Decentralized control of stochastic multi-agent service system

J. George Shanthikumar
Purdue University

Joint work with Huaning Cai, Peng Li and Andrew Lim

Many problems can be thought of as stochastic dynamic resource allocation problems:

- Airline alliances
- Investment firms with multiple trading desks
- Large scale service systems, call centers
- ...

Objective: operate the integrated system efficiently.

Centralized decision making is often not possible:

- different groups specialize in different tasks.

- relevant information is not/can not be shared

- integrating IT systems across different groups is hard.

$\Rightarrow$ decentralized agents

Decentralized control involves many things:

- different agents have knowledge and control of different parts of the system

- different agents can observe different subsets of the state e.g. Witsenhausen (1968), Ho (1980), Rantzer (2009), Bertsekas (2007), Adelman & Mersereau (2008)

- different agents have different incentives

Key issues:

- Each decision maker maximizes its own objectives conditional on the available information.

- In general: local optimality $\neq$ global optimality

How do we efficiently manage such multi-agent systems?

Overview

- Centralized formulation of the pricing-service problem
- Decentralized model
  - Stochastic models for pricing and service agents
  - Coordination mechanism
  - Pricing agent's problem
  - Service agent's problem
- Main results
  - Coordinating through transfer prices
  - Computing transfer prices with info. sharing constraints
- Example

Problem description

- arrivals occur randomly at a rate determined by the prices dynamically controlled by (multiple) pricing agents.

- service times are random and determined by the dynamically controlled service rate

- net profits = revenue from arrivals - service costs

Standard approach: centralized model

$$\begin{cases} \max\limits_{p(\cdot),\,\mu(\cdot)} E\left\{\int_0^T \sum_{j=1}^m p_j(t^-)dN_j(T) - \int_0^T c(x(t),\,\mu(t))dt - h(x(T))\right\} \\ dx(t) = \sum_{j=1}^m dN_j(t) - dJ(t),\ x(0) = x_0 \\ 0 \le x(t) \le C,\ \forall\, t \in [0,\, T], \\ N_i(t) \sim \lambda_i(p_i(t^-)),\ J(t) \sim \mu(t^-). \end{cases}$$

Assumes there is and all-knowing dictator jointly controlling all pricing and service decisions.

It is possible to write down the dynamic programming equations

$$\begin{cases} 0 = V_t(t, x) + \max_{p, \mu} \Big\{ \sum_i \lambda_i(p_i) \Big[p_i - \Big(V(t, x) - V(t, x+1) \Big) \Big] \\ \quad -c(x, \mu) + \mu \Big(V(t, x-1) - V(t, x) \Big) \Big\} \\ V(t, x) = -h(x) \end{cases}$$

but this is irrelevant if no single agent knows all demand models $\lambda_i(p_i)$ and the cost structure $(c(x, \mu), h(x))$ and has the power to jointly control all decision variables.

Decentralized control

Pricing agent i controls $p_i(t)$ and service agent controls $\mu(t)$.

Integrated system profits & dynamics:

$$\begin{cases} E\left\{\int_0^T \sum_{j=1}^{m} p_j(t^-)dN_j(T) - \int_0^T c(x(t),\, \mu(t))dt - h(x(T))\right\} \\ dx(t) = \sum_{j=1}^{m} dN_j(t) - dJ(t),\ x(0) = x_0 \\ 0 \le x(t) \le C,\ \forall\, t \in [0,\, T], \\ N_i(t) \sim \lambda_i(p_i(t^-)),\ J(t) \sim \mu(t^-). \end{cases}$$

Coupling comes through the state variable $x(t)$.

Key issue: what do pricing/service agents know and how do they make decisions?

Pricing agent

- controls $p_i(t)$ and knows demand function $\lambda_i(p_i)$.

- may not know the number of other agents, let alone their demand functions $\lambda_j(p_j)$, their pricing policies $p_j(t, x)$ or the aggregate arrival rate.

