

Basic Facts about Risk Aversion

Question: Consider an individual whose expresses the risk preferences

$[\$2000] \sim 0.5[\$5000]+0.5[\$0]$ and $[\$1000] \sim 0.27[\$5000]+0.73[\$0]$.

If this individual is rational (consistent), which should he prefer among the lotteries

$0.5[\$2000]+0.5[\$1000]$ and $0.4[\$5000]+0.6[\$0]$?

Consider an individual with a twice-differentiable utility function $u(\bullet)$.

Suppose this individual has wealth x plus a gamble that will pay a small random amount $\tilde{\epsilon}$, such that $E(\tilde{\epsilon}) = 0$.

Let δ be the maximum that he would pay to insure against this gamble. So $u(x-\delta) = E(u(x+\tilde{\epsilon}))$.

Assuming that all possible values of $\tilde{\epsilon}$ are near 0, Taylor series approximations yield

$u(x) - \delta u'(x) \approx E[u(x) + u'(x)\tilde{\epsilon} + u''(x)\tilde{\epsilon}^2/2] = u(x) + u''(x) \text{Var}(\tilde{\epsilon})/2$.

and then $\delta \approx -[u''(x)/u'(x)] \text{Var}(\tilde{\epsilon})/2$.

The individual's Arrow-Pratt risk-aversion index at wealth x is $r(x) = -u''(x)/u'(x)$.

So we find that the individual's value for a small zero-expected-value gamble is approximately half of the variance of the gamble times the individual's risk-aversion index.

Notice that this risk-aversion index $r(x)$ would not change if we changed how i 's utility is measured to some other equivalent scale $\hat{u}(x) = Au(x)+B$ where $A>0$ and B are constants.

The reciprocal of the risk aversion index $\tau(x) = 1/r(x)$ is called the risk tolerance index, which has the advantage of being measured in the same units as money (dollars).

An individual whose risk aversion is constant, independent of his given wealth x , must have a utility function that satisfies the differential equation $-u''(x)/u'(x) = R$, for some constant R .

This differential equation is equivalent to $d[\text{LN}(u'(x))]/dx = -R$.

For $R>0$, the solutions to this differential equation with $u'>0$ are $u(x) = B - Ae^{-Rx}$,

where $A>0$ and B are arbitrary scale constants.

Suppose an individual with constant risk aversion R ($u(x) = -e^{-Rx}$) will get a random income \tilde{Y} drawn from a Normal distribution with mean μ and standard deviation σ .

The certainty equivalent (CE) of this lottery \tilde{Y} for this individual is the sure amount of money W that individual would be willing to accept instead of this lottery \tilde{Y} .

Thus, the certainty equivalent W satisfies the equation $u(W) = E(u(\tilde{Y}))$.

Fact: Suppose $u(\bullet)$ is a utility function with constant risk aversion, \tilde{Y}_1 and \tilde{Y}_2 are independent random variables, and $u(W_i) = E(u(\tilde{Y}_i))$ for each i in $\{1,2\}$. Then $u(W_1+W_2) = E(u(\tilde{Y}_1+\tilde{Y}_2))$.

Fact: For an individual with constant risk aversion R , a Normal lottery with mean μ and standard deviation σ has certainty equivalent $W = \mu - 0.5R\sigma^2$.

That is, $-e^{-R(\mu-0.5R\sigma^2)} = \int_{-\infty}^{+\infty} (-e^{-Ry}) \frac{e^{-0.5((y-\mu)/\sigma)^2}}{\sqrt{2\pi}\sigma} dy$. (Proof: use $\int_{-\infty}^{+\infty} \frac{e^{-0.5((y-[\mu-R\sigma^2])/ \sigma)^2}}{\sqrt{2\pi}\sigma} dy = 1$.)

Efficient Risk Sharing in a Syndicate

Question: Individuals 1 and 2 have constant risk tolerances $T_1 = \$20,000$ and $T_2 = \$30,000$.

1 has an investment paying returns drawn from a Normal distribution: $\mu = \$35,000$, $\sigma = \$25,000$.

1 can sell any share θ of this investment to 2 for any price up to her CE of this share.

To maximize 1's total CE, what share θ should 1 offer to sell, for what price? (60% for \$17,250)

Let N denote the set of member of a syndicate or investment partnership.

They hold assets which will yield returns that are a random variable \tilde{Y} with some given probability distribution.

Each individual i in N has a given utility function for money $u_i(\bullet)$.

Let $x_i(y)$ denote the individual i 's planned payoff when the syndicate earns $\tilde{Y}=y$.

To be feasible, we must have $\sum_{i \in N} x_i(y) = y$, $\forall y \in \mathbb{R}$.

Let us consider an efficient allocation rule $(x_i(\bullet))_{i \in N}$ that maximizes $\sum_{i \in N} \lambda_i E(u_i(x_i(\tilde{Y})))$, subject to this feasibility constraint, for some given positive utility weights $(\lambda_i)_{i \in N}$.

To simplify the characterization of the efficient rule, we assume all appropriate differentiability.

For each outcome y , the marginal weighted utility $\lambda_i u_i'(x_i(y))$ must be equal across all i in N .

That is, there must exist some function $v(y)$ such that $\lambda_i u_i'(x_i(y)) = v(y)$, $\forall i \in N$, $\forall y \in \mathbb{R}$ (1).

Differentiating with respect to y , we get $\lambda_i u_i''(x_i(y)) \partial x_i / \partial y = v'(y)$ (2).

Equation (1) implies $\lambda_i = v(y) / u_i'(x_i(y))$.

Substituting this into (2), we get $[u_i''(x_i(y)) / u_i'(x_i(y))] \partial x_i / \partial y = v'(y) / v(y)$.

The bracketted term here is just the Arrow-Pratt index of risk aversion times -1 .

Let $\tau_i(x_i)$ denote the reciprocal of the Arrow-Pratt risk-aversion index, which we may call the risk tolerance of individual i when i gets payoff $x_i(y)$. That is $\tau_i(x_i) = -u_i'(x_i) / u_i''(x_i)$.

Then we get $\partial x_i / \partial y = -(v'(y) / v(y)) \tau_i(x_i)$.

This derivative $\partial x_i / \partial y$ may be called i 's share of the variable risks held by the syndicate.

Feasibility implies $1 = \sum_{j \in N} \partial x_j / \partial y = -(v'(y) / v(y)) \sum_{j \in N} \tau_j(x_j)$.

Thus we get $\partial x_i / \partial y = \tau_i(x_i(y)) / [\sum_{j \in N} \tau_j(x_j(y))]$. (3)

That is, individual i 's share of the syndicate's risks should be proportional to i 's risk tolerance.

The special case of constant absolute risk aversion:

Suppose now that each individual i has the utility function $u_i(w) = -e^{-w/T(i)}$,

for some given parameter $T(i)$. Then $u_i'(w) = (e^{-w/T(i)}) / T(i)$, $u_i''(w) = -(e^{-w/T(i)}) / (T(i))^2$.

So the parameter $T(i) = -u_i'(w) / u_i''(w)$ is i 's constant risk tolerance at all payoff levels.

Then $\partial x_i / \partial y = T(i) / \sum_{j \in N} T(j)$, $\forall y$,

and so $x_i(y) = x_i(0) + y T(i) / \sum_{j \in N} T(j)$.

That is, individuals who have constant risk tolerance should linearly share all risks in proportion to their risk tolerances.

References: Robert Wilson, "The theory of syndicates," *Econometrica* 36(1):119-132 (1968).

Karl Borch, "Equilibrium in a reinsurance market," *Econometrica* 30(3):424-444 (1962).

Optimal risk sharing among partners with constant risk tolerance

Consider a group of individuals who have formed a partnership to share the risky profits from some joint venture or gamble. Suppose each individual j in this group has a constant risk tolerance, denoted by τ_j . Let T denote the sum of all the partners' risk tolerances ($T = \sum_j \tau_j$).

Fact 1. (Optimal sharing rule) These partners can maximize the sum of their certainty equivalents by sharing the risky profits among themselves in proportion to their risk tolerances, with each individual j taking the fractional share τ_j/T of the risky profits.

Fact 2. (Corporate risk tolerance) The maximal sum of the partners' certainty equivalents that can be achieved by optimal risk sharing (as described in Fact 1) is equal to the certainty equivalent of the whole gamble to an individual who has a constant risk tolerance equal to T , the sum of these partners' risk tolerances. Thus, to maximize the sum of their certainty equivalents, the partnership should evaluate gambles according to its total risk tolerance T , whenever the partners have a choice about which gambles to undertake. That is, in making decisions about risky investments, the partnership should act as a corporate person with a risk tolerance equal to the sum of its members' risk tolerances.

Fact 3. (Partners' unanimity) Suppose that the partners are planning to share all risky income in the proportions described in Fact 1. Then any decision by the partnership to change their investments (that is, to change the probability distribution of their total income) will be unanimous, because any change in that increases the sum of the partners' certainty equivalents will also increase each individual's certainty equivalent of his proportional share.

Fact 4. (Coase theorem) If the partners were planning to share risks according to a sharing rule that does not maximize the sum of the partners' certainty equivalents, then any partner j could propose another sharing rule that would increase j 's own certainty equivalent and would not decrease the certainty equivalents of any other partners.

Risk-sharing example. Consider a financial asset yielding a return \tilde{Y} that is drawn from a Normal distribution with mean $\mu = \$35000$, $\sigma = \$25000$. Investor 1 has risk tolerance $T_1 = 20000$, and investor 2 has risk tolerance $T_2 = \$30000$.

When 1 owns the whole asset, his certainty equivalent is

$$CE_1(\tilde{Y}) = 35000 - (0.5/20000) * 25000^2 = 35000 - 15625 = \$19375.$$

If 2 owned the whole asset, her certainty equivalent would be

$$CE_2(\tilde{Y}) = 35000 - (0.5/30000) * 25000^2 = 35000 - 10417 = \$24583$$

So if 1 sold the asset to 2 for price x then they would get

$$CE_1(x) = x, \quad CE_2(\tilde{Y} - x) = \$24583 - x$$

The transaction makes both better off if $19375 \leq x \leq 24583$.

In this range, $x=24583$ would be best for investor 1.

Now consider the possibility that 1 could sell a 50% share to 2 for some price x . Then they get

$$\begin{aligned} CE_1(0.5\tilde{Y}+x) &= 0.5*35000 + x - (0.5/20000)*(0.5*25000)^2 = \\ &= 17500 + x - 3906 = \$13594 + x, \end{aligned}$$

$$\begin{aligned} CE_2(0.5\tilde{Y}-x) &= 0.5*35000 - (0.5/30000)*(0.5*25000)^2 = \\ &= 17500 - x - 2604 = \$14896 - x. \end{aligned}$$

This transaction makes both better off if

$$19375 - 13594 = 5781 \leq x \leq 14896.$$

The sum of their certainty equivalents is $13594 + 14896 = \$28490$

which is higher than either could get from the asset alone.

When the price x is 14896, investor 1 gets all of this certainty equivalent value.

In general, if 1 sells a share θ to investor 2 for a price x , their certainty equivalents are

$$CE_1((1-\theta)\tilde{Y}+x) = (1-\theta)*35000 + x - (0.5/20000)*[(1-\theta)*25000]^2$$

$$CE_2(\theta\tilde{Y}-x) = \theta*35000 - x - (0.5/30000)*(\theta*25000)^2$$

The sum of these is maximal when $(1-\theta)^2/T_1 + \theta^2/T_2$ is minimized, at $\theta = T_2/(T_1+T_2) = 0.6$.

When 1 sells the optimal 60% share to investor 2 for a price x , they get

$$\begin{aligned} CE_1(0.40*\tilde{Y}+x) &= 0.40*35000 + x - (0.5/20000)*(0.40*25000)^2 = \\ &= 14000 + x - 2500 = \$11500 + x, \end{aligned}$$

$$\begin{aligned} CE_2(0.60*\tilde{Y}-x) &= 0.60*35000 - (0.5/30000)*(0.60*25000)^2 = \\ &= 21000 - x - 3750 = \$17250 - x. \end{aligned}$$

This transaction makes both better off if $19375 - 11500 = 7875 \leq x \leq 17250$.

The maximized sum of their certainty equivalents is $11500 + 17250 = \$28750$.

When the price x is 17250, investor 1 gets this maximal certainty-equivalent value.

For a fictitious corporate person with risk tolerance $T_0 = T_1 + T_2 = 50000$, the asset \tilde{Y} would be worth $CE_0(\tilde{Y}) = 35000 - (0.5/50000)*25000^2 = 35000 - 6250 = \28750 .

Moral Hazard with constant risk tolerance, monetary effort cost, 2 actions

A risk-neutral principal is designing an incentive plan for a risk-averse agent.

The agent must choose among two unobservable actions: a_H and a_L .

Each action has a cost to the agent, $c(a_L) = c_L < c(a_H) = c_H$. The agent could earn \bar{w} elsewhere.

Suppose the agent's utility from choosing action a and getting wage w would be $u(w-c(a)) = -\text{EXP}(-(w-c(a))/T)$, where $T>0$ is the agent's constant risk tolerance.

[This model differs from standard textbook models in that effort cost here is monetary, subtracted from income before applying the utility function: $u(w-c(a))$ instead of $u(w)-c(a)$.]

The principal can only observe an outcome \tilde{y} that depends on the agent's action according to the conditional probability distribution $p(y|a_H)$ or $p(y|a_L)$.

