The Economics of Cybersecurity — Lecture 3 Notes

Adam Hastings

January 23, 2024

Pre-Class

- Write title, course number, hours, on blackboard
- Write out sections of discussion

Would you rather question: Would you rather code every day for the rest of your life or never code again?

1 The Market for Lemons

1.1 Opening questions

- What did people think?
- It was "straightforward" microeconomics but it makes assumptions and assumes that the reader knows what these assumptions are. Did anyone struggle to follow any of the arguments being made or struggle to follow the math? I think the math is deceptively simple.
- Akerlof won the Nobel Memorial Prize in Economics for basically this paper alone (maybe a couple others, but on this same topic). Is that surprising to you? (Note: Shared it with Stiglitz—a Columbia professor!)
  - Side note–The Nobel Prize in Economics is officially the Sveriges Riksbank Prize in Economic Sciences in Memory of Alfred Nobel. Not one of the original Nobel Prizes! Administered by a bank!

1.2 Lemons Model

I want to cover this model in full detail because it's such a bite-sized example of how we can use mathmatical modeling to explain and illustrate things. It might feel like excruciating detail since

you already read the paper but like I said it has a lot of hidden assumptions and by going through the model we can air out and examine what these assumptions are.

1.2.1 Asymmetric Information

First, there's the assumption that the market is going to settle on a fixed price p.

We also have two groups. Group 1 has a utility function of

$$U_1 = M + \sum_{i=1}^{N} x_i$$

while Group 2 has a utility funciton of

$$U_1 = M + \sum_{i=1}^{N} \frac{3}{2} x_i$$

Recall that utility is a quantitative representation of goodness, and that more is better.

Which group values cars more? Group 2.

Which group has cars? Group 1.

What does this mean about the inital allocation of goods? (This was a homework quesiton.) It means the initial allocation is Pareto inefficient. This means we can make a Pareto optimization. If Group 2 collectively bought all the cars at some price p where $1 < p < 3/2$, then Group 1's utility is higher because they were paid more for their cars than their cars were worth to them, *and* Group 2's utility is higher because they paid *less* for their cars than the cars were worth to them. *Does that make sense?*

Why call this Group 1 and Group 2? Why not just call them "buyers" and "sellers"? I think this is because calling them "buyers" and "sellers" makes it seem like Group 1 wants to sell cars and Group 2 wants to buy cars. That is true but slightly misleading. Group 1 still values cars (as per their utility funciton), and according to their utility function they still would actually buy a car if it was a good deal. Calling Group 1 the "sellers" makes it sound like they're going to sell no matter the price, which simply isn't true. Framing things using a utility function allows the model remain simple and allows us to assume that Groups 1 and 2 are the same, just with different valuations of cars.

For example, both groups have demand functions! Let's take a look at them.

1.2.2 Group 1

Group 1's demand function looks like this:

$$D_1(p) = Y_1/p \qquad \mu/p > 1$$
$$D_1(p) = 0 \qquad \mu/p < 1$$

What does $\mu/p > 1$ mean? I found it easier to rewrite it as $\mu > p$. Recall that quality and price are not the same units but are on the range of 0 to 2. So $\mu > p$ just means that average quality is higher than the price, and $\mu < p$ means that average quality is lower than the price.

What happens when the average quality is lower than the price? Well this depends on Group 1's utility function. The scaling factor for each car in Group 1's utility function is 1. Hence for *Group 1*, $\mu/p > 1$ is a bad deal. In the case of a bad deal, what is the demand going to be? 0. Hence $D_1(p) = 0, \quad \mu/p < 1$ as above.

What about the other case, where $\mu/p > 1$? Akerlof introduces new variables, Y_1 and Y_2 to represent Group 1 and Group 2's income respetively, including the income that results from selling cars.

So we know that $\mu/p > 1$ is a good deal for Group 1 members, and they're going to want cars at this price, so what will the demand be? Somewhere in the above assumptions made is that cars are good, more cars are better, and that the value of each additional car is equal to its quality. So when the price is a good deal, Group 1 wants as many cars as they can get with their Y_1 income. And the amount they can get is Y_1/p. Does that make sense?

That's how Akerlof constructs the demand curves for Group 1. How does he construct the supply curve?

