., much more costly than simply reducing my wealth by $\$1,000$).

- If $100,000$ owners of $\$250,000$ homes all put $\$1,000$ into the pool, this pool will collect $\$100$ million.

- In expectation, 400 of us will lose our houses $\left(\frac{100,000}{250} = 400\right)$.

- The pool will therefore pay out approximately $250,000 \cdot 400 = \$100$ million and approximately break even.

- Everyone who participated in this pool is better off to be relieved of the risk, though most will pay $\$1,000$ the insurance premium and not lose their house.

- However, there is still some risk that the pool will face a larger loss than the expected $1/400$ of the insured.

- The law of large numbers says this variance gets vanishingly small if the pool is large and the risks are independent. How small?

$$\begin{aligned} V(Loss) &= \frac{P_{Loss}(1-P_{Loss})}{100,000} \frac{0.004(1-0.004)}{100000} = 3.984 \times 10^{-8} \\ SD(Loss) &= \sqrt{3.984 \times 10^{-8}} = 0.0002 \end{aligned}$$

- Using the fact that the binomial distribution is approximately normally distributed when n is large, this implies that:

$$\Pr[Loss \in (0.004 \pm 1.96 \cdot 0.0002)] = 0.95$$

- So, there is a 95% chance that there will be somewhere between 361 and 439 losses, yielding a cost per policy holder in 95% of cases of $\$924.50$ to $\$1,075.50$.

- Most of the risk is defrayed is this pool of $100,000$ policies.

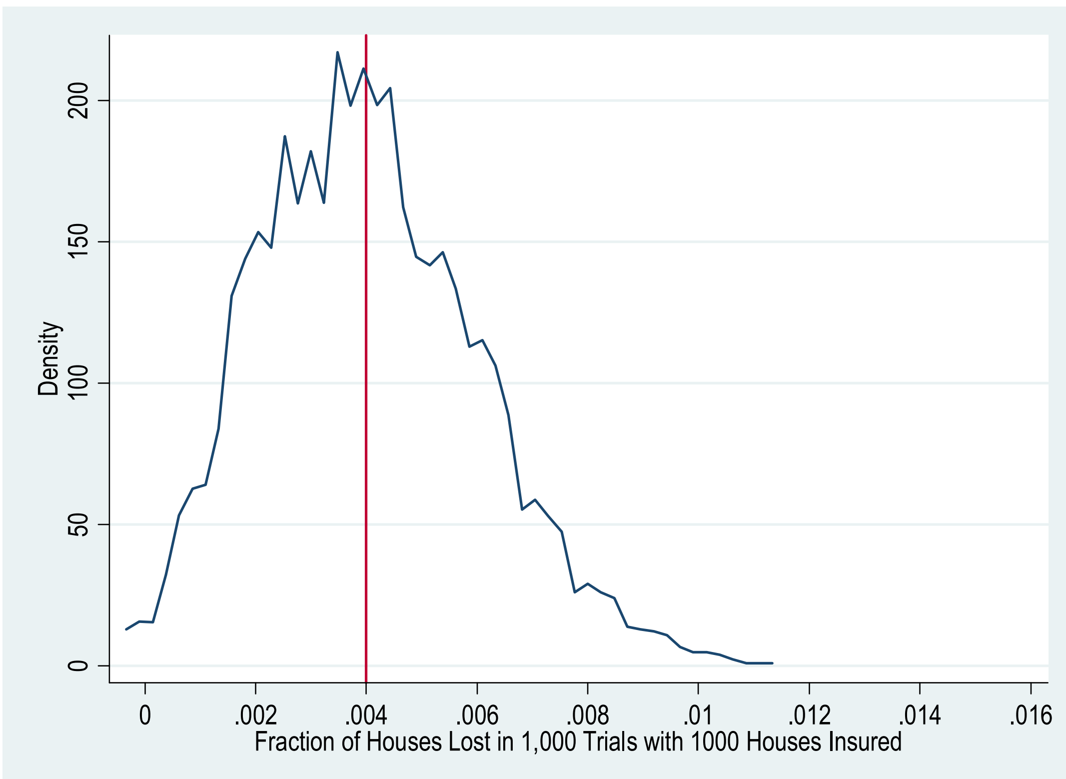
- And as $n \to \infty$, this risk entirely vanishes.

