Basic Informational Economics

Assignment #4 for Managerial Economics, ECO 351M, Fall 2016
Due, Monday October 31 (Halloween).

The Basic Model

One must pick an action, a in a set of possible actions A, before a random variable, X, takes one of its possible values, $X = x$, at which point, the realized utility is $u(a, x)$. Before on picks the action, one observes a signal, S, taking one of its possible values $S = s$. The signal S is correlated with the random variable X, and this is what makes it useful. The joint distribution of X and S is given by $q(x, s) = Prob(X = x, S = s)$. Here we are assuming that the values of X and S are discrete, changing to continuous random variables makes the math harder without adding a great deal of intuition. In any case, the expected utility maximizer's problem is to pick a function from observations to actions, $s \mapsto a(s)$, so as to maximize

$$(\ddagger) \qquad E\,u(a(S), X) = \textstyle\sum_{x,s} u(a(s), x)q(x, s).$$

We will use two pieces of notation: $\pi(s) = Prob(S = s) = \sum_x q(x, s)$ is the distribution of the random signal S; $\beta(x) = Prob(X = x) = \sum_s q(x, s)$ is the original (or prior) distribution of the random variable X.

Solving the Basic Model

There are, theoretically, two ways to solve the problem in equation (‡). The first, wildly impractical way to solve the problem is to look at all possible policies, that is, to look at each and every function $a : S \to A$, for each function, calculate $E\,u(a(S), X)$, and then pick the one giving the maximal value. The second way is known as "crossing the bridge when you come to it," one waits until observing a signal value, $S = s$, and then figures out the best thing to do. Since you only observe one signal, s, this means that you don't need to think about what you would have done if some other signal, $s' \neq s$, had happened.

> An example demonstrates the comparison between the first and the second approach: the first approach involves doing your homework for this class by figuring out what answer you would give to every problem I could possibly assign, and then once you see what problems I actually assign, pick your answers from what you've already done; the second approach waits until you see what problems are actually assigned and then solves those.

All that is left is to "figure out the best thing to do" after seeing $S = s$. We do this by re-writing equation (‡) above. We now use $\pi(s) = \sum_x q(x, s) = Prob(S = s)$. Using this, rewrite $E\,u = \sum_{x,s} u(a(s), x)q(x, s)$ as

$$\textstyle\sum_s \pi(s) \sum_x u(a(s), x) \frac{q(x,s)}{\pi(s)}$$

where the sum is over the s such that $\pi(s) > 0$ so that we're never dividing by 0. Let us now define your 'beliefs,' denoted by β as a mnemonic for beliefs, about the probability that $X = x$ given that we have seen $S = s$. From Bayes Rule, this is

$$Prob(X = x|S = s) = \beta(x|s) := \tfrac{q(x,s)}{\pi(s)},$$

also known as "the **posterior probability** that $X = x$ after $S = s$ has been observed.

After seeing $S = s$, let $a^*(s)$ be the solution to the problem

$$\max_{a \in A} \sum_x u(a, x)\beta(x|s).$$

Here is the important result.

If $a^*(s)$ solves $\max_a \sum_x u(a, x)\, \beta(x|s)$ at each s with $\pi(s) > 0$, then $a^*(\cdot)$ solves the original maximization problem in (‡).

The reason is that $E\, u(a^*(S), X)$ is the weighted average $\sum_s \pi(s) \sum_x u(a^*(s), x)\beta(x|s)$, and the only way to maximize this weighted average is to maximize each part of it that receives a positive weight.

Valuing Information

Finally, let $V^\circ = \max_{a \in A} E\, u(a, X) = \max_{a \in A} \sum_x u(a, x)\beta(x)$ be the maximal expected value when the decision maker has no signal and $V^* = E\, u(a^*(S), X)$ be their maximal expected value when they can observe the value of the signal S before choose their action. We know that $V^* \geq V^\circ$ (look at the weighted average argument).

We define the **value of the information structure** as $V^* - V^\circ$.

Problems

For much of this set of problems, your job will be to figure out the best response to different beliefs and then to put them together in the right way.

1. [Infrastructure investment] $X = G$ or $X = B$ corresponds to the future weather pattern, the actions are to Leave the infrastructure alone or to put in New infrastructure, and the signal, s, is the result of investigations and research into the distribution of future values of X. The utilities $u(a, x)$ are given in the following table where c is the cost of the new infrastructure.

