Dynamic Principal-Agent Problem

Gonzalo Cisternas
MIT

Today

- Two-Period Model
  - Rogerson (ECMA 1985) and Chiappori et al. (EER, 1994)
- Infinite Horizon and Recursive Formulation in Discrete Time
  - Spear and Srivastava (REStud 1987); Phelan and Townsend (REStud, 1991)
- Continuous Time and Brownian Noise
  - Sannikov (REStud 2008)

Two-Period Model

Model

- Risk-neutral principal and risk-averse agent
- Output can takes n possible values $q_1 < q_2 < ... < q_n$
- In each period $t \in \{ 1, 2\}$
  - agent takes action $a \in A \subset \mathbb{R}$ compact at the beginning of the period
  - action induces a prob. distribution π over output in the same period

$$\pi(a) := (\pi_1(a), ...\pi_n(a)),$$

where $\pi_i(a) := \text{Prob}(q_i|a) > 0,\ i = 1, ..., n,$ and $\sum_{i=1}^{n} \pi_i(a) = 1$

- The agent's action affects only current output. Conditional on the action, output is i.i.d. across time
- The agent's action is hidden to the principal; output is public

Model: Cont'd

- At $t \in \{1,2\}$, after observing output, q_t, principal makes a payment I_t
- Agent's utility at $t \in \{1,2\}$

$$U(I_t, a_t) = u(I_t) - g(a_t)$$

where $u' > 0$, $u'' < 0$, and $g' > 0$, $g'' > 0$, $g(0) = 0$. If $A = \{a_L, ..., a_k, ...a_H\}$, $g(a_L) < ... < g(a_K) < ... < g(a_H)$

- $\delta \in (0,1]$ discount factor
- Agent cannot save nor borrow
- Agent has a reservation utility W_0

Contract

- A **contract** is a tuple

$$\mathcal{I} := ((I_1, ..., I_n), (I_{11}, ...I_{1n}), ..., (I_{n1}, ..., I_{nn}))$$

  where

  - I_i is the wage paid at the end of $t = 1$ if q_i was observed
  - I_{ij} is the wage paid at the end of $t = 2$ if the public history was (q_i, q_j)

- The ppal. chooses $\mathcal{I}$ and recommends an action $\vec{a} := (a_0, (a_1, ..., a_n))$ with $a_0 =$ action at $t = 1$, and $a_i =$ effort at $t = 2$ after q_i

The Problem

$$\max_{\mathcal{I},\vec{a}} \quad \sum_{i=1}^{n} \pi_i(a_0) \left[q_i - I_i + \delta \sum_{j=1}^{n} \pi_j(a_i)[q_j - I_{ij}] \right]$$

$$s.t. \quad (PC): \sum_{i=1}^{n} \pi_i(a_0) \left[u(I_i) - g(a_0) + \delta \sum_{j=1}^{n} \pi_j(a_i)[u(I_{ij}) - g(a_i)] \right] \geq W_0$$

$$(IC) \; : \; \vec{a} \in \arg\max_{\vec{\tilde{a}}} \sum_{i=1}^{n} \pi_i(\tilde{a}_0) \left[u(I_i) - g(\tilde{a}_0) + \delta \sum_{j=1}^{n} \pi_j(\tilde{a}_i)[u(I_{ij}) - g(\tilde{a}_i)] \right]$$

- **Today**:
  - look at binary-effort case: $A = \{a_L, a_H\}$, $g(a_L) = 0$ and $g(a_H) = \psi > 0$
  - principal wants to implement $\vec{a}_H$