- may not know cost functions $(c(x, \mu), h(x))$ for service agent or the service policy $\mu(t)$.

- observes $x(t)$ but may not know its true dynamics.

Pricing agent's model

$$dx(t) = dN_i(t) - d\tilde{N}_i(t) - d\tilde{J}_i(t)$$

where

- $N_i(t) \sim \lambda_i(p_i)$ is the arrival process controlled by agent i (accurately specified)

- $\tilde{N}_i(t) \sim \tilde{\lambda}_i(t)$ is agent i's model for arrivals brought by other agents (mis-specified)

- $\tilde{J}_i(t) \sim \tilde{\mu}_i(t)$ is agent i's model for the service system (mis-specified)

Service agent

Observes $x(t)$ and assumes dynamics

$$dx(t) = d\tilde{N}_0(t) - dJ_0(t)$$

where

- $\tilde{N}_0(t) \sim \tilde{\lambda}_0(t)$ is the service agent's aggregate arrival process model; (mis-specified)

- $J_0(t) \sim \mu(t)$ is the service process (accurately specified).

Local optimality does not give global optimality

Key observation: no choice of $\tilde{\lambda}_i(t)$ or $\tilde{\mu}_i(t)$ result in a centrally optimal choice of $p_i(t)$.

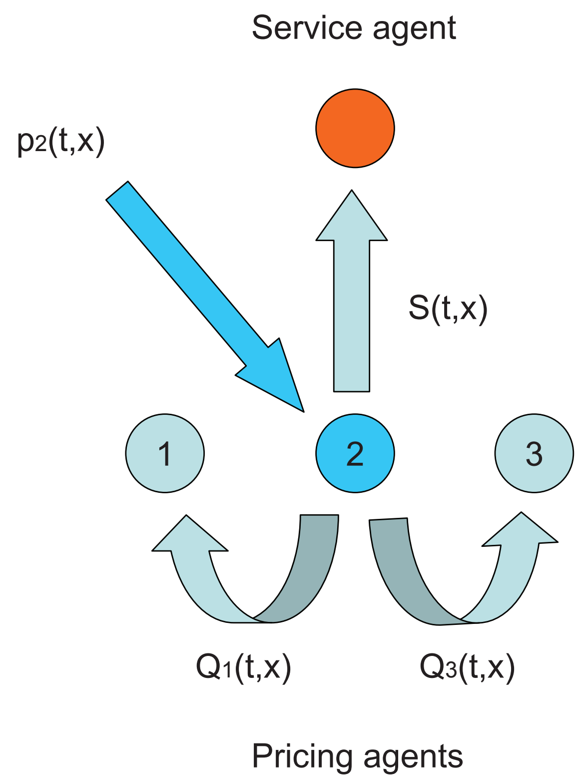
e.g. Pricing agent i's problem:

$$\begin{cases} \max_{p_i} E \int_0^T p_i(t^-) dN_i(t) = E \int_0^T p_i(t^-) \lambda_i(p_i(t^-)) dt \\ dx(t) = dN_i(t) + d\tilde{N}_i(t) - d\tilde{J}_i(t) \\ 0 \leq x(t) \leq C \\ N_i(t) \sim \lambda_i(p_i(t^-)),\ \tilde{N}_i(t) \sim \tilde{\lambda}_i(t),\ \tilde{J}_i(t) \sim \tilde{\mu}_i(t) \end{cases}$$

Trivial case: when $C = \infty$ the pricing agent maximizes revenue rate $p_i \lambda_i(p_i)$ which is not optimal for the system.