	Good	Bad
L	10	6
N	$(10 - c)$	$(9 - c)$

   a. After you see a signal $S = s$, you form your posterior beliefs $\beta(G|s)$ and $\beta(B|s)$. For what values of $\beta(G|s)$ do you optimally choose to leave the infrastructure as it is? Express this inequality as "L if $c > M$," give the value of M in terms of $\beta(G|s)$, and interpret. [You should have something like "leave the old infrastructure as it is if costs are larger than expected gains."]
   Ans. $E\, u(L, (\beta, (1 - \beta))) = 10\beta + 6(1 - \beta) = 6 + 4\beta$, $E\, u(N, (\beta, (1 - \beta))) = 10\beta + 9(1 - \beta) - c = 9 + \beta - c$. Thus, L is the better decision if

$$6 + 4\beta > (\geq) 9 + \beta - c, \text{ that is, if } c > 3 - 3\beta = 3(1 - \beta).$$

   Since $(1 - \beta)$ is the probability that the state ends up Bad, this policy is "L if costs exceed the expected benefits."

   b. If $c = 0.60$, that is, if the new infrastructure costs 20% of the damages it prevents, give the set of $\beta(G|s)$ for which it optimal to leave the old infrastructure in place.
   Ans. $0.60 > 3(1 - \beta)$ iff $0.20 > (1 - \beta)$ iff $\beta > 0.80$.

For the rest of this problem, assume that $c = 0.60$.

c. Suppose that the original or prior distribution has $\beta(G) = 0.75$ so that, without any extra information, one would put in the New infrastructure. We now introduce some signal structures. Suppose that we can run test/experiments that yield $S = s_G$ or $S = s_B$ with $P(S = s_G|G) = \alpha \geq \frac{1}{2}$ and $P(S = s_B|B) = \gamma \geq \frac{1}{2}$. The joint distribution, $q(\cdot, \cdot)$, is

	Good	Bad
s_G	$\alpha \cdot 0.75$	$(1-\gamma) \cdot 0.25$
s_B	$(1-\alpha) \cdot 0.75$	$\gamma \cdot 0.25$

Give $\beta(\cdot|S = s_G)$ and $\beta(\cdot|S = s_B)$. Verify that the average of the posterior beliefs is the prior, that is, verify that $\sum_s \pi(s)\beta(\cdot|x) = \beta(\cdot)$.

Ans. Using Bayes' Law,

$$\beta(\cdot|S = s_G) = (\tfrac{\alpha \cdot 0.75}{\alpha \cdot 0.75 + (1-\gamma) \cdot 0.25}, \tfrac{(1-\gamma \cdot 0.25}{\alpha \cdot 0.75 + (1-\gamma) \cdot 0.25}),$$
$$\beta(\cdot|S = s_B) = (\tfrac{(1-\alpha) \cdot 0.75}{(1-\alpha) \cdot 0.75 + \gamma \cdot 0.25}, \tfrac{\gamma \cdot 0.25}{(1-\alpha) \cdot 0.75 + \gamma \cdot 0.25}),$$

The weighted average is $\pi(s_G) \cdot \beta(\cdot|S = s_G) + \pi(s_B) \cdot \beta(\cdot|S = s_B)$. To calculate this, note that the denominator in $\beta(\cdot|S = s_G)$ is $\pi(s_G)$ while the denominator in $\beta(\cdot|S = s_B)$ is $\pi(s_B)$. This means that the weighted average is

$$((1-\alpha) \cdot 0.75 + \alpha \cdot 0.75, (1-\gamma) \cdot 0.25 + \gamma \cdot 0.25) = (0.75, 0.25).$$

d. Show that if $\alpha = \gamma = \frac{1}{2}$, then the signal structure is worthless.

Ans. Substituting into the beliefs found using Bayes' Law in the previous problem, $\beta(\cdot|s_G) = \beta(\cdot|s_B) = (0.75, 0.25)$. Because there is no change between the prior and the posterior, the optimal plan ignores the signal, hence leads to the utility V° (as defined above).

e. Give the set of $(\alpha, \gamma) \geq (\frac{1}{2}, \frac{1}{2})$ for which the information structure strictly increases the expected utility of the decision maker. [You should find that what matters for increasing utility is having a positive probability of changing the decision.]