Cost Minimization

$$\min_{\mathcal{I}} \quad \sum_{i=1}^{n} \pi_i(a_H) \left[I_i + \delta \sum_{j=1}^{n} \pi_j(a_H) I_{ij} \right]$$

$$s.t. \quad \sum_{i=1}^{n} \pi_i(a_H) \left[u(I_i) - \Psi + \delta \underbrace{\sum_{j=1}^{n} \pi_j(a_H)[u(I_{ij}) - \Psi]}_{W_i} \right] \geq W_0$$

$$\vec{a}_H \in \arg\max_{\vec{\tilde{a}}} \sum_{i=1}^{n} \pi_i(\tilde{a}_0) \left[u(I_i) - g(\tilde{a}_0) + \delta \sum_{j=1}^{n} \pi_j(\tilde{a}_i)[u(I_{ij}) - g(\tilde{a}_i)] \right]$$

- Solve via **backward induction**; histories: $\emptyset$ and q_i, $i = 1, ..., n$
- Break dynamic problem of choosing $\mathcal{I}$ into two static problems
  1. At $t = 2$: given "promised utility" W_i for $t = 2$, find $\vec{I}_{i\cdot}$ that implements a_H at minimum cost in that period
  2. At $t = 1$: find $(\vec{I}, \vec{W})$ that implements a_H at $t = 1$ given that W_i must be delivered after observing q_i

$T = 2$: Static Contract

- Treat W_i as a parameter

$$C_2(W_i) := \min_{(I_{i1},\dots,I_{in})} \sum_{j=1}^{n} \pi_j(a_H) I_{ij}$$

$$s.t. \quad (PC): \sum_{j=1}^{n} \pi_j(a_H) u(I_{ij}) - \psi \geq W_i$$

$$(IC): \sum_{j=1}^{n} \pi_j(a_H) u(I_{ij}) - \psi \geq \sum_{j=1}^{n} \pi_j(a_L) u(I_{ij})$$

- $C_2(W_i) =$ cost of delivering W_i utils to the agent at $t = 2$
- FOC: $\frac{1}{u'(I_{ij})} = \lambda_2 + \mu_2 \left[1 - \frac{\pi_j(a_L)}{\pi_j(a_H)}\right]$; λ_2, μ_2 lagrange multipliers of (PC) and (IC)
- Envelope: $C_2'(W_i) = \lambda_2 = \sum_{j=1}^{n} \frac{\pi_j(a_H)}{u'(I_{ij})}$

$T = 1$: Going Backwards

- The principal solves

$$C_1(\overline{W}) := \min_{(\vec{I},\vec{W})} \sum_{i=1}^{n} \pi_i(a_H)[I_i + \delta C_2(W_i)]$$

$$s.t. \quad \sum_{i=1}^{n} \pi_i(a_H)[u(I_i) + \delta W_i] - \psi \geq \overline{W}$$

$$\sum_{i=1}^{n} \pi_i(a_H)[u(I_i) + \delta W_i] - \psi \geq \sum_{i=1}^{n} \pi_i(a_L)[u(I_i) + \delta W_i]$$

- FOC:

$$(I_i) \quad : \quad -\pi_i(a_H)/u'(I_i) + \lambda_1 \pi_i(a_H) + \mu_1(\pi_i(a_H) - \pi_i(a_L)) = 0$$

$$(W_i) \quad : \quad -\pi_i(a_H)\delta C_2'(W_i) + \lambda_1 \delta \pi_i(a_H) + \mu_1 \delta(\pi_i(a_H) - \pi_i(a_L)) = 0$$

$$\Rightarrow \frac{1}{u'(I_i)} = C_2'(W_i) = \sum_{j=1}^{n} \frac{\pi_j(a_H)}{u'(I_{ij})}$$

Intertemporal Cost Smoothing

Proposition

Consider an optimal contract that implements a_H. Then, the inverse of the agent's marginal utility is a martingale:

$$\frac{1}{u'(I_i)} = \sum_{j=1}^{n} \frac{\pi_j(a_H)}{u'(I_{ij})} \left(= \mathbb{E}^{a_H}\left[\frac{1}{u'(I_{i\cdot})}\right]\right)$$

Cost smoothing: $h(v) := u^{-1}(v)$ wage that delivers v utils to the agent

$$h'(v_i) = \mathbb{E}^{a_H}[h'(v_{i\cdot})]$$

i.e., at the optimum, mg. cost of providing an extra util to the agent must the the same across periods along histories q_i, $i = 1, ..., n$

Implication on Wages: Contract Has Memory

- Result is more general: $1/u'(I_i) = \mathbb{E}^a[1/u'(I_{i\cdot})]$ for all $a > a_L$

Proposition

Consider an optimal contract that implements $a > a_L$. Suppose that $I_i \neq I_j$. Then, there exists $i' \neq j'$ such that $I_{ii'} \neq I_{jj'}$.