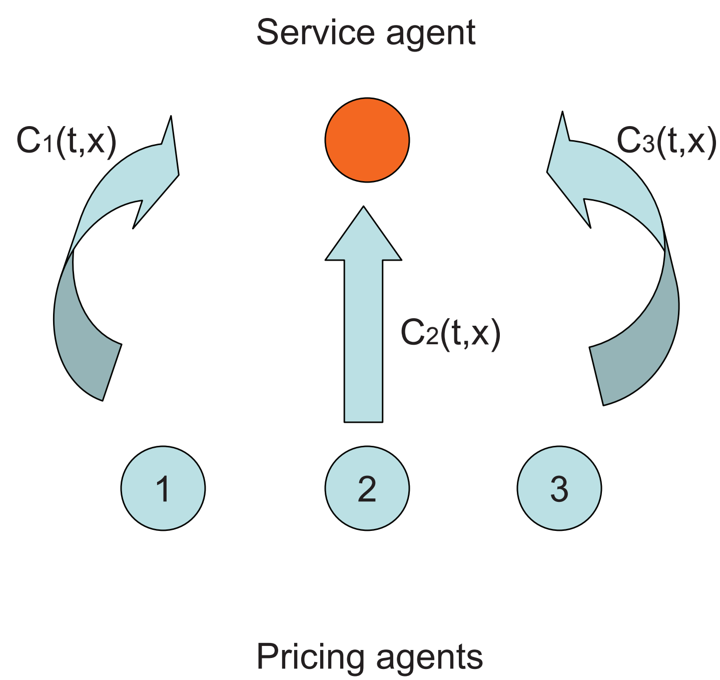
Transfer contracts and revenue sharing

Revenue transfer when pricing agent 2 makes a sale



$$\text{Income for Agent 2} = p_2(t,\, x(t^-)) - \sum_{j \neq 2} Q_j(t,\, x(t^-)) - S(t,\, x(t^-))$$

Revenue transfer when service agent clears a customer



$$\text{Income for service agent} = \sum_{i=1}^{m} C_i(t,\, x(t^-))$$

Revenue sharing/transfer prices

When pricing agent i brings in a customer and changes state from $x \to x+1$ at time t:

- pay $Q_j(t, x)$ to pricing agent j;
- pay $S(t, x)$ to service agent.

$$\Rightarrow \text{total cost:} \quad R_i(t, x) = \sum_{j \neq i} Q_j(t, x) + S(t, x)$$

When service agent completes service and changes state from $x \to x-1$ at time t:

- receive $C_i(t, x)$ from pricing agent i.

$$\Rightarrow \text{total revenue:} \quad R_0(t, x) = \sum_{j} C_j(t, x(t^-))$$

Pricing agent's problem

Given transfer prices C_i, Q_i and

$$R_i(t, x) = \sum_{j \neq i} Q_j(t, x) + S(t, x)$$

pricing agent i solves

$$\begin{cases} \max_{p_i} E \int_0^T \left[p_i(t^-) - R_i(t, x(t^-)) \right] dN_i(t) \\ \quad + \int_0^T Q_i(t, x(t^-)) d\tilde{N}_i(t) - \int_0^T C_i(t, x(t-)) d\tilde{J}_i(t) \\ dx(t) = dN_i(t) + d\tilde{N}_i(t) - d\tilde{J}_i(t) \\ 0 \leq x(t) \leq C \\ \underbrace{N_i(t) \sim \lambda_i(p_i(t^-))}_{\text{accurate}}, \ \underbrace{\tilde{N}_i(t) \sim \tilde{\lambda}_i(t), \ \tilde{J}_i(t) \sim \tilde{\mu}_i(t)}_{\text{mis-specified}} \end{cases}$$

Service agent's problem

Given transfer prices S and

$$R_0(t, x) = \sum_j C_j(t, x(t^-))$$

service agent solves

$$\begin{cases} \min_{\mu(t^-)} E \int_0^T c(x(t), \mu(t))dt - \int_0^T S(t, x(t^-))d\tilde{N}_0(t) \\ \quad - \int_0^T R_0(t, x(t^-))dJ_0(t) + h(x(T)) \\ dx(t) = d\tilde{N}_0(t) - dJ_0(t), \; x(0) = x_0 \\ 0 \le x(t) \le C \\ \underbrace{\tilde{N}_0(t) \sim \tilde{\lambda}_0(t)}_{\text{mis-specified}}, \; \underbrace{J_0(t) \sim \mu(t)}_{\text{accurate}} \end{cases}$$

Weak duality

For any choice of transfer prices

Centralized (opt.) profit

$$\geq \sum_j \{\text{agent' } j\text{'s (max.) profit}\} - \text{service agents (min.) cost}$$

(RHS depends on the choice of transfer contracts)

Is it possible to choose transfer prices that achieve equality?