Ans. The optimal decision before receiving the signals is N. For any $(\alpha, \gamma) \geq (\frac{1}{2}, \frac{1}{2})$, after seeing the signal s_B, N is still optimal. Hence, it is only after seeing the signal s_G that there is any chance to improve expected utility. From the earlier problem, we know that L is the optimal decision for all $\beta(G) > 0.80$, thus we need

$$\tfrac{\alpha \cdot 0.75}{\alpha \cdot 0.75 + (1-\gamma) \cdot 0.25} > 0.80.$$

Rearranging, this inequality is satisfied iff $\frac{3}{4}\alpha + \gamma > 1$.

2. [A more theoretical problem] We say that signal structure A is unambiguously better than signal structure B if every decision maker, no matter what their action set and what their utility functions are, would at least weakly prefer A to B. This problem asks you to examine some parts of this definition.

Let us suppose that the random variable takes on two possible values, x and x', with probabilities $p > 0$ and $p' = (1-p) > 0$, that there are two possible actions, a and b, that the signals take on two values s and s', that the signal structure is given by

	x	x'
s	αp	$(1-\gamma)p'$
s'	$(1-\alpha)p$	$\gamma p'$

and that the utilities are given by

	x	x'
a	$u(a,x)$	$u(a,x')$
b	$u(b,x)$	$u(b,x')$

We assume that $\alpha > \frac{1}{2}$ and $\gamma > \frac{1}{2}$ and we let $V(\alpha,\gamma)$ denote the maximal expected utility when the signals are distributed as above and the utilities are above.

a. Show that $V(\alpha,\gamma) \geq V(\frac{1}{2},\frac{1}{2})$, that is, the information structure with $\alpha > \frac{1}{2}$ and $\gamma > \frac{1}{2}$ is unambiguously better than the structure with $\alpha = \gamma = \frac{1}{2}$.
Ans. If $\alpha = \gamma = \frac{1}{2}$, then $\beta(\cdot|s) = \beta(\cdot|s') = (p,p')$, that is, the posterior beliefs are exactly the same as the prior information. Since information can always be ignored, for any $\alpha > \frac{1}{2}$ and $\gamma > \frac{1}{2}$, we know that the optimal decisions must be at least as good, in terms of expected utility, as $V(\frac{1}{2},\frac{1}{2})$.

b. Give utilities such that $V(\alpha,\gamma) > V(\frac{1}{2},\frac{1}{2})$, that is, show that some decision maker strictly prefers the information structure with $\alpha > \frac{1}{2}$ and $\gamma > \frac{1}{2}$ to the structure with $\alpha = \gamma = \frac{1}{2}$.
Ans. To give this answer simply, it is best to re-organize the problem. The re-organization takes a bit.
Consider the payoff matrix

	x	x'
a	$u(a,x)$	$u(a,x')$
b	$u(b,x)$	$u(b,x')$

If we simultaneously have $u(a,x) > u(b,x)$ and $u(a,x') > u(b,x')$, then, for all beliefs about x and x', the decision a is best, so that all information structures are worthless because they will not change the optimal decision. The same is conclusion about information is true if we simultaneously have $u(a,x) < u(b,x)$ and $u(a,x') < u(b,x')$ because, in this case, b is the best decision "no matter what." To proceed, let us assume that a is the best choice if x is true and b is the best choice if x' is true. There is no harm in this, we could always relabel the choices or the states and make this true. Numerically, there is no loss in simultaneously assuming that

$$u(a,x) > u(b,x) \text{ and } u(b,x') > u(a,x').$$

Next notice that if we add/subtract a constant from the x column or from the x' column in the payoff matrix, we do not change the optimal decision for any beliefs $(\beta,(1-\beta))$. That is, for any κ,κ', positive, negative or zero, the optimal decisions, at any β, for the following the matrix are the same as the optimal decisions for the matrix that we started with.

	x	x'
a	$u(a,x)-\kappa$	$u(a,x')-\kappa'$
b	$u(b,x)-\kappa$	$u(b,x')-\kappa'$

To make the notation simpler, let us set $\kappa = u(b,x)$, $\kappa' = u(a,x')$, and, using both $u(a,x) > u(b,x)$ and $u(b,x') > u(a,x')$, we can take the payoff matrix to be, for some $r, s > 0$,

	x	x'
a	r	0
b	0	s

Now the problem reduces to finding $r, s > 0$ such that for fixed $(\alpha,\gamma) > (\frac{1}{2},\frac{1}{2})$, the decision changes with positive probability. For posterior beliefs $(\beta,(1-\beta))$, the decision a is strictly better than the decision b iff

$$\beta r > (1-\beta)s, \text{ that is, if } \beta > \tfrac{s}{s+r}.$$

Pick $r, s > 0$ such that the initial beliefs (p,p') are equal to $(\frac{s}{s+r},\frac{r}{s+r})$. This means that the optimal decision before information was either a or b, but after information, it becomes strictly optimal to pick a after signal s and b after signal s'.