We're assuming linear utility. If the price $p = 0$, there are going to be 0 cars sold. If the price is $p = 2$, there are going to be N cars sold. This is a straight line with slope $N/2$ (*Draw this on the board with price p on x-axis*). Hence

$$S_1 = \frac{pN}{2} \quad p \leq 2 \tag{1}$$

Note that I think this is a typo! Akerlof writes "S_2" instead of "S_1"! This was the extra credit question in the homework. Good job if you spotted this. If the price $p > 2$, we can just assume that all N cars will be sold.

What is the interpretation of the above? It means that the supply is subject to the price. This is just mathematically encoding the idea that sellers will only sell if the market price is above their car's quality, and will hold onto the car if the market price is below their car's quality.

There's another derived equation here, for average quality:

$$\mu = p/2 \tag{2}$$

Where does this come from? Recall that that at any price p, the cars that are sold will be of quality $x_i \leq p$ (which is valid because price and quality are normalized to the same [0,2] scale). If we draw out the PDF of the quality of cars that will sell at price p, what is the average quality? $p/2$.

1.2.3 Group 2

Now let's look at the supply and demand curves for Group 2. What is a "good deal" for Group 2? How much utility do they get from a car? They get $\frac{3}{2}x_i$. So on average they get $\frac{3}{2}\mu$ for each car. So they will buy when $\frac{3}{2}\mu > p$. As with Group 1, there demand is subject to their own income Y_2 and the price of the cars p.

$$\begin{aligned} D_1 = Y_2/p \qquad & \frac{3}{2}\mu > p \\ D_1 = 0 \qquad & \frac{3}{2}\mu < p \end{aligned}$$

And Group 2 has no cars so

$$S_2 = 0$$

Now we can write the combined demand function $D(p, \mu) = D_1 + D_2$. This follows straightforwardly from the above two demand curves.

If $p < \mu$, then both Group 1 and Group 2 want cars, so $D(p, \mu) = Y_1/p + Y_2/p = (Y_1 + Y_2)/p$.

If $\mu < p < \frac{3\mu}{2}$, then only Group 2 wants cars, so $D(p, \mu) = Y_2/p$.

If $p > \frac{3\mu}{2}$, then $D(p, \mu) = 0$.

This is the full model in the case of the asymmetric information, where Group 1 only sells cars if the market price is beneath their car's value.

But price is p while average quality is $\mu = p/2$. Of the three demand curve cases above, which one does this correspond to?

Let's rewrite the above inequalities to be in terms of p instead of μ:

- Does $p < (p/2)$? No, never for a positive p. So this case doesn't apply.
- Does $\frac{p}{2} < p < \frac{3}{2}(\frac{p}{2})$? No, never.
- The last one holds, though: $p > \frac{3\mu}{2} \equiv p > \frac{3p}{4}$. What is the combined demand in this case? It's 0.

No sales take place despite the fact that there are Group 1 members who have cars they are willing to sell at prices Group 2 members are willing to pay.

In other words, in this model, there are Pareto optimizations that could occur but the market does not produce this outcome. This system remains Pareto inefficient. This is why this situation is called a market failure.

1.2.4 Symmetric Information

Akerlof also shows what happens when there is symmetric information. What changes?

$$\begin{aligned} S(p) &= N & p &> 1 \\ S(p) &= 0 & p &< 1 \end{aligned}$$

It's a step function. Group 1 will sell when the price is greater than 1.

The demand curves are similar:

$$\begin{aligned} D(p) &= (Y_1 + Y_2)/p & p &< 1 \\ D(p) &= (Y_2)/p & 1 < p &< 3/2 \\ D(p) &= 0 & p &> 3/2 \end{aligned}$$

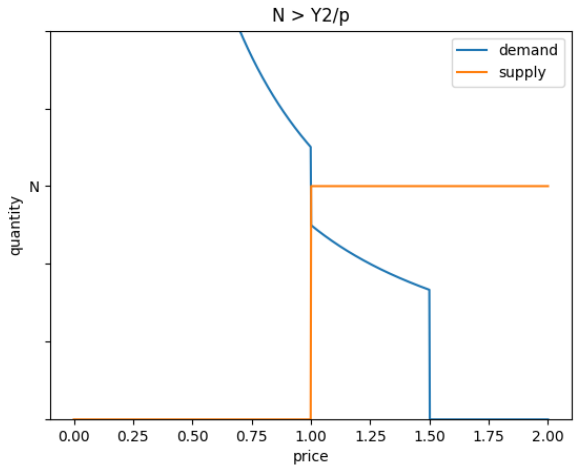
In classic microeconomics, the price of a good is the point at which the supply curve and the demand curve intersect. When does $S(p) = D(p)$?