- **Intuition (dynamic optimization):** if it is optimal to let current utility depends non-trivially on an outcome, smoothing requires future wages to also depend non-trivially on that outcome
- **Economic intuition?**

Spreading Out Risk to Reduce Cost of Incentive Provision

- Recall that

$$\frac{1}{u'(I_i)} = C_2'(W_i)$$

- But $C_2(\cdot)$ is convex:
  - Let $v_{ij} := u(I_{ij})$ and $h := u^{-1}$, objective becomes $\sum_j \pi_j(a_H)h(v_{ij})$ and constraints are linear in v_{ij}'s; choose v_{ij}'s instead of I_{ij}'s
  - If $\vec{v}_{i\cdot}$ is optimal given W_i, $i = 1, 2$, $\lambda\vec{v}_{1\cdot} + (1-\lambda)\vec{v}_{2\cdot}$ is feasible given $\lambda W_1 + (1-\lambda)W_2$
  - $C(\lambda W_1 + (1-\lambda)W_2) \leq \sum_j \pi_j(a_H)h(\lambda v_{1j} + (1-\lambda)v_{2j}) < \lambda C(W_1) + (1-\lambda)C(W_2)$ by strict convexity of h

- Thus, $I_1 > I_2 \Leftrightarrow W_1 > W_2$; i.e., rewards distributed across time

- Static problem $\Rightarrow$ cost of incentive provision comes from introducing risk; dynamic setting $\Rightarrow$ optimal to spread risk over time

Access to Credit

Proposition

Suppose the optimal no-credit contract is in place and that after the outcome of $t = 1$ the agent is suddenly allowed to borrow or save. Then, the agent will save and never borrow

Proof: Optimal $t = 1$ consumption after q_i is given by $I_i - s^*$ where

$$u'(I_i - s^*) - \sum_{i=1}^{n} \pi_j(a_i) u'(I_{ij} + s^*) = 0$$

But $1/u'(I_i) = \sum_{j=1}^{n} \pi_j(a_i)/u'(I_{ij}) > 1/\sum_{j=1}^{n} \pi_j(a_i)u'(I_{ij})$ ($1/x$ is convex). Thus, $\sum_{j=1}^{n} \pi_j(a_i)u'(I_{ij}) > u'(I_i) \Rightarrow s^* > 0$.

Access to Credit: Intuition

- Memory result: intertemp. distribution of incentives - hence, consumption - **across** different output realizations (i.e., across different public histories)
  - When ppal is risk averse, her marginal utility appears in the numerator $\Rightarrow$ really a **risk-sharing** result
- Savings result is about how consumption is distributed over time **along** a public history
  - Wages are too **frontloaded**: agent would like to save and thus delay consumption to tomorrow
- **Intuition:** lowering W_i makes t=2 marginal utility more sensitive to small changes in payments $\Rightarrow$ cheaper to provide incentives

Rogerson (1985): Summary

1. **Memory** $\Rightarrow$ compensation can be a complex function of past outputs
2. Inverse Euler result is quite **general**
   - holds for any compact set of actions A (see Rogerson's proof)
3. Inverse of marginal utility evolving as a **martingale** is a weak restriction on wages
   - Many wage processes can satisfy that property
4. No exploration of optimal a; no optimal contract
5. **Savings**: the agent would like to save if allowed

- In fully dynamic settings one would expect (1), (2) and (5) to hold. To nail down (3) and (4) optimal contract must be solved
  - Next: focus on (1) assuming no access to credit (i.e., ignore (5))