c. Show that for any utilities, if $\alpha' > \alpha \geq \frac{1}{2}$ and $\gamma' > \gamma \geq \frac{1}{2}$, then $V(\alpha',\gamma') \geq V(\alpha,\gamma)$.
Ans. This was very simple above where we had $\alpha = \frac{1}{2}$ and $\gamma = \frac{1}{2}$. The argument here has to be a good bit more subtle. The function $V(\beta) = \max\{\beta r,(1-\beta)s\}$ is convex. Let $Q_{\alpha,\gamma}$ be the distribution assigning mass $\pi(s) = \alpha p + (1-\gamma)p'$ to $\beta := \beta(\cdot|s)$ and mass $\pi(s') = (1-\alpha)p + \gamma p'$ to $\beta' := \beta(\cdot|s')$. The expected utility to the information structure is

$$\pi(s)V(\beta) + \pi(s')V(\beta').$$

Now, in moving from (α,γ) to (α',γ'), the posterior beliefs maintain the same average, (p,p'), but have moved strictly out toward the edges, toward certainty of either x or x'. By the arguments we gave for Jensen's inequality, this means that the expected utility is at least as high under $Q_{\alpha',\gamma'}$ as it was under $Q_{\alpha,\gamma}$.

d. Give utilities such that $V(\alpha',\gamma') \geq V(\alpha,\gamma)$ for $\alpha' > \alpha \geq \frac{1}{2}$ and $\gamma' > \gamma \geq \frac{1}{2}$.
Ans. As before, pick $r, s > 0$ so that the decision maker is indifferent after seeing at least one of the signals when the information structure is given by α and γ. Making the information more precise then changes that decision, hence increases expected utility.

3. [The value of repeated independent observations] This problem continues where the first problem left off. Now suppose that the test/experiment can be run twice and that the results are independent across the trials. Thus, $P(S = (s_G,s_G)|G) = \alpha^2$, $P(S = (s_G,s_B)|G) = P(S = (s_B,s_G)|G) = \alpha(1-\alpha)$, and $P(S = (s_B,s_B)|G) = (1-\alpha)^2$ with the parallel pattern for B.

a. Fill in the probabilities in the following joint distribution $q(\cdot,\cdot)$. Verify that the average of posterior beliefs is the prior belief.

	Good	Bad
(s_G,s_G)		
(s_G,s_B)		
(s_B,s_G)		
(s_B,s_B)		

b. Give $\beta(G|(s_G, s_G))$, $\beta(G|(s_G, s_B))$, $\beta(G|(s_B, s_G))$, and $\beta(G|(s_B, s_B))$.
Ans. Filling in the table yields

	Good	Bad
(s_G, s_G)	$\alpha^2 \cdot 0.75$	$(1-\gamma)^2 \cdot 0.25$
(s_G, s_B)	$\alpha(1-\alpha) \cdot 0.75$	$\gamma(1-\gamma) \cdot 0.25$
(s_B, s_G)	$\alpha(1-\alpha) \cdot 0.75$	$\gamma(1-\gamma) \cdot 0.25$
(s_B, s_B)	$(1-\alpha)^2 \cdot 0.75$	$\gamma^2 \cdot 0.25$

From these
- $\beta(G|(s_G, s_G)) = \frac{3\alpha^2}{3\alpha^2+(1-\gamma)^2}$,
- $\beta(G|(s_G, s_B)) = \frac{3\alpha(1-\alpha)}{3\alpha(1-\alpha)+\gamma(1-\gamma)}$,
- $\beta(G|(s_B, s_G)) = \frac{3\alpha(1-\alpha)}{3\alpha(1-\alpha)+\gamma(1-\gamma)}$, and
- $\beta(G|(s_B, s_B)) = \frac{3(1-\alpha)^2}{3(1-\alpha)^2+\gamma^2}$.