Let Y denote the set of possible values of y , which we assume here to be a finite set.

The principal can promise the agent a wage $w(y)$ that depends on the observable outcome.

Suppose that the principal's expected payoff is much higher when the agent chooses a_H .

So the principal's problem is to design the wage-function $w(\bullet)$ to minimize the expected wage expense, subject to the constraints that the agent should not prefer the outside option \bar{w} or the lower action a_L :

$$\begin{aligned} &\text{Choose } (w(y))_{y \in Y} \text{ to minimize } \sum_{y \in Y} p(y|a_H) w(y) \text{ subject to} \\ &\sum_{y \in Y} p(y|a_H) u(w(y)-c(a_H)) \geq u(\bar{w}) \quad (\text{participation constraint: } \lambda), \\ &\sum_{y \in Y} p(y|a_H) u(w(y)-c(a_H)) \geq \sum_{y \in Y} p(y|a_L) u(w(y)-c(a_L)) \quad (\text{moral-hazard constraint: } \mu). \end{aligned}$$

The participation constraint must be binding, or else the principal could reduce all $w(y)$.

If the moral-hazard constraint were not binding, then the optimal solution would be $w(y)=\bar{w}+c_H$ for all outcomes y , but then the agent would prefer the lower-cost action a_L (with $c(a_L)<c(a_H)$).

So both constraints must be binding at the optimal solution. Let λ and μ denote the Lagrange multipliers of the participation and moral-hazard constraints respectively. The Lagrangean is

$$\begin{aligned} \mathcal{L}(w;\lambda,\mu) = &\sum_y p(y|a_H) w(y) - \lambda[\sum_y p(y|a_H) u(w(y)-c_H) - u(\bar{w})] \\ &- \mu[\sum_y p(y|a_H) u(w(y)-c_H) - \sum_y p(y|a_L) u(w(y)-c_L)]. \end{aligned}$$

The optimality conditions $0 = \partial \mathcal{L} / \partial w(y)$ yield

$$\forall y \in Y: \quad 0 = p(y|a_H) - (\lambda + \mu) p(y|a_H) u'(w(y)-c_H) + \mu p(y|a_L) u'(w(y)-c_L).$$

With constant risk tolerance, $u'(w-c) = \text{EXP}(-(w-c)/T)/T = -u(w-c)/T$.

Thus, summing the optimality conditions over all y in Y , we get

$$0 = 1 + (\lambda + \mu) \sum_y p(y|a_H) u(w(y)-c_H)/T - \mu \sum_y p(y|a_L) u(w(y)-c_L)/T,$$

which with the binding constraints yields $0 = 1 + (\lambda + \mu) u(\bar{w})/T - \mu u(\bar{w})/T = 1 + \lambda u(\bar{w})/T$.

So the participation constraint's Lagrange multiplier is $\lambda = 1/u'(\bar{w}) = -T/u(\bar{w}) = T \times \text{EXP}(\bar{w}/T)$.

With constant risk tolerance T , $u'(w(y)-c_L) = \eta u'(w(y)-c_H)$, where $\eta = \text{EXP}(-(c_H-c_L)/T)$.

So the optimality conditions become $0 = 1 - u'(w(y)-c_H)[\lambda + \mu - \mu \eta p(y|a_L)/p(y|a_H)]$, $\forall y \in Y$.

Then the optimal wage $w(y)$ can be determined from the equations:

$$T \times \text{EXP}((w(y)-c_H)/T) = 1/u'(w(y)-c_H) = \lambda + \mu - \mu \eta p(y|a_L)/p(y|a_H), \quad \forall y \in Y.$$

Thus, the optimal wage $w(y)$ is monotone decreasing in the likelihood ratio $p(y|a_L)/p(y|a_H)$.

With $\mu \leq \lambda/(\eta \max_{y \in Y} p(y|a_L)/p(y|a_H) - 1)$, μ is determined by the requirement that the moral-hazard constraint must be satisfied as a binding equality.

Moral hazard with risk neutrality but limited liability (Tirole's first example)

In a basic moral-hazard model from Tirole's Theory of Corporate Finance (section 3.2), the probability of a project's success depends on the agent's actions. The model's given parameters are $(I, C, R, p_H, p_L, B, \bar{w})$, where I denotes the capital investment cost required for new project, C denotes the agent's collateral assets that he can forfeit if project fails, R denotes the returns from the project if it succeeds, but the returns will be 0 if the project fails, p_H is the probability of success if the agent behaves appropriately, but p_L is the probability of success if the agent misbehaves, and B denotes the value of private benefits that the agent gets by misbehaving.

We assume that $p_L < p_H$, $p_H R > I > p_L R + B$, and $I > C$, so that the project is worthwhile only if agent behaves appropriately, but the agent cannot undertake the project alone. Everyone is risk neutral, but the agent cannot pay more than his collateral C , so the wage cannot be less than $-C$.

We will consider two versions of this model. In version 1, losing C is the worst possible outcome for the agent. So the principal chooses the agent's net wages w_S and w_F , for the cases of success and failure respectively. In version 2, we admit the possibility that, in case of failure, the principal could also subject the agent to punishment that cost the agent $z \geq 0$ with no benefit to the principal. So in the extended version 2, the principal's problem is to choose (w_S, w_F, z) to

$$\begin{aligned} & \text{maximize } \Pi = p_H(R - w_S) + (1 - p_H)(-w_F) - I \text{ subject to} \\ & p_H w_S + (1 - p_H)(w_F - z) \geq \bar{w} \quad (\text{participation constraint; } \lambda), \\ & p_H w_S + (1 - p_H)(w_F - z) \geq B + p_L w_S + (1 - p_L)(w_F - z) \quad (\text{strategic incentive constraint; } \mu), \\ & w_F \geq -C, w_S \geq -C, \text{ and } z \geq 0 \quad (\text{limited liability}). \end{aligned}$$

The principal's problem in version 1 is same except that we drop the variable z (or require $z = 0$).

The principal's expected profit is $\Pi = (p_H R - I) - [p_H w_S + (1 - p_H)w_F] = (p_H R - I) - E\tilde{w}$.

The participation constraint implies $E\tilde{w} = p_H w_S + (1 - p_H)w_F \geq \bar{w} + (1 - p_H)z$.

and the incentive constraint implies $(p_H - p_L)(z + w_S - w_F) \geq B$,

and so $E\tilde{w} = p_H w_S + (1 - p_H)w_F \geq B p_H / (p_H - p_L) + w_F - p_H z$.

The principal wants to minimize his expected wage bill $Ew = p_H w_S + (1 - p_H)w_F$, so it will be $E\tilde{w} = \text{maximum}\{\bar{w} + (1 - p_H)z, B p_H / (p_H - p_L) + w_F - p_H z\}$.

In any case, this formula can be minimized by making w_F as small as possible, so $w_F = -C$.

Now consider version 1, where punishment is not allowed, so $z = 0$.

When $\bar{w} \geq B p_H / (p_H - p_L) - C$, the participation constraint is binding,

and the optimal solution is $w_F = -C$, $w_S = [\bar{w} + (1 - p_H)C] / p_H$, yielding $E\tilde{w} = \bar{w}$.

But when $\bar{w} < B p_H / (p_H - p_L) - C$, the incentive constraint is binding,

and the optimal solution is $w_F = -C$, $w_S = B / (p_H - p_L) - C$, yielding $E\tilde{w} = B p_H / (p_H - p_L) - C$.

Now consider version 2, where punishment $z \geq 0$ is allowed.

When $\bar{w} \geq B p_H / (p_H - p_L) - C$, the participation constraint is binding,

and the optimal solution remains $w_F = -C$, $w_S = [\bar{w} + (1 - p_H)C] / p_H$, $z = 0$, yielding $E\tilde{w} = \bar{w}$.

But when $\bar{w} < B p_H / (p_H - p_L) - C$, increasing z above 0 reduces the expected wage bill until we get

$\bar{w} + (1 - p_H)z = B p_H / (p_H - p_L) - C - p_H z$, so that both constraints bind, and the optimal solution is

$w_F = -C$, $z = B p_H / (p_H - p_L) - C - \bar{w}$, $w_S = \bar{w} + (1 - p_H)B / (p_H - p_L)$,

yielding $E\tilde{w} = p_H \bar{w} + B p_H (1 - p_H) / (p_H - p_L) - (1 - p_H)C$.

Lagrangean analysis of the limited-liability basic moral-hazard example.

The Lagrangean can be written: $\mathcal{L}(w; \lambda, \mu) =$

$$= p_H(R - w_S) + (1 - p_H)(-w_F) - I + \lambda[p_H w_S + (1 - p_H)(w_F - z) - \bar{w}] + \mu[(p_H - p_L)(w_S + z - w_F) - B].$$

We know that the minimal value $w_F = -C$ can always be applied, but w_S will be strictly greater than this lower bound in any optimal solution, and so the Lagrangean's derivative with respect to w_S must satisfy $0 = \partial \mathcal{L} / \partial w_S = p_H[-1 + \lambda + \mu(p_H - p_L) / p_H]$.

If the participation constraint binds but the incentive constraint does not, we get $\lambda = 1$ and $\mu = 0$. This happens in our optimal solutions whenever $\bar{w} > B p_H / (p_H - p_L) - C$.

If the incentive constraint binds but the participation constraint does not, then we must have $\lambda = 0$ and $\mu = p_H / (p_H - p_L)$. This happens in our optimal solutions when $\bar{w} < B p_H / (p_H - p_L) - C$ in version 1 (where $z = 0$ is required).

But in version 2, where $z > 0$ is allowed, both constraints are binding when $\bar{w} < B p_H / (p_H - p_L) - C$. Then with both constraints binding and $z > 0$, the Lagrange multipliers λ and μ satisfy

$$0 = \partial \mathcal{L} / \partial w_S = p_H[-1 + \lambda + \mu(p_H - p_L) / p_H],$$

$$0 = \partial \mathcal{L} / \partial z = -\lambda(1 - p_H) + \mu(p_H - p_L),$$

and so we get $\lambda = p_H$ and $\mu = p_H(1 - p_H) / (p_H - p_L)$.

With w_F at its lower bound, our optimal solutions in all cases also satisfy

$$0 \geq \partial \mathcal{L} / \partial w_F = (1 - p_H)(-1 + \lambda) - \mu(p_H - p_L) = \lambda - 1.$$

[To get the second equality here, notice that $0 = \partial \mathcal{L} / \partial w_S$ implies $(1 - \lambda)p_H = \mu(p_H - p_L)$.]

To see how small changes in the parameters \bar{w} or B would affect the optimal value of the principal's objective, we differentiate the Lagrangean with respect to these parameters:

$$\partial \mathcal{L} / \partial \bar{w} = -\lambda, \quad \partial \mathcal{L} / \partial B = -\mu.$$

Similarly, to see how small changes in C would affect the optimal value of the principal's objective, we can use the fact that $w_F = -C$ in our optimal solution, which implies that

$$\partial \mathcal{L} / \partial C = -\partial \mathcal{L} / \partial w_F = 1 - \lambda.$$

To the **efficiency wage** model: Now consider a special case of this example where a project is managed over a short time period ε . What we called "failure" now can be interpreted as an accident occurring in this short time period, and success is not having an accident. Let us ignore the parameters R and I and instead simply assume that it will always be optimal operate the project and make sure that the agent chooses to behave well (as I is modest but R is large).

Suppose the agent has no collateral, and he can always run away to get the expected value $\bar{v} > 0$.

So punishments that would make him worse off are impossible, and we have $C = -\bar{v}$.

Over the short time period, the probability of an accident is small in any case, but it is smaller if the agent behaves well. So suppose that $p_H = 1 - \alpha\varepsilon$ and $p_L = 1 - \beta\varepsilon$, where $\beta > \alpha > 0$.

Hidden benefits of misbehavior are also small over the short time period, so suppose that $B = D\varepsilon$.

To hire the the agent he must expect some \bar{w} satisfying $\bar{w} \geq \bar{v}$.

Then the expected wage is minimized by paying $w_F = \bar{v}$ in case of an accident, otherwise paying $w_S = \max\{(\bar{w} - \alpha\varepsilon\bar{v}) / (1 - \alpha\varepsilon), \bar{v} + (D\varepsilon) / [(1 - \alpha\varepsilon) - (1 - \beta\varepsilon)]\} = \max\{(\bar{w} - \alpha\varepsilon\bar{v}) / (1 - \alpha\varepsilon), \bar{v} + D / (\beta - \alpha)\}$.

To hire a new agent from outside, we must promise $\bar{w} = \bar{v}$, and then we get $w_S = \bar{v} + D / (\beta - \alpha)$, $E\bar{w} - \bar{v} = (1 - \alpha\varepsilon)D / (\beta - \alpha) \rightarrow D / (\beta - \alpha)$ as $\varepsilon \rightarrow \infty$. Changing agents costs $D / (\beta - \alpha)$ every time!

Becker-Stigler-Shapiro-Stiglitz efficiency wages

Consider a similar moral-hazard problem where the agent is risk neutral, and there are only two possible observations: $y=1$ denotes normal business, and $y=0$ denotes an accident occurring.

We consider a short interval of time ε , in which the probability of an accident is $\alpha\varepsilon$ if the agent chooses to be diligent a_H , and $\beta\varepsilon$ if the agent chooses to shirk a_L , where $\beta > \alpha$.

Choosing a_L also yields a hidden benefit of $D\varepsilon$ to the agent in this period.