Infinite Horizon:

Memory and Continuation Value as State Variable

The Problem

- Let $h^t := (q_1, ..., q_{t-1})$: public history at the beginning of period t
- Choose $I := (I_t)_{t \geq 0}$ and $a := (a_t)_{t \geq 0}$ with $I_t = I(h^t, q_t)$ and $a_t = a(h^t)$ solving

$$\max_{(I,a)} \quad \mathbb{E}^a \left[\sum_{t=0}^{\infty} \delta^t (q_t - I_t) \right]$$

$$s.t. \quad (PC): \ \mathbb{E}^a \left[\sum_{t=0}^{\infty} \delta^t (u(I_t) - g(a_t)) \right] \geq \overline{W}$$

$$(IC) \ : \ \mathbb{E}^a \left[\sum_{t=0}^{\infty} \delta^t (u(I_t) - g(a_t)) \right] \geq \mathbb{E}^{\tilde{a}} \left[\sum_{t=0}^{\infty} \delta^t (u(I_t) - g(\tilde{a}_t)) \right], \ \forall \tilde{a}$$

where $\mathbb{E}^a[\cdot] :=$ prob. dist. over paths $(q_t)_{t \in \mathbb{N}}$ when $(a_t)_{t \geq 0}$ is followed

- But for $T = 2$ we only need to know W_i going forward
    - Actions affect current output only and draws are conditionally i.i.d.

Recursive Formulation

Proposition (Spear and Srivastava)

$(I_t, a_t)_{t\geq 0}$ is an optimal contract given $W_0 = W$ if and only if there is $\mathcal{W} := [\underline{W}, \overline{W}] \subseteq \mathbb{R}$, and functions $F : \mathcal{W} \to \mathbb{R}$, $a : \mathcal{W} \to A$, and $U, I : \mathcal{W} \times \mathbb{R} \to \mathcal{W}$ s.t. for $W \in \mathcal{W}$:

- *$I(W, \cdot)$ and $a(W)$ solve*

$$\max_{(J(\cdot),a)} \quad \mathbb{E}_q^a \left[q - J(q) + \delta F(U(W,q)) \right]$$

$$s.t. \quad (PK) \ : \ \mathbb{E}_q^a \left[u(J(q)) - g(a) + \delta U(W,q) \right] = W$$

$$(IC_L) \ : \ a \in \arg\max_{\tilde{a}} \mathbb{E}_q^{\tilde{a}} \left[u(J(q)) - g(\tilde{a}) + \delta U(W,q) \right], \ \forall \tilde{a} \in A$$

- *$F(W)$ satisfies $F(W) = \mathbb{E}_q^{a(W)} \left[q - I(W,q) + \delta F(U(W,q)) \right]$, and*
- *$F(W)$ is the principal's profit when $W_0 = W$, $I_t(h^t) = I(W_t, q_t)$ and $a_t(h^t) = a(W_t)$, where $W_{t+1} = U(W_t, q_t)$, $W_0 = W$.*

Memory and Continuation Value

- **Advantages**:
  - Complex (non-Markov) problem reduced to a family of static problems
  - Memory is summarized in the value that W takes $\Leftrightarrow$ **recursive structure**
    - Cont. game at $t+1$ given $W_{t+1} = w$ identical to game at t with $W_t = w$
    - actions do not have persistence and conditionally i.i.d. output are crucial
  - Dynamic IC constraint replaced by static -i.e, **one-shot** - IC constraint
    - Recursive structure permits concatenation of such local IC constraints.

      Obs: Authors do not prove the 'if' (i.e., $\Leftarrow$) in the previous proposition
- **Disadvantages**
  - Static problem still complex: variational problem
  - With a continuum of actions, continuum of IC constraints
  - Computationally intensive ▸ Phelan and Townsend (1991) algorithm

Why Continuous Time?