c. Show that if $\alpha = \gamma = \frac{1}{2}$, then the signal structure is worthless.
Ans. In all of the cases above, if $\alpha = \gamma = \frac{1}{2}$, then the posterior distribution is $(0.75, 0.25)$, the same as the prior, hence no decision is ever changed, hence there is no value to the signal structure.

d. Give the set of $(\alpha, \gamma) \geq (\frac{1}{2}, \frac{1}{2})$ for which the information structure strictly increases the expected utility of the decision maker.
Ans. If one followed the strategy of ignoring the second observation, one would be using the optimal strategy for the first problem. For that problem, we know that any α, γ pair satisfying $\frac{3}{4}\alpha + \gamma > 1$ gives a strict expected utility improvement. This must mean that any such α, γ gives a strict expected utility improvement. The question is whether there are any other α, β pairs. The best way to proceed here is to suppose that one starts with one signal only, note that the best strategy is L if the first signal is s_G and N if the first signal is s_B. A strict improvement in expected utility happens if there is a positive probability that the second signal changes the strategy. We handle the cases in order. Note that, from the first problem, the condition for equality of the expected utility with one signal is $\frac{3}{4}\alpha + \gamma = 1$.
- If the first signal is s_G and the second is s_G, then belief that good will happen is $\frac{3\alpha^2}{3\alpha^2+(1-\gamma)^2}$. Now,
$$\frac{3\alpha^2}{3\alpha^2+(1-\gamma)^2} > \frac{3\alpha}{3\alpha+(1-\gamma)},$$
to see why, multiply the top and bottom of $\frac{3\alpha}{3\alpha+(1-\gamma)}$ by α and note that $\alpha(1-\gamma) > (1-\gamma)^2$ because $\alpha > \frac{1}{2} > (1-\gamma)$, hence the denominator has gone up. Thus, the decision will not be changed.
- If the first signal is s_G and the second is s_B, or if the first signal is s_B and the second is s_G, then belief that good will happen is $\frac{3\alpha(1-\alpha)}{3\alpha(1-\alpha)+\gamma(1-\gamma)}$. If this is greater than 0.8, then it will change the decision in the case (s_B, s_G), if it is less than 0.8, then it will change the decision in the case (s_G, s_B). To complete the problem, we only need examine the borderline case that it is equal to 0.80. Note that

$$\frac{3\alpha}{3\alpha+(1-\gamma)} = \frac{3\alpha(1-\alpha)}{3\alpha(1-\alpha)+(1-\alpha)(1-\gamma)} = \frac{3\alpha^2}{3\alpha^2+\frac{(1-\alpha)}{\gamma}\gamma(1-\gamma)} > \frac{3\alpha(1-\alpha)}{3\alpha(1-\alpha)+\gamma(1-\gamma)}$$

where the last inequality comes from $\frac{(1-\alpha)}{\gamma} < 1$ (so we are decreasing the denominator). Thus, if $\beta(G|s_G, s_B) = 0.80$, then $\frac{3\alpha}{3\alpha+(1-\gamma)} > 0.80$, so that the information is definitely valuable.

- If the first signal is s_B and the second is s_B, there is no change in the optimal action since $\beta(G|s_B, s_B) < \beta(G|s_B)$.

e. Explain why the set is larger here than it was in the previous problem.
Ans. This was, implicitly at least, just answered. One can ignore the information in the second signal, if you ever use it, then it is valuable. We do not ignore it if $\frac{3\alpha(1-\alpha)}{3\alpha(1-\alpha)+\gamma(1-\gamma)} < 0.80$ or if $\frac{3\alpha(1-\alpha)}{3\alpha(1-\alpha)+\gamma(1-\gamma)} > 0.80$, and if $\frac{3\alpha(1-\alpha)}{3\alpha(1-\alpha)+\gamma(1-\gamma)} = 0.80$, then we know that the information in the first signal is already useful.