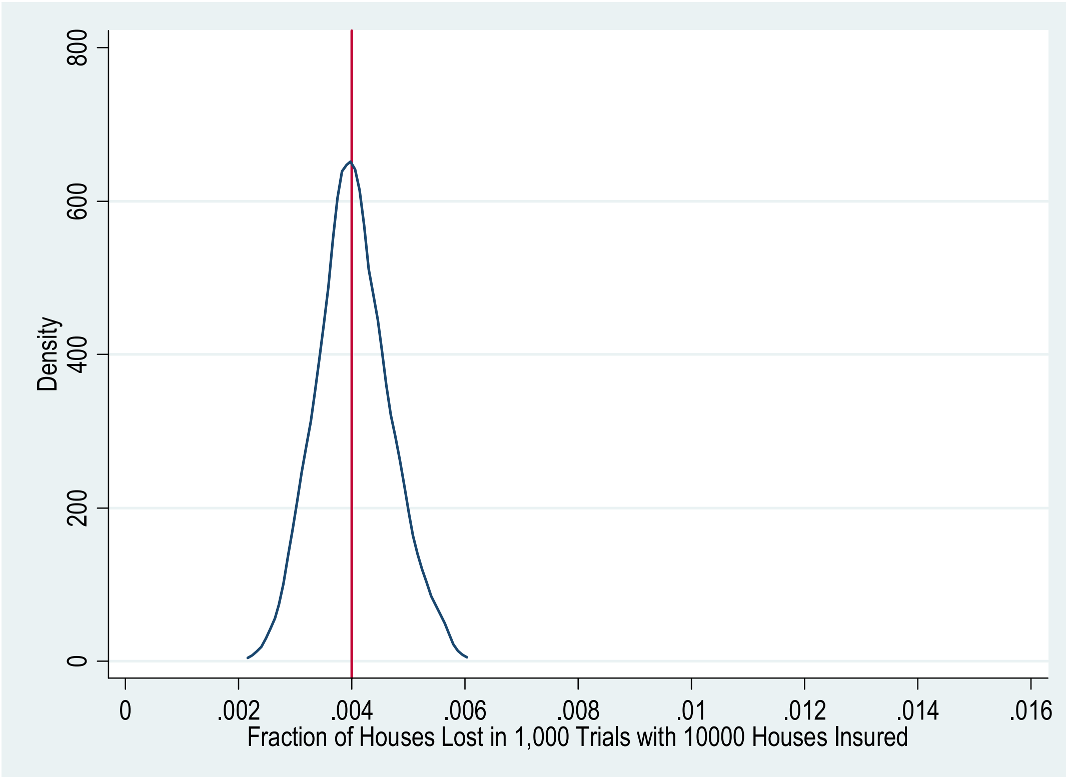
- So, risk pooling generates a Pareto improvement (assuming we establish the insurance mechanism before we know whose house will burn down).

- In class, I will also show a numerical example based on simulation. Here, I've drawn independent boolean variables, each with probability $1/250$ of equalling one (representing a loss). I plot the frequency distribution of these draws for $1,000$ replications, while varying the sample size (number of draws): $1,000$, $10,000$, $100,000$, $1,000,000$, and $10,000,000$.