1. As t progresses, ppal keeps track of W_t. Policy $U(\cdot,\cdot)$ is her choice and induces a law of motion $W_{t+1} = U(W_t, q_t)$

   Cont. time gives a tractable controlled dynamic for $(W_t)_{t\geq 0}$

2. Spear and Srivastava

   2.1 Follow "first-order approach," i.e., replace continuum of static IC with

   $$\frac{\partial}{\partial a}\mathbb{E}_q^a\left[u(I(W,q)) + \delta U(W,q)\right]\Big|_{a=a(W)} = g'(a(W))$$

   but the validity of FOA is not shown

   2.2 One-shot deviation principle is not proven (should hold under appropriate integrability conditions)

$\Rightarrow$ Cont. time the agent's static problem is concave

and one-shot deviation principle is easier to prove

Continuous Time:

Sannikov (REStud, 2008)

Model

- $(X_t)_{t\geq 0}$ cumulative output process

$$X_t = \int_0^t a_s ds + \sigma Z_t, \quad \text{where} \quad Z_t \sim \mathcal{N}(0, \sigma^2 t)$$

- Intuition: relevant "period" is $[t, t+dt)$
  1. At the beginning of the period agent takes action a_t
  2. Realized output, dX_t, given by linear-additive prod. function

$$q_t := dX_t = a_t dt + \sigma dZ_t \text{ where } dZ_t \sim \mathcal{N}(0, \sigma^2 dt)$$

- Agent's flow utility over $[t, t+dt)$ is given by $[u(I_t) - g(a_t)]dt$
  - I_t depends on $(X_s : 0 \leq s < t)$ (or $\leq$; irrelevant in continuous time)
  - $u : \mathbb{R}_+ \to \mathbb{R}_+$, $u(0) = 0$, $u' > 0$, $u'' < 0$ and $g : A \to \mathbb{R}_+$ continuous, increasing and convex with $g(0) = 0$ (Technical: $u'(c) \to 0$ as $c \to \infty$ and there is γ_0 s.t. $g(a) \geq \gamma_0 a$ for all $a \in A$)

The Principal's Problem

- Expected profits

$$\mathbb{E}^a\left[\int_0^\infty e^{-rs}(dX_s - I_s ds)\right] = \mathbb{E}^a\left[\int_0^\infty e^{-rs}(a_s - I_s)ds\right]$$

- Choose $I := (I_t)_{t\geq 0}$ and recommendation $a := (a_t)_{t\geq 0}$ that solve

$$\begin{aligned}
\max_{(I,a)} \quad & \mathbb{E}^a\left[r\int_0^\infty e^{-rs}(a_s - I_s)ds\right] \\
s.t. \quad & (PC):\ \mathbb{E}^a\left[r\int_0^\infty e^{-rs}[u(I_s) - g(a_s)]ds\right] \geq W_0 \\
& (IC):\ a \in \arg\max_{\tilde{a}}\ \mathbb{E}^{\tilde{a}}\left[r\int_0^\infty e^{-rs}[u(I_s) - g(\tilde{a}_s)]ds\right]
\end{aligned}$$

- Let $W_t := \mathbb{E}^a\left[r\int_t^\infty e^{-r(s-t)}[u(I_s) - g(a_s)]ds \middle| (X_0 : 0 \leq s < t)\right]$

Towards a Law of Motion of W_t

- Y random variable; $M_t = \mathbb{E}_t[Y]$ is a martingale; Why?
- Let $Y = r\int_0^\infty e^{-rs}[u(I_s) - g(a_s)]ds$. The following is a martingale

$$G_t = \mathbb{E}_t^a\Big[\underbrace{r\int_0^\infty e^{-rs}[u(I_s) - g(a_s)]ds}_{Y:=}\Big] = \int_0^t e^{-rs}[u(I_s) - g(a_s)]ds + e^{-rt}W_t$$

- **Martingale Representation Thm**: every martingale w.r.t. the Brownian information is an integral against a Brownian motion

$$G_t = r\sigma\int_0^t e^{-rs}Y_s dZ_s \text{ some } (Y_t)_{t\geq 0} \Rightarrow \text{“}dG_t = r\sigma e^{-rt}Y_t dZ_t\text{”}$$