4. You are the only builder of houses for a planned sub-division. There is a %40 chance that the demand function for houses will be low, $P = 300,000 - 300 \cdot Q$, and a %60 chance that it will be high, $P = 400,000 - 250 \cdot Q$. Your cost function is $C(Q) = 1,400,000 + 120,000Q$. You have to make the decision about how many to build before you know whether the demand will be low or high,

a. How many houses should you build to maximize your expected profits? What are your maximal expected profits?
Ans. To unify all the algebra, let

$$\Pi(a, b, c, F) = \max_Q \,(a - bQ)Q - (F + cQ).$$

After a bit of work, you should find that $\Pi = \frac{(a-c)^2}{4b} - F$. For this problem, define

$$V(\beta) = \Pi(a, b, c, F)$$

where $a = 300,000\beta + 400,000(1-\beta)$, $b = 300\beta + 250(1-\beta)$, $c = 120,000$ and $F = 1,400,000$. $V(\beta)$ gives your expected profits when your belief that Low demand will happen is β.
The specific answer for this problem is $V(0.40)$ because, with no information, you believe that Low demand will happen with probability 0.40.

b. How much would you be willing to pay for a "crystal ball" forecast of demand conditions, that is, a perfectly accurate forecast?
Ans. $(0.40V(1) + 0.60V(0)) - V(0.40)$.

c. An econometric forecaster offers to do a study of the market. The forecasts are not guaranteed accurate, rather they have the probability distribution given in the following table.

	Low Demand	High Demand
Low forecast	0.3	0.2
High forecast	0.1	0.4

If you had access to this forecasting service, what would your maximal expected profits be? What would your willingness to pay for this forecast be?
Ans. Beliefs of L are 0.60 after a Low forecast, and 0.20 after a High forecast. Therefore maximal expected profits are $0.50V(0.60) + 0.50V(0.20)$. The most you would be willing to pay for this forecast is $V(0.40) - (0.50V(0.60) + 0.50V(0.20))$.

5. From Ch. 18 of Kreps's *Micro for Managers*, Problem 18.2(a) and (b).

Ans. For a consumer with probability p of a loss, the expected utility of no insurance is

$$EU_{no} = p\sqrt{0 + 10,000} + (1-p)\sqrt{80,000 + 10,000} = 100p + 300(1-p),$$

the expected utility of partial insurance is

$$EU_{partial} = 250p + 290(1-p), \text{ and}$$

the expected utility of total insurance is

$$EU_{total} = \sqrt{80,000 - 11,600 + 10,000} = 280.$$

For Reece's $p = 0.10$, this is a comparison of $10 + 270 = 280$, 286 and 280, so partial insurance is the choice. For Yost's $p = 0.03$, this is a comparison of $3 + 291 = 294$, 288.80 and 280, so no insurance is the choice.

More generally, graph the expected utilities as a function of $p \in [0, 0.40]$, and note the regions in which each is larger. For $0 \leq p < \frac{1}{16}$, no insurance is better, for $\frac{1}{16} < p < \frac{1}{4}$, partial insurance is better, for $\frac{1}{4} < p$, full insurance is best.

6. From Ch. 18 of Kreps's *Micro for Managers*,
   a. Problem 18.3,
   Ans. The point is to manage the adverse selection by not paying any purchasers who die in the first 2 years.
   b. Problem 18.5, and
   Ans. Blue Shield is comparing the average employee, who must at least be healthy enough to work, to the people who come in looking for insurance, and this last selects for people who know that their health care expenses are likely to be high. For the second part of the question, flexibility allows people who are expecting a particular type of health care cost to be large to buy more of that coverage. One effect is that coverage is a better value to the employees so they are happier, another is that the pool of people signing up for any particular benefit is likely to be more expensive to the insurance company, they have adverse selection that way.
   c. Problem 18.6.
   Ans. Bad cars make people want to get rid of them. A car that has been "gotten rid of" more ofter is more likely to be one that people want to be rid of.

7. By putting in effort $0 \leq e \leq \overline{e}$, a business can reduce the probability of fire in the warehouse to $P(e)$ where $P'(e) < 0$ and $P''(e) > 0$. If there is a fire, there will be a loss $L > 0$ and profits will be $(R - L) > 0$, if there is not a fire, they will be R. The utility costs of effort to firm θ, $0 < \theta < 1$, are $(1-\theta)c(e)$ where $c'(e) > 0$ and $c''(e) > 0$. Higher θ's correspond to lower costs of prevention, so these are "better" firms. One simple form of insurance policy costs C and has a deductible D, that is, it reduces losses from L to D by paying $(L - D)$ to an insured firm in case of a fire causing loss L. As is only sensible, $C < D < L$.