Let's break this down by cases. If $p < 1$, then supply is 0 and demand is $(Y_1 + Y_2)/p$, which will be greater than 0 if we assume that these three variables are greater than 0. There will be no intersection of supply and demand in this range.

Likewise, if $p > 3/2$, demand is 0 but supply is N. Again we assume $N > 0$ so there will be no intersection of supply and demand here.

That means that the intersection of supply and demand will be between $p = 1$ and $p = 3/2$. What are the conditions needed for p to be in this price range?

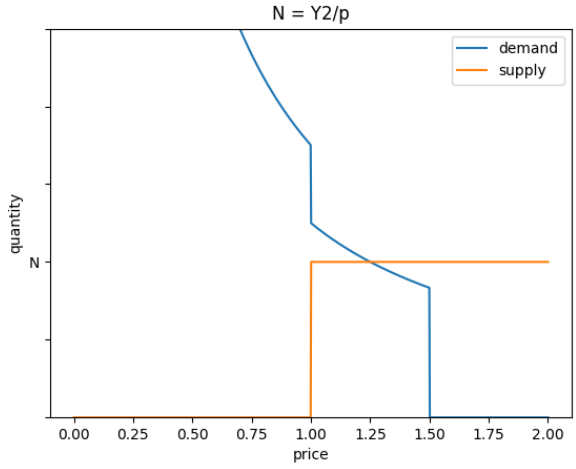
There are three possible cases. Case 1 is that $N > Y_2/p$:



In this case, supply and demand meet when $p = 1$. And hence $N > Y_2/p$,

$$p = 1 \quad \text{if} \quad Y_2 < N \tag{3}$$

Case 2 is when $N = Y_2/p$:

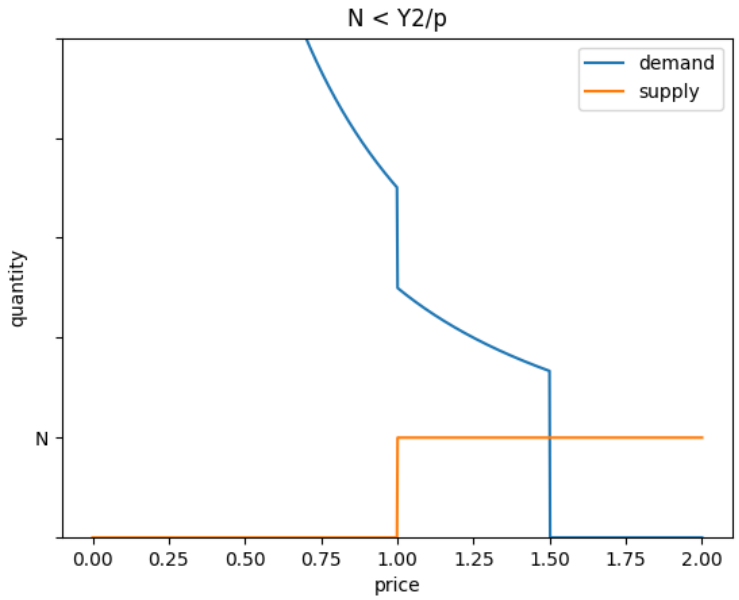


This case requires that $1 < p < 3/2$. So if $N = Y_2/p \implies p = Y_2/N$, then we get

$$1 < \frac{Y_2}{N} < \frac{3}{2}$$
$$1 > \frac{N}{Y_2} > \frac{2}{3}$$

$$\frac{2Y_2}{3} < N < Y_2 \tag{4}$$

Case 3 is when $N < Y_2/p$:



This means that $p = \frac{3}{2}$. Plugging into the above and rearranging yields:

$$N < \frac{2Y_2}{3} \tag{5}$$

But in all cases there is a point at which supply equals demand. In economics this is called a **market clearing price**. And just like the name implies, at this price, *all goods in the market will sell!*

What does this mean for the two groups' utility functions? Both groups are better off after the trade. We have made a Pareto improvement. I think it's even Pareto optimal, meaning that no further optimizations are possible.