Suppose that participation constraints apply ex-post: after the outcome is observed, the agent cannot be made worse off than his outside option of \bar{v} , which is the present-discounted value of his lifetime income in the competitive labor market. The principal wants to minimize the expected cost subject to the ex-post participation constraints and the moral-hazard incentive constraint that the agent should not shirk. Let V_1 and V_0 denote the agent's expected total payoff after observing $y=1$ or $y=0$ respectively. So the principal's problem is

$$\begin{aligned} &\text{choose } (V_1, V_0) \text{ to minimize } (1-\alpha\varepsilon)V_1 + \alpha\varepsilon V_0 \text{ subject to} \\ &(1-\alpha\varepsilon)V_1 + \alpha\varepsilon V_0 \geq D\varepsilon + (1-\beta\varepsilon)V_1 + \beta\varepsilon V_0, \quad V_1 \geq \bar{v}, \quad V_0 \geq \bar{v}. \end{aligned}$$

We could add an ex-ante participation constraint $(1-\alpha\varepsilon)V_1 + \alpha\varepsilon V_0 \geq \bar{w}$, but if the payoff \bar{w} that must be promised to recruit the agent is the same as the payoff \bar{v} that he can get by quitting later, then this constraint is redundant with the ex-post participation constraints $V_1 \geq \bar{v}$ and $V_0 \geq \bar{v}$.

The moral-hazard constraint implies $V_1 - V_0 \geq D/(\beta-\alpha)$.

So the optimal solution is $V_0 = \bar{v}$, $V_1 = \bar{v} + D/(\beta-\alpha)$.

So the agent's expected reward $(1-\alpha\varepsilon)V_1 + \alpha\varepsilon V_0 = \bar{v} + (1-\alpha\varepsilon)D/(\beta-\alpha)$ for a short ε -period of service must be greater than the outside option \bar{v} by a positive bonus even as $\varepsilon \rightarrow 0$.

In a dynamic model, if the problem is repeated with a different agent every ε -period then the principal's cost becomes huge, but the cost can be reduced by using benefits of future employment be part of current incentive-pay.

Consider a stationary solution: if $y=1$ this period then the agent will be paid εw and rehired for next period, but if the $y=0$ this period then the agent is dismissed to the outside option \bar{v} .

So with discount rate r , we get the recursion equation: $V_1 = w\varepsilon + (1-r\varepsilon)((1-\alpha\varepsilon)V_1 + \alpha\varepsilon\bar{v})$.

With $V_1 = \bar{v} + D/(\beta-\alpha)$, this implies $w = r\bar{v} + (r + \alpha - \varepsilon\alpha)D/(\beta-\alpha)$.

Here $r\bar{v}$ is the outside wage rate corresponding to the present-discounted value \bar{v} .

As $\varepsilon \rightarrow 0$, this optimal stationary incentive plan pays the agent an efficiency wage rate that exceeds the outside wage rate by $(r+\alpha)D/(\beta-\alpha)$, but to dismiss the agent when an accident occurs.

This $\varepsilon \rightarrow 0$ is a problem of controlling a Poisson process where accidents occur as a Poisson process with the low rate α when the agent is diligent, but the high rate β when the agent shirks.

When \tilde{X} is a Poisson random variable with mean λ , \tilde{X} can be any nonnegative integer, $P(\tilde{X}=k) = e^{-\lambda} (\lambda)^k / k!$ for any $k \in \{0,1,2,\dots\}$, $E(\tilde{X}) = \lambda$, $\text{Var}(\tilde{X}) = \lambda$, $\text{Stdev}(\tilde{X}) = \sqrt{\lambda}$.

When accidents occur in a Poisson process with rate λ , the number of accidents between any two times t and $t+\delta$ ($\delta > 0$) is a Poisson random variable with mean $\lambda\delta$, and it is independent of the number of accidents before time t .

In any short time interval of length ε , the probability of an accident is approximately $\lambda\varepsilon$, the probability of no accidents is approximately $1 - \lambda\varepsilon$, and the probability of two or more accidents is vanishingly smaller of order ε^2 .

Controlling a Poisson processes with constant risk tolerance, no liability limits

Suppose that an agent with constant risk tolerance T is to be paid $\beta\tilde{X}$, where \tilde{X} is a Poisson random variable with some mean λ .

Then the agent's expected utility is

$$EU = \sum_{k=0}^{\infty} \frac{e^{-\lambda} \lambda^k}{k!} (-e^{-k\beta/T}) = -e^{-\lambda} \sum_{k=0}^{\infty} \frac{(\lambda e^{-\beta/T})^k}{k!} = -e^{-\lambda T(1-e^{-\beta/T})/T}$$

and so the agent's certainty equivalent of this income $\beta\tilde{X}$ is $\lambda T(1-e^{-\beta/T})$.

Now suppose that the agent chooses an action $a \in A$ (where A is his set of feasible actions), and this action affects the means $\Lambda_1(a), \dots, \Lambda_n(a)$ of n independent Poisson random variables $\tilde{X}_1, \dots, \tilde{X}_n$. The agent also pays a personal cost $C(a)$ for his action, and has constant risk tolerance T .

The action and cost $C(a)$ cannot be observed by anybody except the agent, but the principal can observe the Poisson random variables $\tilde{X}_1, \dots, \tilde{X}_n$.

A risk neutral principal gets revenue $\tilde{X}_1\pi_1 + \dots + \tilde{X}_n\pi_n$ from these Poisson random variables.

Consider incentive plans, where the principal pays the agent a wage w according to a linear formula $w(\tilde{X}_1, \dots, \tilde{X}_n) = \alpha + \tilde{X}_1\beta_1 + \dots + \tilde{X}_n\beta_n$.

With constant risk tolerance, the agent's certainty equivalent for the sum of n independent random payments is equal to the sum of the certainty equivalents of these n random payments (this additivity condition only holds with constant risk tolerance).

Let \bar{w} denote the best alternative wage that the agent could earn elsewhere.

So the principal's problem can be written: choose $(\alpha, \beta_1, \dots, \beta_n, a^*)$ to maximize $\Lambda_1(a^*)\pi_1 + \dots + \Lambda_n(a^*)\pi_n - [\alpha + \Lambda_1(a^*)\beta_1 + \dots + \Lambda_n(a^*)\beta_n]$ subject to $\max_{a \in A} \alpha + \Lambda_1(a)T(1-e^{-\beta_1/T}) + \dots + \Lambda_n(a)T(1-e^{-\beta_n/T}) - C(a) = \alpha + \Lambda_1(a^*)T(1-e^{-\beta_1/T}) + \dots + \Lambda_n(a^*)T(1-e^{-\beta_n/T}) - C(a^*) \geq \bar{w}$.

Substituting $\alpha = \bar{w} + C(a^*) - \sum_i \Lambda_i(a^*)T(1-e^{-\beta_i/T})$ from the participation constraint, the principal's expected profit becomes $\sum_i \Lambda_i(a^*)[\pi_i - \beta_i + T(1-e^{-\beta_i/T})] - \bar{w} - C(a^*)$.

If we can differentiate with respect to the action a , first-order conditions of the incentive constraint are $\sum_i \Lambda_i'(a^*)T(1-e^{-\beta_i/T}) = C'(a^*)$.

Extending the Poisson process over time, the number of i -arrivals in a time period of length δ may be denoted by $\tilde{X}_i(\delta)$, and it is a Poisson random variable with mean $\delta\Lambda_i(a)$ (=variance).

Results of separate time intervals are independent. With constant risk tolerance, they can be analyzed independently, so the optimal wage is linear in $\tilde{X}_1(\delta), \dots, \tilde{X}_n(\delta)$ over the whole period.

Now fix $\mu_i(a)$ and σ^2 , and consider small positive numbers ϵ .

Given any $\epsilon > 0$, let $\Lambda_i(a) = [\sigma^2 + \epsilon\mu_i(a)]/\epsilon^2$, and let $\tilde{Y}_i(\delta) = \epsilon\tilde{X}_i(\delta) - \delta\sigma^2/\epsilon$.

Then $E(\tilde{Y}_i(\delta)) = \delta\mu_i(a)$, and $\text{Var}(\tilde{Y}_i(\delta)) = \epsilon^2\text{Var}(\tilde{X}_i(\delta)) = \delta[\sigma^2 + \epsilon\mu_i(a)] \rightarrow \delta\sigma^2$ as $\epsilon \rightarrow 0$.

$\Lambda_i(a)$ here becomes large as $\epsilon \rightarrow 0$, and Poisson random variables with large means are approximately Normal.

So as $\epsilon \rightarrow 0$, $\tilde{Y}_i(\delta)$ approaches a Normal random variable,

and \tilde{Y}_i becomes a Brownian-motion process with drift $\mu_i(a)$ and volatility σ .

Wages that are linear in $\tilde{Y}_i(\delta)$, say $A\delta + \sum_{i=1}^n B_i\tilde{Y}_i(\delta)$, are also linear in $\tilde{X}_i(\delta)$, with coefficients:

$$A\delta + \sum_{i=1}^n B_i\tilde{Y}_i(\delta) = A\delta + \sum_{i=1}^n B_i[\epsilon\tilde{X}_i(\delta) - \delta\sigma^2/\epsilon] = (A\delta - n\delta\sigma^2/\epsilon) + \sum_{i=1}^n (B_i\epsilon)\tilde{X}_i(\delta).$$

Holmstrom-Milgrom Control of Brownian Motion (Econometrica 55(2):303-328, 1987):

Suppose the agent is risk-averse with constant risk tolerance, subject to an ex-ante participation constraint as in our moral-hazard Example 1.

But now suppose that the observable is a Normal random variable with mean $\mu(a)$ that depends on his action, and variance σ^2 that does not depend on his action.

For two Normals with different means and the same variance, the likelihood ratios go to infinity in the tails, and so the multiplier of our moral-hazard constraint must be $\mu=0$: it is costless!

The optimal solution achieves first-best nonlinearly: pay \bar{w} except for an infinite punishment in an event that has infinitesimal probability, infinitesimally smaller when he chooses high μ .

Something seems wrong in this above solution.

Holmstrom-Milgrom changed the problem to allow the agent to get feedback as the Normal outcome evolves, by considering the problem of controlling the drift of a Brownian motion.

For such a problem, they showed that linear incentive pay becomes optimal.

For important applications, they explicitly considered multidimensional Brownian motion.

We may (w.l.o.g.) restrict attention to Brownian motion $(\tilde{Y}_1(t), \dots, \tilde{Y}_n(t))$ where the various components $\tilde{Y}_i(t)$ are independent and have the same volatility:

In any time interval from t to $t+\delta$ ($\delta>0$), the changes $\tilde{Y}_i(t+\delta) - \tilde{Y}_i(t)$ are Normal random variables, with mean $\mu_i(a)\delta$ and variance $\sigma^2\delta$, independent across i , and independent of the path to time t .

Here the drifts $\mu_i(a)$ are functions of the agent's hidden action (a) during this interval.

The agent also pays a hidden cost of effort at rate $c(a)$ over time.

The agent's choices of action after time t can depend on the past $(\tilde{Y}_1(t), \dots, \tilde{Y}_n(t))$.

We may assume that the process begins at $(\tilde{Y}_1(0), \dots, \tilde{Y}_n(0)) = (0, \dots, 0)$.

A linear incentive plan for the agent over the period $[0, \Omega]$ specifies constants (A, B_1, \dots, B_n) and promises to pay the agent $w(\tilde{Y}) = A + \sum_i B_i \tilde{Y}_i(\Omega)$.

The agent's final utility payoff is $u(w(\tilde{Y}) - C(a))$, where $C(a) = \int_0^\Omega c(a(t))dt$,

and $u(\bullet)$ is a constant-risk-tolerant utility function. We may let the ending time be $\Omega=1$.

Holmstrom and Milgrom show that such simple linear incentive plans are optimal for the principal to maximize any linear function of the final $(\tilde{Y}_1(\Omega), \dots, \tilde{Y}_n(\Omega))$

subject to an ex-ante participation constraint at time 0 and a moral-hazard constraint that the constant-risk-tolerant agent always chooses an action that maximizes his expected utility:

Choose $w(\bullet)$ to maximize $E(\sum_i \pi_i \tilde{Y}_i(\Omega) - w(\tilde{Y}) | a^*)$ subject to

$a^* \in \operatorname{argmax}_a E(u(w(\tilde{Y}) - C(a)) | a)$, and $E(u(w(\tilde{Y}) - C(a^*)) | a^*) \geq u(\bar{w})$.

Linearity makes $w(\tilde{Y}) = A + \sum_i B_i \tilde{Y}_i(\Omega)$ a Normal random variable, so the agent's constraints can be rewritten with the formula for certainty equivalents of Normals with constant risk tolerance:

$a^* \in \operatorname{argmax}_a A + \sum_i B_i \mu_i(a) - C(a) - (0.5/T)\sigma^2 \sum_i B_i^2, A + \sum_i B_i \mu_i(a^*) - C(a^*) - (0.5/T)\sigma^2 \sum_i B_i^2 \geq \bar{w}$.

The binding participation constraint yields $A = \bar{w} + C(a^*) + (0.5/T)\sigma^2 \sum_i B_i^2 - \sum_i B_i \mu_i(a^*)$.