- Ito's rule: $dG_t = e^{-rt}r[u(I_t) - g(a_t)]dt - re^{-rt}W_t dt + e^{-rt}dW_t$

$$\boxed{dW_t = r(W_t - [u(I_t) - g(a_t)])dt + rY_t\sigma dZ_t}$$

Incentive Compatibility

- What's the Brownian motion? From the principal's perspective

$$\begin{aligned} \sigma dZ_t &= dX_t - a_t dt \\ \Rightarrow dW_t &= \underbrace{(rW_t - [u(I_t) - g(a_t)])}_{\text{promise keeping}} dt + r \cdot \underbrace{Y_t}_{\text{risk/incentives}} \cdot \underbrace{[dX_t - a_t dt]}_{\text{agent controls } dX_t} \end{aligned}$$

- Analogy: RHS is $U(W_t, q_t)$; choosing $U \Rightarrow$ choosing (I_t, Y_t)
- Consider **one-shot** deviation from a_t: $\tilde{a}_t$ over $[t, t + dt)$ and a_s, $s > t$
  - Output changes by $(\tilde{a}_t - a_t)dt \Rightarrow$ agent gains $rY_t(\tilde{a}_t - a_t)dt$
  - Extra cost? $g(\tilde{a}_t) - g(a_t)$
  - Agent won't deviate if: $rY_t a_t - g(a_t) \geq rY_t\tilde{a}_t - g(\tilde{a}_t)$
  - To implement a_t, ppal chooses Y_t s.t. $rY_t a_t - g(a_t) \geq rY_t a' - g(a') \ \forall a'$
- This is a necessary condition for IC (local IC), but sufficiency holds

Bellman Equation

- Risk is costly: $\gamma(a) := \min\{y \in [0,\infty) : a \in \arg\max_{a' \in A} \; ya' - g(a')\}$

$$dW_t = (rW_t - [u(I_t) - g(a_t)])dt + r\gamma(a) \cdot \underbrace{[dX_t - a_t dt]}_{=\sigma dZ_t \text{ as agent follows } a_t}$$

- Value to the principal $F(\cdot)$ then satisfies **ODE**

$$rF(W) = \sup_{a,I}\{a - I + r(W - u(I) + g(a))F'(W) + \frac{1}{2}r^2\gamma(a)^2\sigma^2 F''(W)\}$$

- Agent can always guarantee $W \geq 0$. Why? Boundary conditions
  - $F(0) = 0$: $W = 0 \Rightarrow I_t \equiv 0 \Rightarrow a_t \equiv 0 \Rightarrow F(0) = 0$
  - Ppal can always **retire** the agent: offer constant payment $u^{-1}(W)$ and no production occurs. This costs the principal $F_0(W) = -u^{-1}(W)$

- Look for F over the largest $\mathcal{W} := [0, W_{gp}]$ s.t. $F \geq F_0$ over $\mathcal{W}$ with

$$F(0) = 0, \; \underbrace{F(W_{gp}) = F_0(W_{gp})}_{\text{value matching}} \text{ and } \underbrace{F'(W_{gp}) = F_0'(W_{gp})}_{\text{smooth pasting}}$$

Sannikov's Verification Theorem

Theorem (**Sannikov**, 2008)

1. $\exists!$ *solution* $F \geq F_0$ *to the HJB equation satisfying the boundary conditions for some* $W_{gp} \geq 0$. *Such* F *is concave and it corresponds to the principal's value function.*

2. *For* $W^o \in (0, W_{gp})$, *and as long as* $W_t \in (0, W_{gp})$, $I_t = I(W_t)$ *and* $a_t = a(W_t)$, *with* $a(\cdot)$ *and* $I(\cdot)$ *the maximizers in the HJB, and*

$$dW_t = r(W_t - u(I_t) + g(a_t))dt + r\gamma(a_t)[dX_t - a_t dt],\ t > 0,\ W_0 = W^o$$

with $\gamma(a_t) > 0$, *until the time* τ *at which* W_τ *hits* 0 *or* W_{gp}. *In that case, agent gets constant consumption equal to* $-F_0(W_\tau)$ *thereafter.*