1.3 Closing questions

Is this mindblowing or what? From these super simple models, Akerlof proves that in the presence of asymmetric information, markets are not an efficient way to allocate goods, and everyone is worse off than they could otherwise be as a result.

This is a great example of how to show a point using a mathematical model. A few quesitons come up:

1. Do you think Akerlof created the model first and then found out that it had this interesting property? (I highly doubt it. More likely that he had the intuition, and then created the model to support the intuition).

- What do we think about creating models to "mathify" an intuition we have? Is this just adding math as we see fit to make our arguments seem more impressive and more bulletproof?

2. What did we think about the example applications that Akerlof gives? (Insurance, hiring practices, cost of dishonesty, credit markets in underdeveloped economies)

3. What did we think about the proposed solutions? (brand repuation/chains, seller guarantees, credentials/licensing) (I thought they were pretty basic and obvious. That's fine though—not his job to fix in this paper!)

4. This paper was rejected twice before publication because the reviewers pointed out that if the model was correct, then no used cars would ever be sold. In other words, the real-world evidence that used cars do in fact sell means that this model is incorrect.

   - What do we think of this complaint?
   - Does this mean that one of the assumptions that Akerlof makes is wrong? Which one?
   - A surprising takeaway might be that your model doesn't even need to accurately reflect the real world to be useful. What are your thoughts on that?

2 The Economics of Information Security Investment

The other paper we read was in many ways very similar. It made a mathematical model of something, and then demonstrated how this model exhibits interesting behavior.

2.1 Opening discussion

- Focuses on the risk of securing a "dataset". But seems to be a much more general model than that? Seems to cover any cases where you have something worth protecting.

2.2 Model description

Three parameters:

- λ: The loss given that a breach occurs. Finite, less than some very large number M.
- t: Threat. Probability of a threat occurring. $t \in [0, 1]$. Pinned at a fixed value $t > 0$ (since no amount of investment is going to change t).
- v: Vulnerability. Probability that a threat is successful. $v \in [0, 1]$. Hence $v = 0 \implies$ perfect security, and $v = 1 \implies$ zero security i.e. public information.

Expected loss is therefore $\lambda t v$. The potential loss $L = t\lambda$.

Assumption: Investment can reduce the the vulnerability. Denoted by $z > 0$, in same units (i.e. dollars) as λ. Let $S(z, v)$ denote the probability that information set with vulnerability v will be breached.

Then the authors make three assumptions:

- A1: $S(z, 0) = 0$ for all z. *Ask: Can someone interpret what this means?* (It means that completely secure data remains perfectly secure regardless of the investment)
- A2: $\forall v, S(0, v) = v$. *Ask: Can someone interpret what this means?* (It means that if there is no investment, the probability of a breach is just equal to its vulnerability score.
- A3: $\forall v \in (0, 1), \forall z, S_z(z, v) < 0$ and $S_{zz}(z, v) > 0$, where S_z denotes the partial derivative w.r.t. z and S_{zz} denotes the partial derivative of S_z w.r.t. z. Interpretation: there are diminishing marginal returns to investment (example: $f(x) = 1/x$). Also assume that as $\forall v \in (0, 1,), z \to \infty \implies lim\ S(z, v) \to 0$, i.e. the probability of a breach can be arbitrarily close to 0 with sufficient investment.

What are our thoughts on these assumptions?

Next the authors assume that firms are risk neutral. *What does this mean?* (It was defined in a footnote). It means that you are indifferent towards risk as long as the expected value of interest remains constant. E.g. you would just as likely accept \$10 as you would a 50/50 chance between \$20 and \$10.