Then the principal's expected profit is $E(\sum_i \pi_i \tilde{Y}_i(\Omega) - w(\tilde{Y}) | a^*) = \sum_i \pi_i \mu_i(a^*) - \sum_i B_i \mu_i(a^*) - A$

$= \sum_i \pi_i \mu_i(a^*) - C(a^*) - (0.5/T)\sigma^2 \sum_i B_i^2 - \bar{w}$, which we want to maximize

subject to the incentive constraint, which reduces to: $a^* \in \operatorname{argmax}_a \sum_i B_i \mu_i(a) - C(a)$.

Technical key to their proof: Such a Brownian motion can be viewed as the limit of a linear function of multidimensional Poisson process. Large-mean Poissons are approximately Normal. Consider some small $\varepsilon > 0$. For any action a , let the Poisson processes start at $\tilde{x}_i(0|a,\varepsilon) = 0$ for all i . For any time interval t to $t+\delta$, let $\tilde{x}_i(t+\delta|a,\varepsilon) - \tilde{x}_i(t|a,\varepsilon)$ (= arrivals of type i) be a Poisson random variable with mean $\Lambda_i(a,\varepsilon)\delta = (\sigma^2 + \varepsilon\mu_i(a))/\varepsilon^2$, independently of the past before t , and independently across all $i=1,\dots,n$. Then let $\tilde{y}_i(t|a,\varepsilon) = \varepsilon \tilde{x}_i(t|a,\varepsilon) - t\sigma^2/\varepsilon$. So $\tilde{y}_i(t|a,\varepsilon)$ has frequent discontinuous upward jumps of size ε , but between these jumps it is decreasing at the rate $-\sigma^2/\varepsilon$.

Then $E(\tilde{y}_i(t+\delta|a,\varepsilon) - \tilde{y}_i(t|a,\varepsilon)) = \varepsilon \delta (\sigma^2 + \varepsilon\mu_i(a))/\varepsilon^2 - \delta \sigma^2/\varepsilon = \mu_i(a)\delta$,

and $\text{Var}(\tilde{y}_i(t+\delta|a,\varepsilon) - \tilde{y}_i(t|a,\varepsilon)) = \varepsilon^2 \delta (\sigma^2 + \varepsilon\mu_i(a))/\varepsilon^2 = \sigma^2\delta + \varepsilon\mu_i(a)\delta \rightarrow \sigma^2\delta$ as $\varepsilon \rightarrow 0$.

So as $\varepsilon \rightarrow 0$, this \tilde{y} process converges to the Brownian motion \tilde{Y} that we wanted to study.

A wage that is linear in the \tilde{y}_i , say $A + \sum_{i=1}^n B_i \tilde{y}_i$, would also be linear in the \tilde{x}_i , with coefficients: $A + \sum_{i=1}^n B_i \tilde{y}_i(t|a,\varepsilon) = A + \sum_{i=1}^n B_i [\varepsilon \tilde{x}_i(t|a,\varepsilon) - t\sigma^2/\varepsilon] = (A - nt\sigma^2/\varepsilon) + \sum_{i=1}^n (B_i \varepsilon) \tilde{x}_i(t|a,\varepsilon)$

Consider a discrete-time approximation to the multidimensional Poisson process problem ($\varepsilon > 0$).

In any short time interval, there are $n+1$ events that could occur: no arrival, or one arrival of some type $i \in \{1,\dots,n\}$. (Two or more arrivals in a short time interval has vanishingly small probability.)

The optimal incentive plan in this short time interval, given any participation constraint, would pay some amount for each of these possible events.

Changing the participation constraint would only add a constant to all payments, because of constant risk tolerance.

At any point in time, past payments would not affect the agent's preferences over gambles for additional income and effort-cost, again because of constant risk tolerance.

Thus, the optimal incentive plan for the dynamic model can be decomposed to identical incentive problems in each short time interval, plus a constant to meet the ex-ante participation constraint.

But $\tilde{x}_i(\Omega)$ measures the number of times that a type- i arrival occurred, for each of which the agent is paid the same amount.

So the final payment to the agent is a linear function of $(\tilde{x}_1(\Omega|a,\varepsilon), \dots, \tilde{x}_n(\Omega|a,\varepsilon))$,

and so it is also a linear function of $(\tilde{y}_1(\Omega|a,\varepsilon), \dots, \tilde{y}_n(\Omega|a,\varepsilon))$.

Now take the limit as $\varepsilon \rightarrow 0$, and the limit of the optimal plans pays the agent as a linear function of the Brownian-motion endpoint $(\tilde{Y}_1(\Omega), \dots, \tilde{Y}_n(\Omega))$.

First adverse-selection problem: one privately-informed agent with two possible types

The agent could be one of two types, θ_H or θ_L where $\theta_H > \theta_L$. His type is his private information.

The principal thinks the probability of θ_H is p_H , and the probability of θ_L is $p_L = 1 - p_H$.

The agent's effort q and wage w are both observable numbers which may depend on the agent's reported type, but the agent can misrepresent his type. Effort must be nonnegative $q \geq 0$.

When the agent's type is θ , any wage w and effort $q \geq 0$ yield payoff $w - \theta q$ for the agent and yield profit $\pi(q|\theta) - w$ for the principal. The agent's alternative pays \bar{w} for effort 0.

With prime ' for $\partial/\partial q$, suppose $\pi(0|\theta) = 0$, $\pi'(q|\theta) > 0$, $\pi''(q|\theta) \leq 0$, $\forall \theta, \forall q$, and $\pi'(0|\theta_L) > \theta_L$.

The principal's problem is to choose a contract-menu (q_H, w_H, q_L, w_L) to

maximize $p_H(\pi(q_H|\theta_H) - w_H) + p_L(\pi(q_L|\theta_L) - w_L)$ subject to $w_L \in \mathbb{R}$, $w_H \in \mathbb{R}$, $q_L \geq 0$, $q_H \geq 0$,

$$w_L - \theta_L q_L \geq \bar{w}, \quad [\text{L-participation, } \varepsilon]$$

$$w_H - \theta_H q_H \geq \bar{w}, \quad [\text{H-participation, } \lambda]$$

$$w_L - \theta_L q_L \geq w_H - \theta_L q_H, \quad [\text{H|L-incentive, } \mu]$$

$$w_H - \theta_H q_H \geq w_L - \theta_H q_L. \quad [\text{L|H-incentive, } \delta]$$

The Lagrangean can be written:

$$\begin{aligned} \mathcal{L}(w, q; \lambda, \mu, \delta, \varepsilon) &= p_H(\pi(q_H|\theta_H) - w_H) + p_L(\pi(q_L|\theta_L) - w_L) + \varepsilon[w_L - \theta_L q_L - \bar{w}] + \lambda[w_H - \theta_H q_H - \bar{w}] \\ &\quad + \mu[w_L - \theta_L q_L - w_H + \theta_L q_H] + \delta[w_H - \theta_H q_H - w_L + \theta_H q_L] \\ &= p_H \pi(q_H|\theta_H) - q_H [(\lambda + \delta)\theta_H - \mu\theta_L] + p_L \pi(q_L|\theta_L) - q_L [(\varepsilon + \mu)\theta_L - \delta\theta_H] \\ &\quad + w_H [-p_H + \lambda + \delta - \mu] + w_L [-(1 - p_H) + \varepsilon + \mu - \delta]. \end{aligned}$$

The first-order Lagrange optimality conditions for $w_L \in \mathbb{R}$ and $w_H \in \mathbb{R}$ are

$$0 = \partial \mathcal{L} / \partial w_L = -p_L + \varepsilon + \mu - \delta, \text{ and so } \varepsilon + \mu = p_L + \delta;$$

$$0 = \partial \mathcal{L} / \partial w_H = -p_H + \lambda + \delta - \mu, \text{ and so } \lambda + \delta = p_H + \mu.$$

With these equation, the Lagrangean simplifies to

$$\mathcal{L} = p_H \{ \pi(q_H|\theta_H) - q_H [\theta_H + (\theta_H - \theta_L)\mu/p_H] \} + p_L \{ \pi(q_L|\theta_L) - q_L [\theta_L + (\theta_L - \theta_H)\delta/p_L] \}.$$

With $\theta_L < \theta_H$ and $q_H \geq 0$, the L-participation constraint is implied by the H-participation and H|L-incentive constraints. So L-participation is a redundant constraint, and its multiplier is $\varepsilon = 0$.

The incentive constraints (L|H first, then H|L) imply: $\theta_H(q_L - q_H) \geq w_L - w_H \geq \theta_L(q_L - q_H)$.

With $\theta_H > \theta_L$, this implies $q_L - q_H \geq 0$ and $w_L - w_H \geq 0$, so $q_L \geq q_H$ and $w_L \geq w_H$.

[If $q_L > q_H$ then: $w_H - t q_H \geq w_L - t q_L \Leftrightarrow t \geq (w_L - w_H)/(q_L - q_H)$. So any cost type $t > \theta_H$ would prefer (q_H, w_H) over (q_L, w_L) , while any cost type $t < \theta_L$ would prefer (q_L, w_L) .]

Suppose q_H and q_L maximize the Lagrangean. $\pi'(0|\theta_L) > \theta_L \geq \theta_L + (\theta_L - \theta_H)\delta/p_L$, so $q_L > 0$.

With $\varepsilon = 0$, satisfying $0 = \partial \mathcal{L} / \partial w_L$ requires $\mu > 0$, and so the H|L-incentive constraint is binding.

So we have $\mu - \delta = p_L$ and $\lambda = p_H + \mu - \delta = 1$. Now we have two cases (separating and pooling):

In case 1, we have $\delta = 0$. Then we get $\mu = p_L$, $\lambda = 1$, $\pi'(q_L|\theta_L) = \theta_L$,

$\pi'(q_H|\theta_H) \leq \theta_H + (\theta_H - \theta_L)p_L/p_H$ and $q_H \geq 0$ with at least one equality (comp. slackness),

$w_H = \bar{w} + \theta_H q_H$, and $w_L = \theta_L q_L + w_H - \theta_L q_H$ (to make the λ and μ constraints binding).

If these conditions yield $q_H \leq q_L$, then the δ -constraint will be satisfied and this is the solution.

Otherwise, we have case 2 where $\delta > 0$. So both incentive constraints bind, and so $q_H = q_L = q^* > 0$.

The pooling q^* satisfies $\pi'(q^*|\theta_H) = \theta_H + (\theta_H - \theta_L)\mu/p_H$ and $\pi'(q^*|\theta_L) = \theta_L - (\theta_H - \theta_L)\delta/p_L$.

So q^* can be computed from $p_H \pi'(q^*|\theta_H) + p_L \pi'(q^*|\theta_L) = \theta_H$, because $\mu - \delta = p_L = 1 - p_H$.

Then $\delta = p_L [\theta_L - \pi'(q^*|\theta_L)] / (\theta_H - \theta_L) > 0$ requires $\pi'(q^*|\theta_L) < \theta_L$ for this pooling solution.

Motivating an agent with a linear type drawn from a continuous distribution.

Suppose that the agent's type \tilde{t} is a random variable drawn from an interval $[A,B]$.

The agent's type is his cost of effort, and his utility for income w and effort a is $w - t q$.

Consider any contract $(w(\bullet),q(\bullet))$ where the terms of trade for each type t would be $(w(t),q(t))$.

Let $U(w,q|t) = w(t) - t q(t)$ denote the expected utility of type t under this contract.

For any pair of possible types t and s in $[A,B]$, the $(s|t)$ -informational incentive constraint says

$$U(w,q|t) = w(t) - t q(t) \geq w(s) - t q(s) = U(w,q|s) + (s-t)q(s).$$

Similarly, the $(t|s)$ -incentive constraint implies $U(w,q|s) \geq U(w,q|t) + (t-s)q(t)$

So the $(t|s)$ and $(s|t)$ constraints together imply $(s-t)q(t) \geq U(w,q|t) - U(w,q|s) \geq (s-t)q(s)$.

So when $s > t$ we must have $q(t) \geq q(s)$, and so $q(t)$ is a decreasing function of the cost-type t .

Applying these inequalities over many small steps from t up to B , we get the

informational-rent equation: $U(w,q|t) = U(w,q|B) + \int_t^B q(s) ds$.

The expected income of type t is then $w(t) = U(w,q|t) + tq(t) = U(w,q|B) + \int_t^B q(s) ds + tq(t)$.

Suppose the principal's beliefs about the agent's type are described by the cumulative distribution

$F(t) = P(\tilde{t} \leq t)$, and $f(t) = F'(t)$ is the continuous probability density of this distribution, with $f(t) > 0$

for all t in $[A,B]$. Here $F(B)=1$, $F(A)=0$, and $P(a \leq \tilde{t} \leq b) = F(b) - F(a) = \int_a^b f(t) dt$ whenever $a \leq b$.

Then the expected wage bill is

$$\begin{aligned} \int_A^B w(t) f(t) dt &= \int_A^B [U(w,q|B) + \int_t^B q(s) ds + tq(t)] f(t) dt \\ &= U(w,q|B) + \int_A^B \int_t^B q(s) ds f(t) dt + \int_A^B t q(t) f(t) dt \\ &= U(w,q|B) + \int_A^B \int_A^s f(t) dt q(s) ds + \int_A^B s q(s) f(s) ds \\ &= U(w,q|B) + \int_A^B F(s) q(s) ds + \int_A^B s q(s) f(s) ds \\ &= U(w,q|B) + \int_A^B q(t) [t + F(t)/f(t)] f(t) dt. \end{aligned}$$

So the incentive-compatible expected wage $E(w(\tilde{t}))$ looks like what the principal would have to pay without incentive constraints if the cost of each type t were increased to a virtual cost $t + F(t)/f(t)$. This virtual-cost formula expresses the fact that, when we ask more effort from any type t , we increase the amount that we must pay all types below t , because of incentive constraints.