What is the impact of model mis-specification by the agents?

Main Result: General system

Consider the transfer contracts

$$\begin{aligned} Q_j(t, x) &= V_j(t, x) - V_j(t, x+1) \\ S(t, x) &= W(t, x+1) - W(t, x) \\ C_j(t, x) &= V_j(t, x-1) - V_j(t, x) \end{aligned}$$

where

$$\begin{cases} 0 = \dfrac{d}{dt} V_i(t, x) + \max\limits_{p_i} \lambda_i(p_i) \Big\{ p_i - \Big[V(t, x) - V(t, x+1) \Big] \Big\} \\ V_i(T, x) = 0 \end{cases}$$

and

$$\begin{cases} 0 = \dfrac{d}{dt} W(t, x) + \min\limits_{\mu} \Big\{ c(x, \mu) - \mu \Big[V(t, x) - V(t, x-1) \Big] \Big\} \\ W(T, x) = h(x) \end{cases}$$

Note: $V(t, x)$ is the value function for the centralized problem.

- Contracts coordinate the system;

- Strong duality:

  Centralized (opt.) profit

  $$= \sum_j \{\text{agent' } j\text{'s (max.) profit}\} - \text{service agents (min.) cost}$$

- Robustness to mis-specification:

  – $p_j^*(t)$ is independent of agent j's assumptions about the other pricing agents and the service agent

  – $\mu^*(t)$ is independent of the service agent's assumptions about the pricing agents.

Intuition: Optimal contracts

$W(t, x)$ is the value function for the service agent and $V_j(t, x)$ as the value function of pricing agent j under optimal TP's.

When pricing agent j makes a sale: $x \to x + 1$

- $W(t, x+1) - W(t, x) =$ service agent's expected cost increase
- $V_i(t, x + 1) - V_i(t, x) =$ exp. revenue decrease for p. agent i.

When service agent clears a customer: $x \to x - 1$

- $V_i(t, x - 1) - V_i(t, x) =$ value of extra space to p. agent i.

Under optimal transfer prices:

- when pricing agent i makes a sale:
  - compensate other pricing agents their loss in revenue

$$Q_j(t, x) = V_j(t, x) - V_j(t, x+1)$$

  - compensate the service agent the cost of processing another customer

$$S(t, x) = W(t, x+1) - W(t, x)$$

- When service clears a customer, each pricing agent pays the service agent its valuation of the extra resource (1 spot)

$$C_j(t, x) \;=\; V_j(t, x-1) - V_j(t, x)$$

Intuition: Robustness to mis-specification

If an agent is compensated his valuation of the shared resource, he/she is indifferent as to whether it is used by another agent

$\Rightarrow$ indifference to rate of consumption by others

$\Rightarrow$ robustness to mis-specification.

Additional comments

- TP's summarize what is important about the other agents: Details such as the number of other agents and dynamics are not required.

- TP's can be interpreted as Lagrange multipliers for a stochastic dynamic resource allocation problem. Convexity was not used to establish strong duality.

Key gap: How do we compute transfer prices without needing to solve the centralized problem?

i.e. how do we compute optimal contracts when there is no mechanism designer.