Numerical Examples

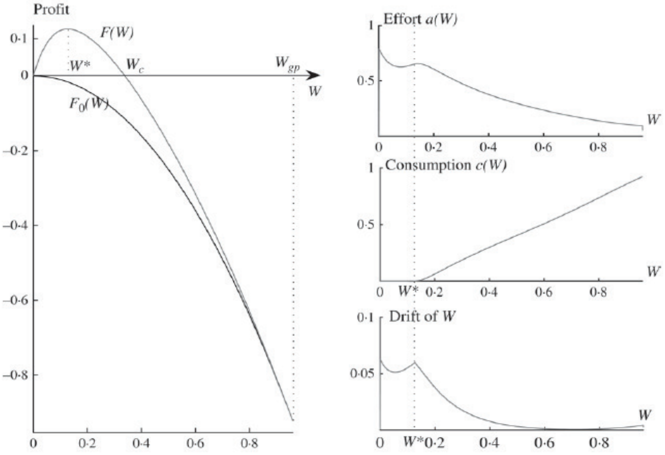

FIGURE 1

Function F for $u(c) = \sqrt{c}$, $h(a) = 0{\cdot}5a^2 + 0.4a$, $r = 0{\cdot}1$ and $\sigma = 1$. Point W^* is the maximum of F

Reference: Sannikov (2008)

Bibliography

1. Chiappori, Macho, Rey, Salanie (1994): "Repeated Moral Hazard: The Role of Memory, Commitment, and the Access to Credit Markets," *European Economic Review*

2. Phelan and Townsend (1991): "Computing Multi-period Information-Constrained Optima," *Review of Economic Studies*

3. Rogerson (1985): "Repeated Moral Hazard," *Econometrica*

4. Sannikov (2008): "A Continuous-Time Version of the Principal-Agent Problem," *Review of Economic Studies*

5. Spear and Srivastava (1987): "Computing Multiperiod Constrained Optima," *Review of Economic Studies*

Computation: Phelan and Townsend (Restud, 1991)

$$F_T(w) = \max_{\Pi(a,q,c,w')} \frac{1}{1+r} \sum_{A\times Q\times C\times W_{T-1}} \Pi(a,q,c,w')(r(q-c) + F_{T-1}(w'))$$

$$\sum_{A\times Q\times C\times W_{T-1}} \Pi(a,q,c,w') = 1$$

The probabilities add up to 1

$$\forall a,q,c,w',\ \Pi(a,q,c,w') \geq 0$$

The probabilities are nonnegative

$$\forall \hat{a},\hat{q},\hat{c} \sum_{W_{T-1}} \Pi(\hat{a},\hat{q},\hat{c},w') = P(\hat{q} \mid \hat{a}) \sum_{Q\times C\times W_{T-1}} \Pi(\hat{a},q,c,w')$$

The joint probability distribution over a, q and c is consistent with the conditional probabilities P(q|a) with which output q is realized given effort a[1]

$$\sum_{A\times Q\times C\times W_{T-1}} w'\,\Pi(a,q,c,w') - w = r \sum_{A\times Q\times C\times W_{T-1}} (w - u(c) + h(a))\Pi(a,q,c,w') \quad \text{(PK)}$$

$$\forall a,\hat{a},\ \sum_{Q\times C\times W_{T-1}} (w' - rh(a))\Pi(a,q,c,w') \geq \sum_{Q\times C\times W_{T-1}} (w' - rh(\hat{a}))\frac{P(q \mid \hat{a})}{P(q \mid a)}\Pi(a,q,c,w') \quad \text{(IC)}$$

▸ Back to Memory and Continuation Value