Example: Akerlof's Lemons. Suppose the "agent" is the seller of a unique object, of which the "principal" is the only potential buyer. The seller's type is the value of the object to him, which depends on his unverifiable private information about its quality. Then $q(t)$ can be reinterpreted as the probability of his selling the good if he acts like type t , which must satisfy $0 \leq q(t) \leq 1$, and $w(t)$ is his expected revenue from selling if he acts like type t .

Suppose \tilde{t} is drawn from a Uniform distribution on the interval from 0 to 100, but the value of the object to the buyer also depends on the quality (which the buyer would learn only after the transaction) and would be $1.5\tilde{t}$. That is, the object would always be worth 50% more to the buyer.

If (w,q) satisfies the incentive constraints and $U(w,q|t) \geq 0$, the buyer's expected gain from trade is $\int_0^{100} [1.5tq(t) - w(t)] f(t) dt = \int_0^{100} [1.5t - t - F(t)/f(t)] q(t) f(t) dt - U(w,q|100) = \int_0^{100} [1.5t - 2t] q(t) dt / 100 - U(w,q|100) \leq 0$. The buyer can only expect to lose if any $q(t) > 0$.

Facts about Uniform distributions. Suppose that \tilde{X} is a random variable drawn from a Uniform distribution on the interval from A to B , where $A < B$. Then $E(\tilde{X}) = (A+B)/2$, and $\forall \theta \in [A,B]$: $F(\theta) = P(\tilde{X} \leq \theta) = P(\tilde{X} < \theta) = (\theta - A)/(B - A)$, $f(\theta) = F'(\theta) = 1/(B - A)$, $F(\theta)/f(\theta) = \theta - A$, $E(\tilde{X} | \tilde{X} \leq \theta) = E(\tilde{X} | \tilde{X} < \theta) = (A + \theta)/2$, $E(\tilde{X} | \tilde{X} \geq \theta) = E(\tilde{X} | \tilde{X} > \theta) = (\theta + B)/2$.

Let's do the analogous result for the case where **the agent is a buyer** of some object.

Here the agent's type t is interpreted as his valuation of the object.

Now let $q(t)$ denote the probability of the agent buying the object if her type is t , and let $\omega(t)$ denote expected amount that the agent will have to pay if her type is t .

(If $\hat{w}(t)$ denotes the price that the type- t agent will pay if she buys, and if she would pay nothing if she does not buy the object, then our $\omega(t)$ is equal to $q(t) \hat{w}(t)$.)

So the expected gains from trade for a type- t buyer are $U(\omega, q|t) = t q(t) - \omega(t)$.

An incentive compatible trading plan must satisfy, for all types s and t in the interval $[A, B]$,

$$U(\omega, q|t) = t q(t) - \omega(t) \geq t q(s) - \omega(s) = U(\omega, q|s) + (t-s)q(s).$$

Then the envelope theorem yields $U'(\omega, q|t) = q(t)$ and so $U(\omega, q|t) = U(\omega, q|A) + \int_A^t q(s) ds$.

So the expected payment from type t is $\omega(t) = t q(t) - U(\omega, q|t) = t q(t) - \int_A^t q(s) ds - U(\omega, q|A)$.

The overall expected payment from the buyer, before her type is not known, is

$$\begin{aligned} \int_A^B \omega(t) f(t) dt &= \int_A^B [t q(t) - \int_A^t q(s) ds - U(\omega, q|A)] f(t) dt \\ &= \int_A^B t q(t) f(t) dt - \int_A^B \int_A^t q(s) ds f(t) dt - U(\omega, q|A) \\ &= \int_A^B s q(s) f(s) ds - \int_A^B \int_s^B f(t) dt q(s) ds - U(\omega, q|A) \\ &= \int_A^B [s f(s) - (1-F(s))] q(s) ds - U(\omega, q|A) = \int_A^B q(t) [t - (1-F(t))/f(t)] f(t) dt - U(\omega, q|A) \end{aligned}$$

Now consider a trading problem where both the seller and buyer have independent private values for the object that they may trade. Let 1 denote the seller, and let 2 denote the buyer.

The seller's type t_1 is what the object is worth to him, and the buyer's type t_2 is what the object would be worth to her. Suppose that each trader's type is his or her private information, but each thinks the other's type is drawn from a cumulative distribution F on the interval $[A, B]$. $f = F'$.

Now we just put subscripts 1 and 2 on all our past analysis. Let $q_1(t_1)$ and $\omega_1(t_1)$ denote the conditional probability of trade and the conditional expected payment if the seller's type is t_1 .

Let $q_2(t_2)$ and $\omega_2(t_2)$ denote the conditional probability of trade and the conditional expected payment if the buyer's type is t_2 . When we don't know either type, the expected payment is

$$\int_A^B \omega_1(t_1) f(t_1) dt_1 = \int_A^B \omega_2(t_2) f(t_2) dt_2 = \int_A^B \int_A^B \omega(t_1, t_2) f(t_1) f(t_2) dt_1 dt_2$$

With incentive compatibility for seller and buyer, this equation becomes:

$$\begin{aligned} U_1(\omega, q|B) + \int_A^B q_1(t_1) [t_1 + F(t_1)/f(t_1)] f(t_1) dt \\ = \int_A^B q_2(t_2) [t_2 - (1-F(t_2))/f(t_2)] f(t_2) dt_2 - U_2(\omega, q|A). \end{aligned}$$

Here $U_1(\omega, q|B)$ is the expected gains from trade for the highest type of seller, and $U_2(\omega, q|A)$ is the expected gains from trade for the lowest type of buyer, which are the types least eager to trade.

Let $q(t_1, t_2)$ denote the probability of trade occurring when the seller is t_1 and the buyer is t_2 .

So $q_1(t_1) = \int_A^B q(t_1, t_2) f(t_2) dt_2$ and $q_2(t_2) = \int_A^B q(t_1, t_2) f(t_1) dt_1$.

Then with participation constraints for the two least-eager-to-trade types, we get

$$\begin{aligned} 0 \leq U_1(\omega, q|B) + U_2(\omega, q|A) \\ = \int_A^B \int_A^B q(t_1, t_2) \{ [t_2 - (1-F(t_2))/f(t_2)] - [t_1 + F(t_1)/f(t_1)] \} f(t_1) f(t_2) dt_1 dt_2. \end{aligned}$$

Consider the case of Uniformly distributed types on $[0, 1]$. Then $A=0$, $B=1$, $F(t) = t$, $f(t) = 1$,

and so we get $0 \leq \int_A^B \int_A^B q(t_1, t_2) \{ [2t_2 - 1] - [2t_1] \} dt_1 dt_2 = 2 E\{q(\tilde{t}_1, \tilde{t}_2)[\tilde{t}_2 - \tilde{t}_1 - 1/2]\}$.

Thus, the total gains from trade $t_2 - t_1$ must have a conditional expected value of at least $1/2$ when trade occurs. But $E(\tilde{t}_2 - \tilde{t}_1 | \tilde{t}_2 > \tilde{t}_1) = [\int_0^1 \int_0^t (t-s) 2 ds dt] = 1/3$.

Examples Suppose the agent's has 2 possible cost-types: $\theta_L=1$ or $\theta_H=2$, each with probability $1/2$. The principal's value of effort q from an agent of type θ is $\pi(q|\theta) = (1+\theta)q^{0.5}$, so the principal's gain from trade when paying w for effort q is $2\sqrt{q} - w$ if $\tilde{\theta}=\theta_L$, $3\sqrt{q} - w$ if $\tilde{\theta}=\theta_H$. We can apply case 1 from our previous analysis of adverse-selection problem with two types. The Lagrange multiplier of the H|L incentive constraint is $\mu = p_L = 0.5$, and the Lagrangian is: $\mathcal{L} = 0.5\{3q_H^{0.5} - q_H[2+(2-1)0.5/0.5]\} + 0.5\{2q_L^{0.5} - q_L[1]\}$. To maximize this over (q_H, q_L) , we need $0 = 3(0.5)q_H^{-0.5} - 3$ and $0 = 2(0.5)q_L^{-0.5} - 1$, and so $q_H = 0.25$ and $q_L = 1$. Because $q_H < q_L$, we know that the case 1 assumptions about binding constraints will be OK here. The values of w_H and w_L are determined by the binding constraint equations: To make H-participation binding, $w_H = \theta_H q_H = 2 \times 0.25 = 0.5$. To make H|L-incentive binding, $w_L = \theta_L q_L + w_H - \theta_L q_H = 1 \times 1 + 0.5 - 1 \times 0.25 = 1.25$. [If we had changed $\pi(q_H|\theta_H)$ to be $A_H \sqrt{q_H}$ with some $A_H > 6$, then the above analysis would have yielded $q_H > q_L$, and so the pooling case 2 would have applied.]

A continuous example: Now suppose that $\tilde{\theta}$ is drawn from a Uniform distribution on $[1,2]$, keeping everything else the same as in the previous example, with $\pi(q|\theta) = (1+\theta)q^{0.5}$. So $F(\theta) = (\theta-1)/(2-1)$, $f(\theta) = 1/(2-1) = 1$, $F(\theta)/f(\theta) = \theta-1$, for any θ in $[1,2]$.

Then the principal's expected gains from trade are $\int_1^2 [\pi(q(\theta)|\theta) - w(\theta)] f(\theta) d\theta = \int_1^2 \{\pi(q(\theta)|\theta) - q(\theta)[\theta + F(\theta)/f(\theta)]\} f(\theta) d\theta - U(w, q|2) = \int_1^2 [(\theta+1)q(\theta)^{0.5} - q(\theta)(2\theta-1)] d\theta - [w(2) - 2q(2)]$.

To maximizing the integrand at every θ , we want $0 = 0.5(\theta+1)q^{-0.5} - (2\theta-1)$, which yields $q(\theta) = [0.5(\theta+1)/(2\theta-1)]^2$.

Because this $q(\theta)$ is monotone decreasing in θ , we know that it is actually feasible.

[If this $q(\theta)$ were increasing over any part of the interval, then it would not be feasible, and we would have an "irregular" case which is more complicated to solve (e.g: Myerson, 1981).]

To get $U(w, q|2) = 0$, we let $w(2) = 2q(2) = 2 \times 0.25 = 0.5$.

Then the information-rent equations give us

$$U(w, q|\theta) = \int_{\theta}^2 q(t) dt = 0.25[0.25t + 0.75\text{LN}(2t-1) - 2.25/(4t-2)]|_{\theta}^2,$$

$$w(\theta) = \theta q(\theta) + \int_{\theta}^2 q(t) dt, \text{ which yields } w(2) = 0.25 \text{ and } w(1) = 1.456$$

Auction: There are n possible suppliers of a service worth V to the buyer. Each supplier i has a cost type $\tilde{\theta}_i$ independently drawn from a Uniform distribution on an interval $[A, B]$.

So $F(\theta_i) = (\theta_i - A)/(B - A)$, $f(\theta_i) = 1/(B - A)$, $F(\theta_i)/f(\theta_i) = \theta_i - A$, for any θ_i in $[A, B]$.

An auction is planned, to select the supplier who actually sells to the buyer.

Let $q_i(\theta_1, \dots, \theta_n)$ denote the conditional probability that i will sell to the buyer, and let $w_i(\theta_1, \dots, \theta_n)$ denote the conditional expected payment to i , given these types $\tilde{\theta}_1 = \theta_1, \dots, \tilde{\theta}_n = \theta_n$.

Integrating over all types other than θ_i , let $Q_i(\theta_i)$ denote the marginal probability that i will sell, and let $W_i(\theta_i)$ denote the expected payment to i , given his own type $\tilde{\theta}_i = \theta_i$.

[If $n=2$ then $Q_1(\theta_1) = \int_A^B q_1(\theta_1, \theta_2) f(\theta_2) d\theta_2$, $W_1(\theta_1) = \int_A^B w_1(\theta_1, \theta_2) f(\theta_2) d\theta_2$.]

Expected payment to i is $EW_i(\tilde{\theta}_i) = U_i(B) + \int_A^B Q_i(t)[t + t_i - A]f(t)dt = U_i(B) + E[Q_i(\tilde{\theta}_i)(2\tilde{\theta}_i - A)]$

Suppose the buyer never pays more than B , so $U_i(w, q|B) = 0$. Then the buyer's expected profit is

$$\sum_i E\{Q_i(\tilde{\theta}_i)[V - (2\tilde{\theta}_i - A)]\} = \int_A^B \dots \int_A^B \sum_i q_i(\theta_1, \dots, \theta_n)[V - (2\theta_i - A)] f(\theta_1) \dots f(\theta_n) d\theta_1 \dots d\theta_n.$$

To maximize expected profit, buy from i (let $q_i=1$) only when $\theta_i = \min\{\theta_1, \dots, \theta_n\} < (V+A)/2$.

Incentive-efficiency in trading problems with one-sided private information.

Consider a seller who has a single indivisible asset that he could sell to a unique buyer.

The seller has private information about the quality of his asset that is summarized by his type θ .

Here $\tilde{\theta} \in \Theta = \{A, A+\delta, A+2\delta, \dots, A+k\delta = B\}$ is the seller's value of keeping his asset.