Computing the transfer prices

Recall: Contracts

$$\begin{cases} Q_j(t,\,x) &=& V_j(t,\,x) - V_j(t,\,x+1) \\ S(t,\,x) &=& W(t,\,x+1) - W(t,\,x) \\ C_j(t,\,x) &=& V_j(t,\,x-1) - V_j(t,\,x) \end{cases}$$

where

$$\begin{cases} 0 = \frac{d}{dt}V_i(t,\,x) + \max_{p_i} \lambda_i(p_i)\Big\{p_i - \Big[V(t,\,x) - V(t,\,x+1)\Big]\Big\} \\ V_i(T,\,x) = 0 \end{cases}$$

and

$$\begin{cases} 0 = \frac{d}{dt}W(t,\,x) + \min_{\mu} \Big\{c(x,\,\mu) - \mu\Big[V(t,\,x) - V(t,\,x-1)\Big]\Big\} \\ W(T,\,x) = h(x) \end{cases}$$

Problem: Requires knowledge of the value function for the centralized problem.

A fixed point representation

$$
\begin{aligned}
Q_j(t, x) &= V_j(t, x) - V_j(t, x+1) \\
S(t, x) &= W(t, x+1) - W(t, x) \\
C_j(t, x) &= V_j(t, x-1) - V_j(t, x)
\end{aligned}
$$

where

$$
\begin{cases}
0 = \dfrac{d}{dt} W(t, x) + \min\limits_{\mu} \Big\{ c(x, \mu) \\
\qquad -\mu \Big[\sum\limits_{j=1}^{m} C_j(t, x) + W(t, x) - W(t, x-1) \Big] \Big\} \\
\qquad + \tilde{\lambda}_0(t, x) \Big\{ W(t, x+1) - W(t, x) - S(t, x) \Big\} \\
W(T, x) = h(x)
\end{cases}
$$

A fixed point representation (cont')

$$
\begin{cases}
0 = \dfrac{d}{dt} V_i(t,\, x) + \max\limits_{p_i} \lambda_i(p_i) \Big\{ p_i - \Big[\sum\limits_{j \neq i} Q_j(t,\, x) + S(t,\, x) \Big] \\
\qquad - \Big[V_i(t,\, x) - V_i(t,\, x+1) \Big] \Big\} \\
\qquad + \tilde{\lambda}_i(t,\, x) \Big\{ Q_i(t,\, x) - \Big[V_i(t,\, x) - V_i(t,\, x+1) \Big] \Big\} \\
\qquad + \tilde{\mu}_i(t,\, x) \Big\{ V_i(t,\, x-1) - V_i(t,\, x) - C_i(t,\, x) \Big\} \\
V_i(T,\, x) = 0
\end{cases}
$$

Fixed point representation

Find $Q = [Q_1, \cdots, Q_m]$, $C = [C_1, \cdots, C_m]$ and S such that:

$$\begin{aligned} Q_j(t, x) &= V_j(t, x) - V_j(t, x+1) \\ S(t, x) &= W(t, x+1) - W(t, x) \\ (C_j(t, x) &= V_j(t, x-1) - V_j(t, x) = Q_j(t, x-1)) \end{aligned}$$

where

$$\begin{aligned} V_1(t, x) &\equiv \text{value function for pricing agent 1} \\ &\vdots \\ V_m(t, x) &\equiv \text{value function for pricing agent } m \\ W(t, x) &\equiv \text{value function for service agent} \end{aligned}$$