Firms are going to make the investment that maximizes their personal benefit. To do this, the authors define an expected benefit of investment in information security, EBIS. Recall that the original expected loss was $vt\lambda = vL$. So with investment the new loss is defined as

$$\text{EBIS}(z) = [v - S(z, v)]L$$

This is how much you expect to lose with investment z. But this doesn't take into account the actual investment itself. What firms really want to minimize is expected *net* benefit of investment in information security, or E*N*BIS, which is

$$\text{ENBIS}(z) \qquad = \qquad \text{EBIS}(z) - z \qquad = \qquad [v - S(z, v)]L - z$$

What firms want to do is find the $z^*(v)$ that minimizes ENBIS.

This is a concave fucntion (upside-down cup) starting at 0 and reaching some maximum before decreasing back to 0.

The derivative with respect to z is then

$$ENBIS_z(z^*) = S_z(z^*, v)L - 1 = 0$$

$$\implies \quad -S(z^*, v)L = 1$$

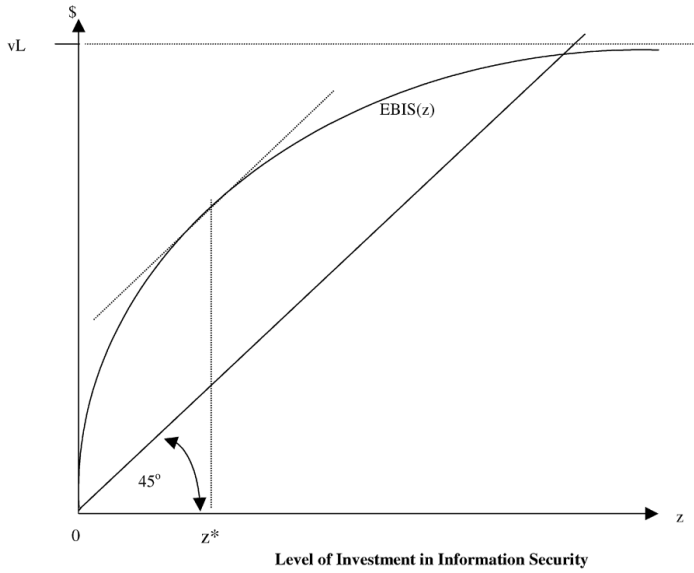


Fig. 1. The benefits and cost of investment in information security.

Initial thoughts:

- Wise to let λ cover all possible types of losses. Keeps the model simple.
- I think the model's biggest limitation is that you need to know the function $S(z, v)$. Very difficult (maybe impossible!) to do so. (but a few classes of candidate $S(z, v)$ functions are provided!)
- Economic modeling like this has a huge similarity with machine learning. What is it? (In both cases, the challenge is usually defining a loss function, and then achieving optimum points either by direct analysis or by computation).
- That brings up another point—-this is a "one round" (aka "one shot") game (unlike in ML).
  - What do we think of this? Is this realistic?
  - Security is an ongoing process. Attackers always responding to Defenders' behavior...
  - But at the same time, much easier to mathematically work with one shot games than to deal with differential equations.
  - What are the alternatives? Giant formulae? Really painful derivations? Simulation perhaps...? Many differential equations have closed form solutions.
  - A quote in a footnote: "A model is supposed to reveal the essence of what is going on: your model should be reduced to just those pieces that are required to make it work" (Varian). Thoughts?

2.3 Model Exploration

One thing this paper has that The Market for Lemons paper does not is model exploration.

In Section II of the paper (where the model is defined), the shape of $S(z, v)$ is left pretty undefined (some constraints on the sign of its first and second derivatives though).

2.3.1 Candidate 1 for $S(z, v)$

Two candidates for $S(z, v)$ are considered.

Candidate 1 is $S^I = \frac{v}{(\alpha z+1)^\beta}$, with $\alpha > 0, \beta \geq 1$. Should we check that this satisfies the above conditions?