For each θ in Θ , $p(\theta)$ denotes the probability (assessed by the buyer) of the seller being type θ ,

$g(\theta)$ denotes the expected value of the seller's asset to the buyer when the seller's type is θ .

The probability $q(\theta)$ of trading the asset and the expected payment $w(\theta)$ from buyer to seller can

depend on the seller's self-reported type, so a trading plan specifies numbers $w(\theta)$ and $q(\theta)$

such that $w(\theta) \in \mathbb{R}$, $0 \leq q(\theta) \leq 1$, $\forall \theta \in \Theta$.

The agent could misrepresent his type, however. A plan $(q(\theta), w(\theta))_{\theta \in \Theta}$ is incentive compatible iff it gives the seller no incentive to misrepresent his type, that is:

$$w(\theta) - \theta q(\theta) \geq w(t) - \theta q(t), \quad \forall \theta \in \Theta, \quad \forall t \in \Theta. \quad [t|\theta\text{-incentive constraint}]$$

When they meet to bargain, the buyer wants to maximize $U_2(w, q) = \sum_{t \in \Theta} p(t)[g(t)q(t) - w(t)]$,

but the seller wants to maximize $U_1(w, q|\theta) = w(\theta) - \theta q(\theta)$ for his true type θ .

We don't know the true type $\tilde{\theta}$. So let us say that an incentive-compatible trading plan (w, q) is incentive-efficient iff there is no other incentive-compatible plan (\hat{w}, \hat{q}) that is better for the buyer

$[U_2(\hat{w}, \hat{q}) > U_2(w, q)]$ and all types of seller $[U_1(\hat{w}, \hat{q}|t) > U_1(w, q|t), \quad \forall t \in \Theta]$.

So (w, q) is incentive-efficient iff there exist weights $\lambda(\theta) \geq 0$ that make it optimal for

$$[*] \quad \begin{aligned} & \text{maximize } \sum_{\theta \in \Theta} p(\theta)[g(\theta)q(\theta) - w(\theta)] + \sum_{\theta \in \Theta} \lambda(\theta)[w(\theta) - \theta q(\theta)] \\ & \text{subject to } w(\theta) \in \mathbb{R}, \quad 0 \leq q(\theta) \leq 1, \quad \forall \theta \in \Theta; \\ & w(\theta) - \theta q(\theta) \geq w(t) - \theta q(t), \quad \forall \theta \in \Theta, \quad \forall t \in \Theta. \end{aligned} \quad [\alpha(t|\theta)]$$

With these one-dimensional types entering linearly into the agent's utility function, it is enough to check that no type wants to misrepresent one step higher or one step lower. (If types t and $t+\delta$ don't want to imitate each other then higher types would prefer to act like $t+\delta$, lower like t .)

So we only need to consider Lagrange multipliers $\alpha(\theta+\delta|\theta) \geq 0$, $\alpha(\theta|\theta+\delta) \geq 0$, $\forall \theta \in \Theta$.

Then (with $0 = \alpha(B|B+\delta) = \alpha(B+\delta|B) = \alpha(A|A-\delta) = \alpha(A-\delta|A)$) the Lagrangean becomes

$$\begin{aligned} \mathcal{L}(w, q; \lambda, \alpha) = & \sum_{\theta \in \Theta} \{p(\theta)[g(\theta)q(\theta) - w(\theta)] + \lambda(\theta)[w(\theta) - \theta q(\theta)]\} \\ & + \sum_{\theta < B} \alpha(\theta+\delta|\theta)[w(\theta) - \theta q(\theta) - w(\theta+\delta) + \theta q(\theta+\delta)] \\ & + \sum_{\theta < B} \alpha(\theta|\theta+\delta)[w(\theta+\delta) - (\theta+\delta)q(\theta+\delta) - w(\theta) + (\theta+\delta)q(\theta)] \end{aligned}$$

The first-order optimality conditions for the type-conditional expected payment $w(t)$ become

$$0 = \partial \mathcal{L} / \partial w(t) = -p(t) + \lambda(t) + \alpha(t+\delta|t) - \alpha(t|t-\delta) + \alpha(t-\delta|t) - \alpha(t|t+\delta), \quad \forall t \in \Theta.$$

So $\lambda(t) = p(t) + \alpha(t|t-\delta) + \alpha(t|t+\delta) - \alpha(t+\delta|t) - \alpha(t-\delta|t) \geq 0 \quad \forall t$, and $\sum_t \lambda(t) = \sum_t p(t) = 1$.

With these equations, the Lagrangean simplifies to $\mathcal{L}(w, q; \lambda, \alpha) = \sum_{\theta \in \Theta} p(\theta)q(\theta)[g(\theta) - v(\theta, \alpha)]$ where $v(\theta, \alpha) = \theta + \delta[\alpha(\theta|\theta-\delta) - \alpha(\theta|\theta+\delta)]/p(\theta)$ is the θ -seller's virtual value of his asset.

Suppose we want to maximize the buyer's expected payoff, subject to incentive constraints and participation constraints that each type of seller must expect nonnegative gains from trade.

The participation constraint for the highest type $w(B) - Bq(B) \geq 0$ implies (with incentive compatibility) that all other types' participation constraints are satisfied. So we can give the B-participation constraint Lagrange multiplier $\lambda(B)=1$, and let $\lambda(\theta)=0$ for all $\theta < B$.

Then [*] can be solved in many cases with $\alpha(\theta-\delta|\theta)=0$ and $\alpha(\theta+\delta|\theta) = F(\theta) = \sum_{t \leq \theta} p(t) \quad \forall \theta$, provided that $v(\theta, \alpha) = \theta + \delta F(\theta-\delta)/p(\theta)$ is monotone increasing in θ (regularity).

Let's consider a simple example, from **MWG section 23.F**.

The seller's cost type is either $L=20$ or $H=40$, but the seller's asset is worth 10 more to the buyer, so $g(L) = 30$, $g(H) = 50$. For now, let's keep $p(H)=p_H$ as a parameter, with $p(L)=1-p_H$.

The Lagrangean now becomes

$$\mathcal{L}(w,q;\lambda,\alpha) = p_H q_H \{50 - [40 + (40-20)\alpha(H|L)/p_H]\} + p_L q_L \{30 - [20+(20-40)\alpha(L|H)/p_L]\}.$$

We must have $\lambda_H = p_H + \alpha(H|L) - \alpha(L|H) \geq 0$, $\lambda_L = p_L + \alpha(L|H) - \alpha(H|L) \geq 0$.

For all incentive compatible trading plans, we have $q_H \leq q_L$.

Choosing $q_H < 1$ can be a useful signal of the high type. But $0 < q_H < 1$ can be an optimal solution of the Lagrangean if and only if $50 = 40 + (40-20)\alpha(H|L)/p_H$, so $\alpha(H|L) = 0.5p_H$.

To complete a solution with $\alpha(H|L)=0.5p_H$, we have two cases to consider.

In case 1, we have incentive-efficient trading plans where the $L|H$ constraint is not binding.

So $\alpha(L|H)=0$. Then $\lambda_H = p_H + \alpha(H|L) = 1.5p_H \geq 0$ and $\lambda_L = p_L - \alpha(H|L) = 1 - 1.5p_H \geq 0$, which yields a solution when $p_H < 2/3$.

Then the Lagrangean simplifies to $p_H q_H (50 - 50) + p_L q_L (30 - 20)$, which is maximized by $q_L=1$.

In this case, with this (λ, α) , any incentive compatible plan in which the $H|L$ -incentive constraint is binding and $q_L=1$ will be incentive-efficient. (In fact, these are all the incentive-efficient plans.)

For example, consider the incentive-compatible plan that is best for sellers among all plans such that buyers get nonnegative expected gains from each type of seller (the best safe plan):

$$w_L = 30q_L, q_L=1, w_H = 50q_H, q_H = (30-20)/(50-20) = 1/3, \text{ so } U_L=30-20, U_H=(50-40)/3.$$

This plan is incentive-efficient when $p_H \leq 2/3$.

But when $p_H > 2/3$, this plan is interim Pareto-dominated by the pooling plan

$$q_L = q_H = 1, w_L = w_H = 50p_H + 30(1-p_H) = 30+20p_H > 43.333... \text{ when } p_H > 2/3.$$

In case 2, $p_H > 2/3$ and we have incentive-efficient plans where both incentive constraint bind.

For both incentive-constraints to bind, we must have pooling $q_H=q_L$ and $w_H=w_L$.

We can still use $\alpha(H|L) = 0.5p_H$ to eliminate q_H from the Lagrangean, but now we must use $\alpha(L|H)>0$ to prevent λ_L from becoming negative. We can get a solution with

$$\alpha(H|L) = 0.5p_H \text{ and } \alpha(L|H) = \alpha(H|L) - p_L = 1.5p_H - 1, \text{ so that } \lambda_L = 0 \text{ and } \lambda_H = 1.$$

Then the Lagrangean simplifies to $p_L q_L \{30 - [20 - 20(1.5p_H - 1)/(1 - p_H)]\}$, maximized by $q_L=1$.

With this (λ, α) in this case, any pooling plan with $q_H=q_L=1$ and $w_H=w_L$ (so that $q_L=1$ and both incentive constraints bind) will be incentive-efficient. (These are all the incentive-efficient plans.)

Now consider a market where there are many competitive buyers, any one of whom could use all the seller's supply. So buyers' competition should drive their expected profits to 0 in equilibrium.

If $p_H \leq 2/3$, then the best safe plan ($q_L=1, w_L=30, q_H=1/3, w_H=50/3$) is incentive-efficient and may be considered a competitive equilibrium for this market. But what if $p_H > 2/3$?

For example, suppose $p_H = 0.8$, so $50p_H + 30(1-p_H) = 46$.

If all buyers were expected to offer this best-safe plan, then a deviating buyer who offered ($\hat{q}=1, 46 > \hat{w} \geq 43.34$) could attract all sellers and make a positive profit on average.

But if all buyers competitively offered the break-even pooling plan ($q=1, w=46$), then a deviating buyer could gain by attracting H-sellers with $(\hat{q}, \hat{w})=(0.9, 48\hat{q})$, so $\hat{w} - 20\hat{q} < 26$, $\hat{w} - 40\hat{q} > 6$, $\hat{w} < 50\hat{q}$.

The right definition of equilibrium in such markets has remained unclear.

A Uniform bilateral trading problem like Akerlof's

Let's consider a bilateral trading problem where agent 1 is the seller of some unique object which he owns, and agent 2 is the only possible buyer of this object.

Depending on the object's quality, it may be worth as little as \$40 to agent 1 and \$60 to agent 2 (if its quality is low) or as much as \$100 to agent 1 and \$120 to agent 2 (if its quality is high).

Player 1 knows the quality of the object. Let 1's cost type \tilde{t}_1 is his value of keeping the object.

With any quality, the object would be worth \$20 more to player 2 than to player 1.

That is, given 1's type \tilde{t}_1 , the value of the object to player 2 would be $g(\tilde{t}_1) = \tilde{t}_1 + 20$.

Player 2's belief about \tilde{t}_1 is described by a Uniform distribution on the interval \$40 to \$100.

Game where buyer bids Suppose first that agent 2 can offer to buy for any positive price r , and then agent 1 will accept or reject the offer. If the offer is rejected then they each get profit 0.

If the offer is accepted then 1's profit is $r - \tilde{t}_1$ and 2's profit is $g(\tilde{t}_1) - r$.

In a subgame-perfect equilibrium, agent 1 will accept if $\tilde{t}_1 < r$, but agent 1 will reject if $\tilde{t}_1 > r$.

Agent 2's expected profit from offering any price r is $Y(r) = P(\tilde{t}_1 < r) [E(g(\tilde{t}_1) | \tilde{t}_1 < r) - r]$.

For any number r between 40 and 100, this expected profit is

$$\begin{aligned} Y(r) &= P(\tilde{t}_1 < r) [E(\tilde{t}_1 + 20 | \tilde{t}_1 < r) - r] = P(\tilde{t}_1 < r) [(E(\tilde{t}_1 | \tilde{t}_1 < r) + 20) - r] = \\ &= [(r - 40)/(100 - 40)] [(40 + r)/2 + 20 - r] = (r - 40)(80 - r)/120 = (-3200 + 40r - r^2)/120. \end{aligned}$$

This quadratic formula is maximized by letting $r = 60$.

(The buyer cannot gain by bidding less than 40 or more than 100, because a bid below 40 would be surely rejected, and a bid above 100 would be worse than the surely-accepted bid of 100.)

So in the unique subgame-perfect equilibrium of this game, agent 2 offers to buy for \$60, and agent 1 accepts if $\tilde{t}_1 < 60$. The probability of trade is $P(\text{trade}) = (60 - 40)/(100 - 40) = 1/3$.

Game where seller bids Suppose now that agent 1 can offer to buy for any positive price r , and then agent 2 will accept or reject the offer. If the offer is rejected then they each get profit 0.

If the offer is accepted, then 1's profit is $r - \tilde{t}_1$ and 2's profit is $g(\tilde{t}_1) - r$.

In this game, the price is named by the agent who has private information, and so signaling effects give us many equilibria.

We may also reinterpret this as a market, with many sellers (1) and enough identical buyers (2) to buy the entire supply, where each seller 1, knowing his own type, publicly commits to a price that he chooses. Then he must sell at this price (if any buyer accepts his bid) or keep the object.