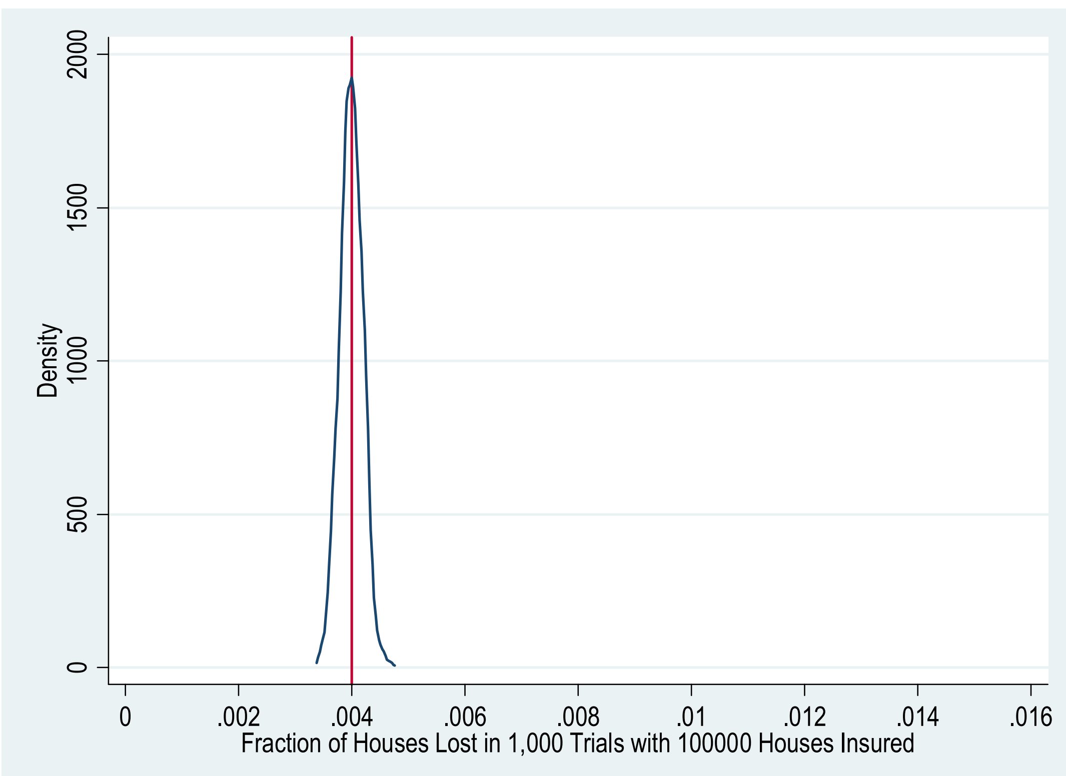
- This simulation shows that as the number of independent risks gets large (that is, the sample size grows), the odds that the number of losses will be more than a few percentage points from the mean contracts dramatically.

- With sample size $10,000,000$, there is virtually no chance that the number of losses would exceed $1/250 \cdot N$ by more than a few percent. Hence, pooling of independent risks effectively eliminates these risks – a Pareto improvement.





Density
0
200
400
600
800
0
.002
.004
.006
.008
.01
.012
.014
.016
Fraction of Houses Lost in 1,000 Trials with 10000 Houses Insured



Density
0
500
1000
1500
2000
0
.002
.004
.006
.008
.01
.012
.014
.016
Fraction of Houses Lost in 1,000 Trials with 100000 Houses Insured

7.2 Risk spreading

- When does this 'pooling' mechanism above not work? When risks are not independent:
  - Earthquake
  - Flood
  - Epidemic
- When a catastrophic event is likely to affect many people simultaneously, it's (to some extent) **non-diversifiable**.
- This is why many catastrophes such as floods, nuclear war, etc., are specifically not covered by insurance policies.
- But does this mean there is no way to insure?
- Actually, we can still 'spread' risk providing that there are some people likely to be unaffected.
- The basic idea here is that because of the concavity of the (risk averse) utility function, taking a little bit of money away from each person incurs lower social costs than taking a lot of money from a few people.
- Many risks cannot be covered by insurance companies, but the government can intercede by transferring money among parties. Many examples:
  - Victims compensation fund for World Trade Center.
  - Medicaid and other types of catastrophic health insurance.
  - All kinds of disaster relief.
- Many of these insurance 'policies' are not even written until the disaster occurs – there was no market. But the government can still spread the risk to increase social welfare.
- For example, imagine 100 people, each with VNM utility function $u(w) = \ln(w)$ and wealth 500. Imagine that one of them experiences a loss of 200. His utility loss is