   The problem for firm θ without insurance is

$$E\Pi(\theta) = \max_{0 \leq e \leq \overline{e}} [P(e)(R - L) + (1 - P(e))R] - (1-\theta)c(e),$$

   while the problem for firm θ with insurance is

$$E\Pi_{ins}(\theta) = \max_{0 \leq e \leq \overline{e}} [P(e)((R - C) - D) + (1 - P(e))(R - C)] - (1-\theta)c(e).$$

a. For a firm with a given θ, compare their optimal effort if they have insurance to their optimal effort if they do not have insurance. Explain. [The difference you find is due to what is called 'moral hazard.']

Ans. Let $\mathrm{M}(e, \theta) = E\,\Pi(\theta) - E\,\Pi_{ins}(\theta)$ and consider the problem

$$f(e, t) = E\,\Pi_{ins}(\theta) + t \cdot \mathrm{M}(e, \theta).$$

For $t = 1$, we have the problem without insurance, for $t = 0$, we have the problem with insurance. The part of $mess(e, \theta)$ that includes e is $D \cdot P(e)$, to the part of $f(\cdot, \cdot)$ that includes both t and e is $(D - L)tP(e)$. Since $P'(e) < 0$ and $(D - L) < 0$, this is supermodular, so when $t = 1$, e^* is higher, that is, the un-insured work harder at preventing fires.

b. For a given insurance policy with cost C and deductible D, which firms, indexed by their θ's, buy insurance? What happens to their riskiness after they buy insurance?

Ans. In both $E\,\Pi(\theta)$ and $E\,\Pi(\theta)$, the objective function (thing being maximized), is supermodular in θ and e, that is, the higher θ, lower-cost-of-effort firms put in more e. By putting in more e, they lower their risk of fire. This means that the high θ firms find insurance less valuable for two reasons: they have lower risk; and they have lower potential cost savings to having lower effort after they buy insurance. The ones who do by insurance, by the previous part of the problem, slack off on their e.

c. An actuarially fair insurance policy is one in which the cost exactly covers the expected payments. In more detail, if p is the average probability of loss among the firms that buy insurance, and $(L - D)$ is what the insurance company pays out every time there is a fire, then $C = p(L - D)$. Verbally explain why the previous results imply that an actuarially fair policy may not insure very many firms.

Ans. An actuarially fair policy selects for the low θ firms, these are the firms that will put in less e whether or not they have insurance, meaning that premiums must be high. But high premiums discourage exactly the lower risk firms that the insurance company would like to sell policies to.

8. From Ch. 18 of Kreps's *Micro for Managers*, Problem 18.8.
9. Mazzucato argues, in Ch. 2 of *The Entrepreneurial State*, that when one looks across firms, the correlation between R&D expenditures and growth is very small unless one considers a special subset of the firms. Why should the selection bias exist and how does she solve it?

   Ans. In her discussion of "Myth 1," she notes that there must be a supermodular relation between R&D and other productive and growth-enhancing activities. For example, only among the 'persistent patenters' in the pharmaceutical industry does one find a positive correlation between R&D and growth rates, and this selects for firms that are set up to take advantage of innovation.
10. Mazzucato argues, in Ch. 3 of *The Entrepreneurial State*, that venture capitalists tend to come into the process of turning new scientific knowledge into new products very late in the process and that they tend to have a time horizon much shorter than is usually needed to bring the products to market. What mechanisms does she describe for venture capitalists that allow them to get out, with a profit, before the product comes to market?

    Ans. Apologies, this is in Ch. 2, in "Myth 3: Venture Capital is Risk-Loving." She discusses venture capitalists' bias toward three-to-five year windows, which makes them less likely to invest in longer-term projects. However, the venture capitalists can often arrange to sell the firms and get out of the business *before* there is a product. They provide a bridge between the early stages in firm development and the product stage, providing the managerial expertise needed to make the initial sales successful.

11. Mazzucato argues, in Ch. 7 of *The Entrepreneurial State*, that high aggregate levels of R&D is not enough for a country to benefit from innovation. What other factors does she name that have the property that $\partial^2 Ben(R,x)/\partial R \partial x > 0$? Here R measures R&D and x measures the other factors.

    Ans. In the case of the Danish wind turbine industry, she argues that the government brought in the large manufacturers which allowed them to gain experience with the technology and develop supply chains. The two x's here are the diffusion of experience and knowledge, and the supply chains. In the case of the German wind turbine industry, she discusses the guaranteed 20-year investment incentive horizons, which reduces the uncertainty about entering or expanding in the market. Here the x is the support of a market for the results of R&D. More generally, she credits the existence/non-existence of a "long-term vision for energy transition" as a crucial x.