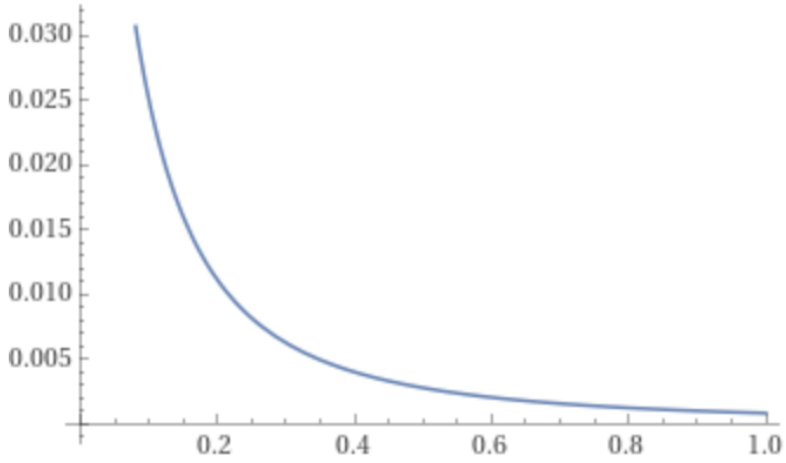


Figure 1: Class I candidate function for $S(z, v)$ (with z along the x-axis)

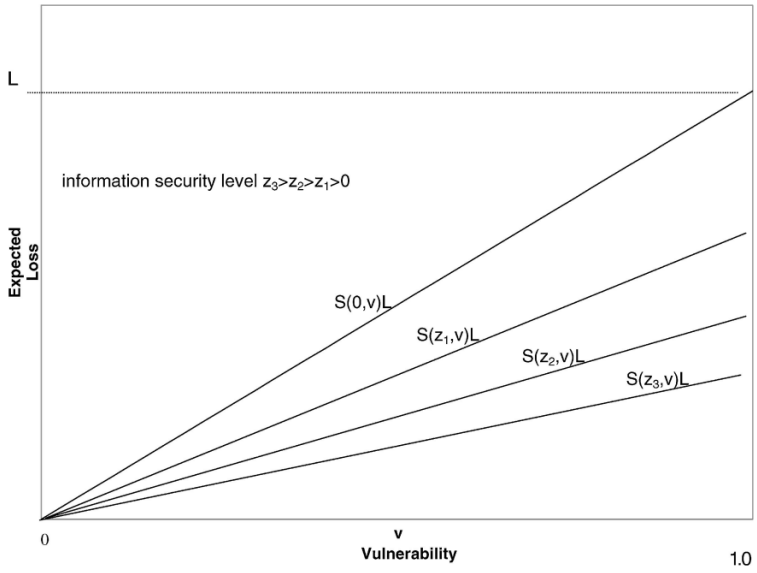


Fig. 2. Expected value of information loss, $S(z, v)\,L$, as vulnerability increases at different levels of investment in information security (for Class I).

Figure 2: Class I: Expected Loss as a function of vulnerability v. Higher investments (higher levels of z) produce less expected loss.

Since we now have a definition of $S(z, v)$, we can compute the derivative and actually find the optimal security investment, which in this case is

$$z^{I*}(v) = \frac{(v\beta\alpha)^{1/(\beta+1)} - 1}{\alpha}$$

which I will not derive here and leave as an exercise for the student (it might be in the appendix).

Plotting this as a function of v gives the following:

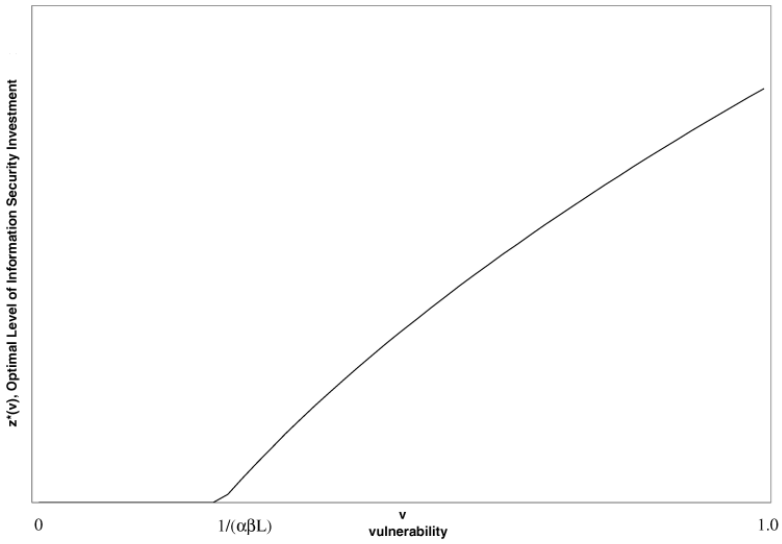


Fig. 3. Optimal value of security investments as a function of vulnerability, $z^*(v)$ for Class I.