$$L = u(300) - u(500) = -0.511.$$

- Now, instead consider if we took this loss and distributed it over the entire population:

$$L = 100 \cdot [\ln(498) - \ln(500)] = 100 \cdot [-0.004] = -0.401.$$

The aggregate loss (-0.401) is considerably smaller than the individual loss (-0.511). (This comes from the concavity of the utility function.)

- Hence, risk spreading may improve social welfare, even if it does not defray the total amount of risk faced by society.

- Does risk spreading offer a Pareto improvement? No, because we must take from some to give to others.

7.3 Risk transfer

- Third idea: if utility cost of risk is declining in wealth (constant absolute risk aversion for example implies declining relative risk aversion), this means that *less wealthy people could pay more wealthy people to bear their risks* and both parties would be better off.

- Again, take the case where $u(w) = \ln(w)$. Imagine that an individual faces a 50 percent chance of losing \$100. What would this person pay to eliminate this risk? It will depend on his or her initial wealth.

- Assume that initial wealth is 200. Hence, expected utility is

$$u(L) = 0.5 \ln 200 + 0.5 \ln 100 = 4.952$$

Expected wealth is \$150. The certainty equivalent of this lottery is $\exp[4.592] = \$141.5$. Hence, the consumer would be willing to pay up to \$8.50 to defray this risk.

- Now consider a person with the same utility function with wealth $1,000$. Expected utility is

$$u(L) = 0.5 \ln 1000 + 0.5 \ln 900 = 6.855.$$

Expected wealth is \$950. The certainty equivalent of this lottery is $\exp[6.855] = \$948.6$. Hence, the consumer would be willing to pay only \$1.40 to defray the risk.

- The wealthy consumer could fully insure the poor consumer at psychic cost \$1.40 while the poor consumer would be willing to pay \$8.50 for this insurance. Any price that they can agree between $(\$1.40, \$8.50)$ represents a pure Pareto improvement.

- Why does this form of risk transfer work? Because the logarithmic utility function exhibits declining absolute risk aversion—the wealthier someone is, the lower their psychic cost of bearing a fixed monetary amount of risk. Is this realistic? Probably.

- Example: Lloyds of London used to perform this risk transfer role:

  - Took on large, idiosyncratic risks: satellite launches, oil tanker transport, the Titanic.
  - These risks are not diversifiable in any meaningful sense.
  - But companies and individuals would be willing to pay a great deal to defray them.
  - Lloyds pooled the wealth of British nobility and gentry ('names') to create a super-rich consumer that in aggregate was much more risk tolerant than even the largest company.
  - For over a century, this idea generated large, steady inflows of cash for the 'names' that underwrote the Lloyds' policies.
  - Then they took on asbestos liability...
  - [For a fascinating account of how Lloyds bankrupted the British nobility, have a look at the 1993 *New Yorker* article by Julian Barnes, "The Deficit Millionaires" on the class website. This article doesn't have much economic content, but it's pretty gripping.]

7.4 Insurance markets: Conclusion

- Insurance is potentially an extremely beneficial financial/economic institution, which can make people better off at low or even zero cost aggregate cost (in the case of risk pooling).

- We'll discuss shortly why insurance markets do not work as well in reality as they might in theory (though they still create enormous social value despite imperfections).

MIT OpenCourseWare
http://ocw.mit.edu

14.03 / 14.003 Microeconomic Theory and Public Policy
Fall 2010

For information about citing these materials or our Terms of Use, visit: http://ocw.mit.edu/terms.