Let's look first for an equilibrium where there is some price r such that agent 2 would surely accept an offer to sell for r but would surely reject an offer to sell for any price higher than r .

In this equilibrium, agent 1 will offer r if $\tilde{t}_1 < r$.

For agent 2 to accept the offer r , 2's expected profit from accepting r must not be negative, so $0 \leq E(V_2(\tilde{t}_1) | \tilde{t}_1 < r) - r = E(\tilde{t}_1 + 20 | \tilde{t}_1 < r) - r = (40 + r)/2 + 20 - r$, which implies $r \leq 80$.

For agent 2 to reject any offer to sell at a price higher than r , such a trade must be unprofitable for agent 2 when she makes the worst inference about agent 1, which is that his type is 40, in which case the object would be worth $40 + 20 = \$60$ to agent 2.

So we can construct such an equilibrium for any r such that $60 \leq r \leq 80$.

In such an equilibrium, types higher than r may be expected to make some offer higher than 120, which agent 2 could never profitably accept.

An offer between r and 120 may be rejected by agent 2 because this surprise offer may lead agent 2 to believe that 1's type is 40, in which case the object is only worth 60 to agent 2. Among these almost-pooling equilibria, agent 1 most prefers the equilibrium with $r = 80$.

In this equilibrium, the probability of trade is $\Pr(\text{trade}) = \Pr(\tilde{t}_1 < 80) = (80 - 40)/(100 - 40) = 2/3$.

Reinterpretation in market: We'd get excess demand if $r < 80$, and the market clears only at $r = 80$. So an uninformed Walrasian auctioneer who posts a price to clear the market would choose $r = 80$.

There are many other equilibria where different types of seller (1) choose different prices.

Let's look for an equilibrium in which some types of agent 1 would offer to sell for \$70, but all higher types would offer to sell for \$100, and agent 2 would be sure to accept \$70 but her probability of accepting \$100 would be between 0 and 1. To find this equilibrium, we have two unknowns to find: let q denote the probability that agent 2 would accept an offer of \$100, and let θ denote the highest type of agent 1 that would offer \$70.

For agent 2 to be willing to randomize between accepting and rejecting \$100, her expected profit from accepting it must be 0, and so

$$0 = E(g_2(\tilde{t}_1) | \tilde{t}_1 > \theta) - 100 = E(\tilde{t}_1 + 20 | \tilde{t}_1 > \theta) - 100 = (\theta + 100)/2 + 20 - 100, \text{ and so } \theta = 60.$$

For agent 1 to offer \$70 below when his type is below θ but \$100 when his type is above θ , we need that $70 - t_1 \geq q(100 - t_1)$ when $t_1 < \theta$, and $70 - t_1 \leq q(100 - t_1)$ when $t_1 > \theta$.

These inequalities imply $70 - \theta = q(100 - \theta)$, and so $q = (70 - 60)/(100 - 60) = 1/4$.

In this equilibrium, $\Pr(\text{trade}) = \Pr(\tilde{t}_1 < \theta) + \Pr(\tilde{t}_1 > \theta)q = (20/60) + (40/60)(1/4) = 1/2$.

This is also a market equilibrium, with no excess demand at either price: $E(g_2(\tilde{t}_1) | \tilde{t}_1 \leq \theta) = 70$.

There is a separating equilibrium in which each possible type t_1 of agent 1 would offer to sell for $r(t_1) = t_1 + 20$, and the probability of agent 2 accepting would depend on the offer r according to the formula $Q(r) = e^{-(r-60)/20}$, for any $r \geq 60$. The derivation is as follows:

For $r = t_1 + 20$ to maximize $Q(r)(r - t_1)$, we need $0 = Q'(r)(r - t_1) + Q(r)$ when $r = t_1 + 20$, and so $-1/20 = Q'(r)/Q(r) = d/dr \ln(Q(r))$.

In our usual notation for incentive-compatible trading plans, this separating equilibrium has, for any seller's type t , the type-conditional probability of trade $q(t) = Q(t+20) = e^{-(t-40)/20}$ and the type-conditional expected payment $w(t) = (t+20)q(t) = (t+20)e^{-(t-40)/20}$.

This function $q()$ could also be derived from the incentive constraints, for all t in $[40, 100]$: $w(t) - tq(t) = \max_s w(s) - tq(s) = \max_s q(s)(s+20-t)$, which yields $0 = q'(t)(t+20-t) - q(t)$. The low-end boundary condition $q(40)=1$ identifies the best separating equilibrium.

If we compare different types expected gains from trade in these three equilibria, we find that low types ($t_1 \leq 73$) prefer the pooling equilibrium where all trade occurs at the price 80, middle types ($74 \leq t_1 \leq 94$) prefer the two-price equilibrium where trade occurs at 70 or 100, and high type ($t_1 \geq 95$) prefer the separating equilibrium. But the separating equilibrium is actually interim pareto-dominated by the following semi-separating equilibrium:

Low types in $[40, 60]$ are pooled together and sell for \$70 with probability $q=1$, but any higher type $t > 60$ separates and sells with a probability $q(t) < 1$ at price $t+20$.

For $t=60$ to be indifferent between selling for price 70 with probability 1 or selling for price $60+20$ with a lower probability, that probability must be $q(60)=0.5$. Then as before, we get $q'(t)/q(t) = -1/20$ and so $q(t) = 0.5e^{-(t-60)/20}$. (Compare: $e^{-(t-40)/20} = 0.368 < 0.5$.)

Equilibrium in an insurance market with adverse selection

Consider a large insurance market with competitive supply of insurance policies. Customers arrive in the market as a continuous flow over time. Each customer privately knows his own type t in some finite set of types T , which he can misrepresent if he not given an incentive for honesty. Let $p(t)$ be the fraction of type- t customers among new arrivals into the market over any period. Each customer can buy one insurance policy from one insurer.

Suppose that there is a given finite set M of possible policies, but the value of a policy to a customer and its expected cost to the insurer will depend on the customer's type.

Let $V(m,t)$ denote the (certainty-equivalent) value of policy m to a customer of type t , and let $C(m,t)$ denote the expected cost of policy m to the insurer when the customer is type t .

Suppose that M includes the noninsurance alternative m_0 with $V(m_0,t) = 0 = C(m_0,t) \quad \forall t \in T$.

To describe the anticipated plan of transactions in such an insurance market, we may let $\mu(m|t)$ denote the probability that a customer of type t will get a policy m in the market, and let $x(t)$ denote the expected price that a customer of type t will pay to buy his policy in the market. We assume that each customer only buys one policy.

So μ satisfies the probability constraints $\mu(m|t) \geq 0 \quad \forall m \in M, \forall t \in T$, and $\sum_{m \in M} \mu(m|t) = 1, \quad \forall t \in T$.

Expected gains from trade for a type- t customer will be $U(\mu,x|t) = \sum_{m \in M} \mu(m|t)V(m,t) - x(t)$.

Expected profit for insurers from a type- t customer will be $x(t) - \sum_{m \in M} \mu(m|t)C(m,t)$.

Because customers can misrepresent their types, such a market plan must be incentive compatible, that is, it must satisfy the informational incentive constraints

$$\sum_{m \in M} \mu(m|t)V(m,t) - x(t) \geq \sum_{m \in M} \mu(m|s)V(m,t) - x(s), \quad \forall t \in T, \forall s \in S.$$

When a customer meets an insurer in this competitive market, the customer knows his own type t and the customer knows that he can get the equilibrium payoff $U(\mu,x|t)$ from many other insurers.

But if insurers expected to lose money by selling to type- t customers in this market, then any one insurer could increase profits by encouraging these type- t customers to go get the payoff $U(\mu,x|t)$ from some other insurer. This argument suggests that a market equilibrium should satisfy the safe-profitability condition that insurers do not lose money on an type of customer

$$x(t) - \sum_{m \in M} \mu(m|t)C(m,t) \geq 0, \quad \forall t \in T.$$

We are assuming that insurance companies are risk neutral and have no limit on the number of customers that they can insure, and so each insurer could supply the entire market, but we are also assuming that the market is actually divided among many such companies.

So if the insurers were making positive profits on average, then any one insurer would have an incentive to reduce his prices slightly and attract all the customers in the market.

Thus, competition among insurers should be good for customers, giving them the best possible menu of prices for the policies that they get.

In fact, there exists a safely-profitable incentive-compatible market plan (μ, x) that is best for all customers in the sense that, for any other $(\hat{\mu}, \hat{x})$ that is safely profitable and incentive compatible, we must have $U(\mu, x|t) \geq U(\hat{\mu}, \hat{x}|t)$, $\forall t \in T$.

This best safely-profitable incentive-compatible plan (μ, x) can be found by solving the following optimization problem:

$$[1] \quad \text{choose } (\mu, x) \text{ to maximize } \sum_{t \in T} p(t) [\sum_{m \in M} \mu(m|t)V(m, t) - x(t)] \text{ subject to}$$

$$x(t) \in \mathbb{R}, \mu(m|t) \geq 0, \forall t \in T, \forall m \in M; \sum_{m \in M} \mu(m|t) = 1, \forall t \in T; \quad \{\text{probability}\}$$

$$\sum_{m \in M} \mu(m|t)V(m, t) - x(t) - [\sum_{m \in M} \mu(m|s)V(m, t) - x(s)] \geq 0, \forall t \in T, \forall s \in S; \{\text{incentive, } \alpha(s|t)\}$$

$$x(t) - \sum_{m \in M} \mu(m|t)C(m, t) \geq 0, \forall t \in T. \quad \{\text{safe-profitability, } \gamma(t)\}$$

To verify that an optimal solution of [1] is best for all types of customers, suppose to the contrary that there is some other safely-profitable incentive-compatible plan $(\hat{\mu}, \hat{x})$ that some types prefer. Let $S(\hat{\mu}, \hat{x})$ denote the set of types that prefer this plan, so $S(\hat{\mu}, \hat{x}) = \{s \in T: U(\hat{\mu}, \hat{x}|s) > U(\mu, x|s)\}$. Then let $(\bar{\mu}, \bar{x})$ be the plan that gives each type its preference among (μ, x) and $(\hat{\mu}, \hat{x})$; that is: $\bar{\mu}(m|t) = \hat{\mu}(m|t)$ and $\bar{x}(t) = \hat{x}(t)$ if $t \in S$, and $\bar{\mu}(m|t) = \mu(m|t)$ and $\bar{x}(t) = x(t)$ otherwise. It is straightforward to check that $(\bar{\mu}, \bar{x})$ is also incentive compatible and safely profitable, and so it would be a better solution to problem [1] if $S(\hat{\mu}, \hat{x})$ were nonempty. (We assume all $p(t) > 0$.)

An optimal solution (μ, x) of [1] is supported by nonnegative Lagrange multipliers (α, γ) such that

$$0 = \partial \mathcal{L} / \partial x(t) = -p(t) - \sum_{s \in T} \alpha(s|t) + \sum_{s \in T} \alpha(t|s) + \gamma(t), \quad \forall t \in T;$$

$$\alpha(s|t) [\sum_{m \in M} \mu(m|t)V(m, t) - x(t) - \sum_{m \in M} \mu(m|s)V(m, t) + x(s)] = 0, \quad \forall t \in T, \forall s \in T.$$

Then the optimal μ maximizes the (α, γ) -Lagrangian \mathcal{L} subject to the probability constraints:

$$\mathcal{L}(\mu, x; \alpha, \gamma) = \sum_{t \in T} \sum_{m \in M} \mu(m|t) \{ [p(t) + \sum_{s \in T} \alpha(s|t)]V(m, t) - \sum_{s \in T} \alpha(t|s)V(m, s) - \gamma(t)C(m, t) \}.$$

For each t , the solution puts positive weight $\mu(m|t) > 0$ only on m that maximize the virtual gains $\{ [p(t) + \sum_{s \in T} \alpha(s|t)]V(m, t) - \sum_{s \in T} \alpha(t|s)V(m, s) - \gamma(t)C(m, t) \}$ over all m in M .

In fact, [1] is a linear programming problem. Its Lagrange multipliers (α, γ) solve a dual problem:

$$[2] \quad \text{minimize } \sum_{t \in T} \max_{m \in M} \{ [p(t) + \sum_{s \in S} \alpha(s|t)]V(m, t) - \sum_{s \in S} \alpha(t|s)V(m, s) - \gamma(t)C(m, t) \}$$

$$\text{subject to } \alpha(s|t) \geq 0, \gamma(t) \geq 0, \forall t \in T, \forall s \in T;$$

$$\gamma(t) - p(t) - \sum_{s \in S} \alpha(s|t) + \sum_{s \in S} \alpha(t|s) = 0, \quad \forall t \in T. \quad \{x(t)\}$$

Notice that the $\gamma(t)$ are nonnegative, and they sum to 1 because

$$\sum_{t \in T} \gamma(t) = \sum_{t \in T} [p(t) + \sum_{s \in S} \alpha(s|t) - \sum_{s \in S} \alpha(t|s)] = \sum_{t \in T} p(t) = 1.$$

So γ looks like a probability distribution over the types, but it may be different from p .

(The definition of the best safely-profitable incentive-compatible plan does not depend on p , and so the optimal solution to [1] is actually independent of the distribution p , as long as all $p(t) > 0$.)

The solution to [2] may depend on p , however.)