What's a "plain English" way of interpreting this chart? If we assume that our function S is of this Class I form then for low values of vulnerability, it may make sense not to invest in security at all! This is somewhat surprising. After the point $1/(\alpha\beta L)$, the optimal investment is monotonically increasing.

If we assume that L is very high, where does that shift this breakeven point? It moves the curve to the left, i.e. this "grace zone" shrinks.

Likewise, what happens if we L is very low, meaning that even if a loss happened it wouldn't be very bad? It shifts this curve rightward. Depending on just how far, it might mean that $1/(\alpha\beta L) > 1$, meaning that the optimal investment is 0 regardless of v!. This requires a low L, a low β, and a very low α.

Anyone remember what α is referred to here? It's a measure of the effectiveness of your investment (same with β). So if you don't expect a big loss, and the investment isn't going to be very effective, you might be better off not investing in security as well.

2.3.2 Candidate 2 for $S(z, v)$

$$S^{II}(z, v) = v^{\alpha z+1}$$

where $\alpha > 0$. The graph looks like this:

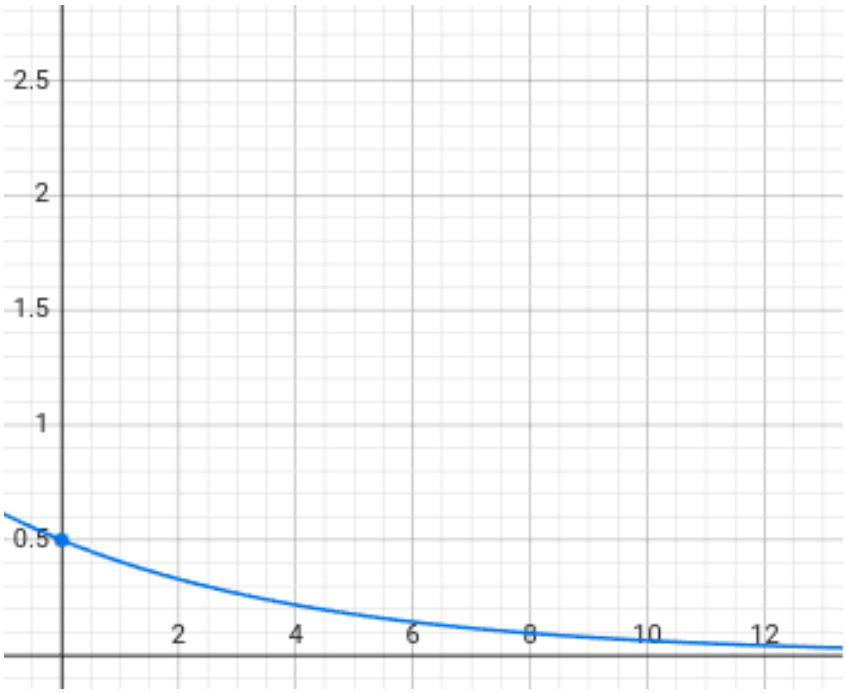


Figure 3: Class II candidate function for $S(z, v)$ (with z along the x-axis). Note that $0 \leq v \leq 1$.

Check: *Does this satisfy the above three assumptions?* Yes.

Looks like this for various values of z:

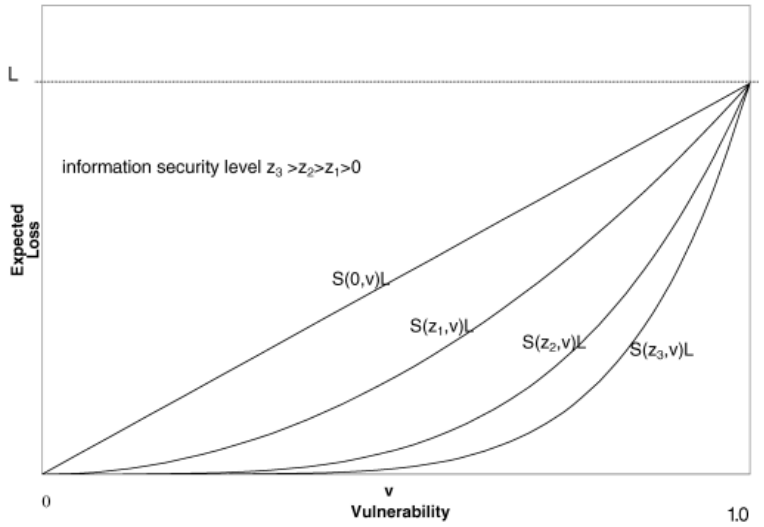


Fig. 4. Expected value of information loss, $S(z, v)\,L$, as vulnerability increases at different levels of investment in information security (for Class II).