Consider an optimization problem that differs from [1] in that the safe-profitability constraints are replaced by an expected-profitability condition that uses the probability distribution γ :

- [3] choose (μ, x) to maximize $\sum_{t \in T} p(t) [\sum_{m \in M} \mu(m|t)V(m,t) - x(t)]$ subject to
- $x(t) \in \mathbb{R}, \mu(m|t) \geq 0, \forall t \in T, \forall m \in M; \sum_{m \in M} \mu(m|t) = 1, \forall t \in T;$ {probability}
 - $\sum_{m \in M} \mu(m|t)V(m,t) - x(t) \geq \sum_{m \in M} \mu(m|s)V(m,t) - x(s), \forall t \in T, \forall s \in S;$ {incentive, $\alpha(s|t)$ }
 - $\sum_{t \in T} \gamma(t)[x(t) - \sum_{m \in M} \mu(m|t)C(m,t)] \geq 0.$ { γ -expected profitability}

The optimal solution (μ, x) for problem [1] is still optimal for this problem [3], because it satisfies all optimality conditions with the same Lagrange multipliers α that we got for [1]. The only difference is that the $\gamma(t)$ are parameters in [3], but were Lagrange multipliers in [1].

Now, returning to the theory of competitive equilibrium in such insurance markets, suppose that there are many different insurers all offering policies according to the best safe plan (μ, x) .

Each customer of type t gets expected payoff $U(\mu, x|t) = \sum_{m \in M} \mu(m|t)V(m,t) - x(t)$.

Although each type t 's rate of arrival into the market is proportional to $p(t)$ (by assumption), different types may circulate longer than other types.

Think about a version of Gresham's law: that bad types may circulate more than good types. Types that spend a longer time in the market then will be over-represented in the stock of searching customers at any point in time.

So we can assume that, among the customers who are searching for an insurer at any point in time, the fraction of type- t individuals is $\gamma(t)$.

With this population searching for policies, imagine an insurer who is thinking about making a short-term offer to sell some policy n in M for some price y .

Let $S(n, y)$ denote the set of types that prefer this offer, so $S(n, y) = \{s \in T: V(n, s) - y > U(\mu, x|s)\}$.

We can now show that any such offering must expect to lose money:

Claim: If $S(n, y) \neq \emptyset$ then $\sum_{s \in S(n, y)} \gamma(s)[y - C(n, s)] < 0$.

To prove this claim, suppose to the contrary that some types s strictly gain from (n, y) , compared to what they get from (μ, x) , and the insurer gets nonnegative expected profit from these types.

Then (μ, x) would not be optimal for the γ -weighted problem [3], because the objective could be strictly increased by switching all the types who strictly prefer the new (n, y) -offer over to this offer, while leaving all the other types with the same treatment that they got in (μ, x) . (Here we are using the fact that, under (μ, x) , the insurers expect nonnegative profit from every type, so diverting some types over to the new offer does not undermine the profitability of (μ, x) for all other types. Incentive compatibility is preserved as we divert only the types that want to go.)

So the following $(\bar{\mu}, \bar{x})$ would be better than (μ, x) as a solution for problem [3]:

if $U(\mu, x|t) \geq V(n, t) - y$ then $\bar{\mu}(m|t) = \mu(m|t) \forall m$ and $\bar{x}(t) = x(t)$;

if $U(\mu, x|t) < V(n, t) - y$ then $\bar{\mu}(n|t) = 1, \bar{\mu}(m|t) = 0 \forall m \neq n$, and $\bar{x}(t) = y$.

But (μ, x) is an optimal solution of problem [3], and so the claim must be true.

Thus, when the shopping population has the type-distribution γ , the customers' best safely-profitable plan (μ, x) that solves problem [1] becomes a competitive equilibrium where no supplier can hope to earn positive profits by offering any policy at any other price.

Now let's add **liquidity or limited-liability constraints** on the customers: Suppose that the amount that each type t can pay is bounded by his assets $A(t)$. So problem [1] becomes

$$\begin{aligned}
[1a] \quad & \text{choose } (\mu, x) \text{ to maximize } \sum_{t \in T} p(t) [\sum_{m \in M} \mu(m|t)V(m,t) - x(t)] \text{ subject to} \\
& x(t) \in \mathbb{R}, \mu(m|t) \geq 0, \forall t \in T, \forall m \in M; \sum_{m \in M} \mu(m|t) = 1, \forall t \in T; & \{\text{probability}\} \\
& \sum_{m \in M} \mu(m|t)V(m,t) - x(t) - [\sum_{m \in M} \mu(m|s)V(m,t) - x(s)] \geq 0, \forall t \in T, \forall s \in S; & \{\text{incentive, } \alpha(s|t)\} \\
& x(t) - \sum_{m \in M} \mu(m|t)C(m,t) \geq 0, \forall t \in T; & \{\text{safe-profitability, } \gamma(t)\} \\
& x(t) \leq A(t), \forall t \in T. & \{\text{liquidity}\}
\end{aligned}$$

An optimal solution of [1a] is supported by nonnegative Lagrange multipliers (α, γ) such that, for each type t in T :

$$\text{if } x(t) < A(t) \text{ then } 0 = \partial \mathcal{L} / \partial x(t) = -p(t) - \sum_{s \in T} \alpha(s|t) + \sum_{s \in T} \alpha(t|s) + \gamma(t),$$

$$\text{if } x(t) = A(t) \text{ then } 0 \leq \partial \mathcal{L} / \partial x(t) = -p(t) - \sum_{s \in T} \alpha(s|t) + \sum_{s \in T} \alpha(t|s) + \gamma(t),$$

$$\text{and } \alpha(s|t) [\sum_{m \in M} \mu(m|t)V(m,t) - x(t) - \sum_{m \in M} \mu(m|s)V(m,t) + x(s)] = 0, \forall t \in T, \forall s \in T.$$

Then the optimal μ maximizes the (α, γ) -Lagrangian \mathcal{L} subject to the probability constraints:

$$\mathcal{L}(\mu, x; \alpha, \gamma) = \sum_{t \in T} \sum_{m \in M} \mu(m|t) \{ [p(t) + \sum_{s \in T} \alpha(s|t)] V(m,t) - \sum_{s \in T} \alpha(t|s) V(m,s) - \gamma(t) C(m,t) \}.$$

For each t , the solution puts positive weight $\mu(m|t) > 0$ only on m that maximize the virtual gains $\{ [p(t) + \sum_{s \in T} \alpha(s|t)] V(m,t) - \sum_{s \in T} \alpha(t|s) V(m,s) - \gamma(t) C(m,t) \}$ over all m in M .

The $\gamma(t)$ are all nonnegative, but their sum may be greater than 1 (if liquidity binds), because

$$\sum_{t \in T} \gamma(t) \geq \sum_{t \in T} [p(t) + \sum_{s \in S} \alpha(s|t) - \sum_{s \in S} \alpha(t|s)] = \sum_{t \in T} p(t) = 1.$$

Let $\Gamma = \sum_{t \in T} \gamma(t)$, and let $\hat{\gamma}(t) = \gamma(t)/\Gamma$. So the $\hat{\gamma}(t)$ numbers look like probabilities of types.

As before, let us replace the safe-profitability constraints with an expected-profitability condition that uses the probability distribution $\hat{\gamma}$ (instead of p):

$$\begin{aligned}
[3a] \quad & \text{choose } (\mu, x) \text{ to maximize } \sum_{t \in T} p(t) [\sum_{m \in M} \mu(m|t)V(m,t) - x(t)] \text{ subject to} \\
& x(t) \in \mathbb{R}, \mu(m|t) \geq 0, \forall t \in T, \forall m \in M; \\
& \sum_{m \in M} \mu(m|t) = 1, \forall t \in T; & \{\text{probability}\} \\
& \sum_{m \in M} \mu(m|t)V(m,t) - x(t) \geq \sum_{m \in M} \mu(m|s)V(m,t) - x(s), \forall t \in T, \forall s \in S; & \{\text{incentive, } \alpha(s|t)\} \\
& \sum_{t \in T} \hat{\gamma}(t) [x(t) - \sum_{m \in M} \mu(m|t)C(m,t)] \geq 0. & \{\hat{\gamma}\text{-expected profitability}\}
\end{aligned}$$

The optimal solution (μ, x) for problem [1a] is still optimal for this problem [3a], because (μ, x) satisfies all optimality conditions for [3a] with the same Lagrange multipliers α as for [1a], when Γ is the Lagrange multiplier for the $\hat{\gamma}$ -expected profitability constraint (as $\Gamma \hat{\gamma}(t) = \gamma(t) \forall t$).

So now let us suppose that the market is operating competitively, with many different insurers all offering policies according to (μ, x) , so that each customer of type t gets expected payoff

$U(\mu, x|t) = \sum_{m \in M} \mu(m|t)V(m,t) - x(t)$, and let us also suppose that, at any point of time, the fraction of each type t in the shopping population is $\hat{\gamma}(t)$. Then as before, it is impossible for any firm to earn positive profits by a short-term offer of any contract that attracts only types t that get more than $U(\mu, x|t)$ from the new offer. (If any such offer were profitable for the firm, then it would create a contradiction of the fact that the safely-profitable (μ, x) is optimal for [3a].)

An optional exercise to verify a randomized equilibrium of the competitive screening game in MWG 13.D, where two firms simultaneously offer menus of employment contracts:

Consider a discrete example of our basic labor-market model. There are two types of workers: low types with reservation wage $\theta_L=2$ and productivity $g(\theta_L)=3$, and high types with reservation wage $\theta_H=4$ and productivity $g(\theta_H)=5$. The fraction of high-type workers in the population is $p_H=0.8$, and so the expected productivity of a randomly-sampled worker is $E(g(\tilde{\theta})) = 0.2 \times 3 + 0.8 \times 5 = 4.6$. Each worker has one unit of labor to sell, but he can accept a contract to sell only some fraction q of it to one firm. So an employment contract specifies the fraction q of the worker's labor that he will sell, and a total salary w that he will be paid. So when a worker of type θ accepts contract (q,w) , the worker's payoff is $w - \theta q$, and the firm's payoff is $g(\theta)q - w$. We may say that a menu of contracts $((q_L, w_L), (q_H, w_H))$ is incentive compatible if, among these two contracts, low types weakly prefer (q_L, w_L) and high types weakly prefer (q_H, w_H) . Given any menu $((q_L, w_L), (q_H, w_H))$, applying the first contract to a low-type worker would give the firm payoff $y_L = 3q_L - w_L$ and give the worker payoff $u_L = w_L - 2q_L$, and applying the second contract to a high-type worker would give the firm payoff $y_H = 5q_H - w_H$ and give the worker payoff $u_H = w_H - 4q_H$.

We may say that the H|L incentive constraint is binding in such an incentive compatible menu if low types would be indifferent among the two contracts. (Read "H|L" as "H-given-L.")

(a) Show that, for any incentive-compatible menu $((q_L, w_L), (q_H, w_H))$, there is an amount \hat{w}_L such that $((1, \hat{w}_L), (q_H, w_H))$ is incentive compatible, the low-type workers would be indifferent between $(1, \hat{w}_L)$ and (q_L, w_L) , and firms would prefer hiring low-type workers under the $(1, \hat{w}_L)$ contract than the (q_L, w_L) contract. (Thus we can assume without loss of generality that firms only offer such incentive compatible menus where low types work full-time and the H|L constraint is binding.)

(b) Show that, when $((1, w_L), (q_H, w_H))$ is an incentive-compatible menu in which the H|L incentive constraint is binding, the firm's profits (y_L, y_H) from hiring each type of worker (in its intended contract) can be expressed as a linear function of the two worker-types' payoffs (u_L, u_H) .

(c) Suppose firm 1 will announce a randomly generated menu $((1, \tilde{w}_L), (\tilde{q}_H, \tilde{w}_H))$ such that: \tilde{w}_L is drawn from a uniform distribution over the interval from 3 to 4.6 (from $g(\theta_L)$ to $Eg(\tilde{\theta})$), $\tilde{w}_L - 2 = \tilde{w}_H - 2\tilde{q}_H$ (so the low types would be indifferent between the two contracts), and $0.2(3 - \tilde{w}_L) + 0.8(5\tilde{q}_H - \tilde{w}_H) = 0$ (so the menu applied to all workers would yield expected profit 0). With these two equations, \tilde{q}_H and \tilde{w}_H depend linearly on \tilde{w}_L . Show that the resulting menu is incentive compatible. Then show that the two types' expected payoffs under such a random menu $\tilde{u}_L = \tilde{w}_L - 2$ and $\tilde{u}_H = \tilde{w}_H - 4\tilde{q}_H$ are also uniform random variables over certain intervals. Show the formulas for the cumulative probabilities $F_L(u) = P(\tilde{u}_L \leq u)$ and $F_H(u) = P(\tilde{u}_H \leq u)$.

(d) Given that firm 1 is behaving according to the randomized strategy in (c), show that firm 2 could get zero expected profit by choosing any incentive-compatible menu $((1, w_L), (q_H, w_H))$ that satisfies the conditions $w_L - 2 = w_H - 2q_H$, $0.2(3 - w_L) + 0.8(5q_H - w_H) = 0$, and $3 \leq w_L \leq 4.6$.

(e) Now suppose that firm 2 considers any incentive-compatible menu $((1, w_L), (q_H, w_H))$ in which the H|L incentive constraint is binding. Using your formulas from (b) and (c), show that firm 2's expected profit per worker $0.2 F_L(u_L) y_L + 0.8 F_H(u_H) y_H$ cannot be positive.