Then analysis reveals that there is an optimum investment level at

$$z^{II*}(v) = \frac{\ln(1/-\alpha v L(\ln v))}{a \ln v}$$

Plotting this as a function of v gives the following:

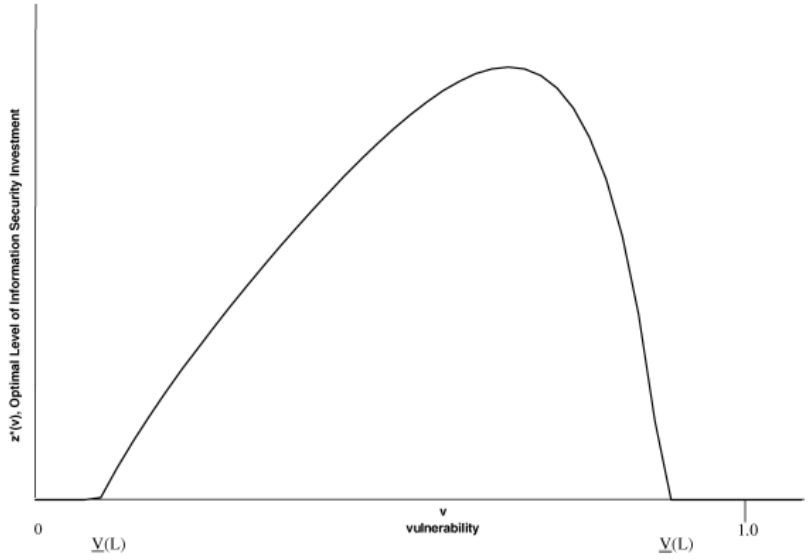


Fig. 5. Optimal value of security investments as a function of vulnerability, $z^*(v)$ for Class II.

What's the "plain English" interpretation of this chart? Recall that we are assuming that v is not really a controllable parameter. The interesting takeaway is that for highly vulnerable datasets, there actually is a point at which it makes sense to *not invest at all!* I.e. it's so vulnerable that you're better off just accepting as a loss because it would be too expensive to try to secure.

What's the implication? According to the authors, this concept (which they call Proposition 2 in the paper) means that firms should be careful to concentrate security resources

Proposition 3 states that if we assumes the security breach probability function S belongs to either Class I or Class II. Then it is determined (in the appendix) that optimal investment is $z^*(v) < (1/e)vL$. In other words, the optimal investment is less than $1/e$ of the expected loss vL. *Anyone know what* $1/e$ *is, or remember it from the paper?* It's about 0.3679, or 36.79%. So if you expect to lose \$100 with no security investment, and assume Class I or Class II for S, then you should under no circumstances ever invest more than \$36.79 into security. Does anyone else find this surprising?

2.4 Concluding discussion

Questions:

- Why might we choose S^I over S^{II}? What is the "plain English" difference between them?
- How practically useful is this? Do you think that firms are using this model to determine how much to invest in cybersecurity?
  - I don't think I've ever heard of anyone using this in the real world...
  - Yet this paper has almost 2000 citations. What does that mean?
  - Why? Is the model too abstract or too simple to accurately capture the nature of information security? Or is it too hard to come up with good estimates of values like t and λ?in real world situations? Is it something else?
- Despite all this, the thing I like about this paper is that it is able to make high-level impactful statements, like e.g. for S^I "a firm can be better off concentrating its resources on high-vulnerability information sets".
- Perhaps just describing the shape of functions and reaching conclusions like "an optimum exists" is useful.
  - Maybe just being able to list out assets in terms of high, medium, and low risk can—when combined with the model—be useful in deciding where to make security investments.
- I think the authors list out these limitations well in the paper.