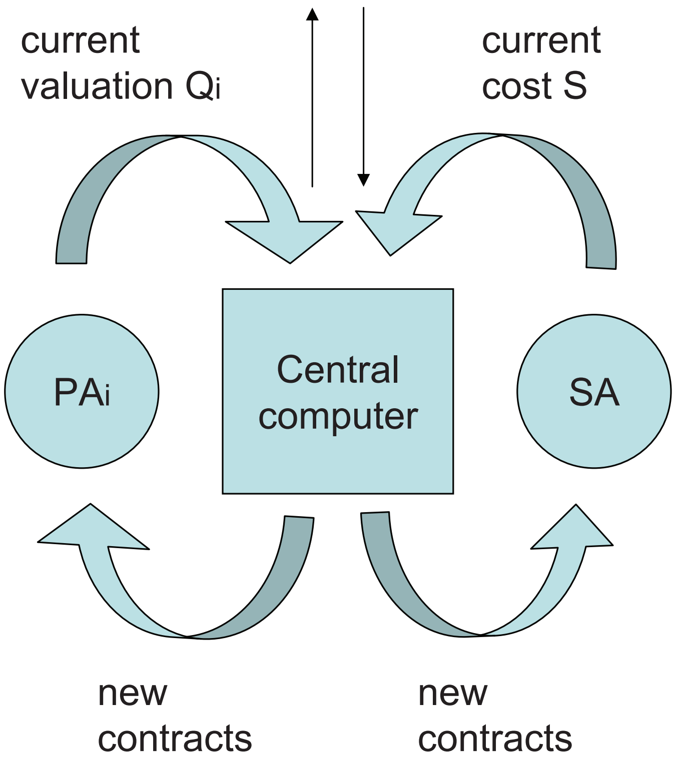
Natural idea: Iterate of Q, C and S.

Observations:

- The ODEs are DP equations for the pricing and service problems under contracts C, Q and S.

- The fixed point equation relates optimal contracts to value functions of the associated decentralized problems.

Interpretation as a negotiation



Main result

Under "standard" technical assumptions, the algorithm converges to the optimal transfer prices.

Comments:

- Assumes truth telling in reporting of valuations in each round.

- Iterative algorithm is analogous to updating dual variables. Convexity is not required for convergence or strong duality. (However, we assume each agent solves a DP at each iteration exactly)

- All that is communicated between agents are contracts constructed by aggregating transfer prices.
  (The number of agents and their models are not needed)

Example

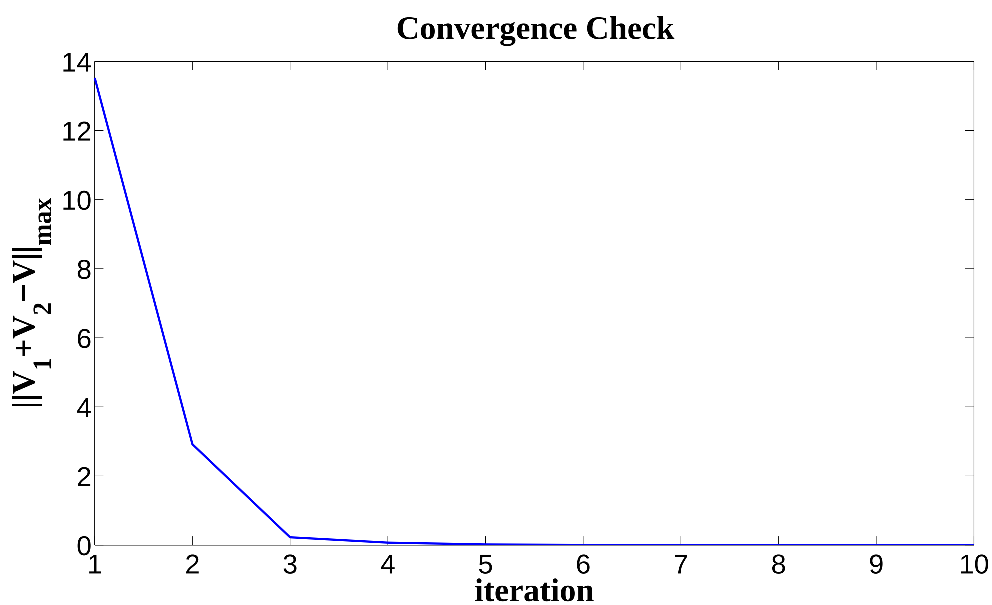
- Single pricing and service agent;

- pricing agent: demand rate $\lambda(p) = ae^{-bp}$;
  assumes $\tilde{\mu} \equiv 0$ (mis-specified);

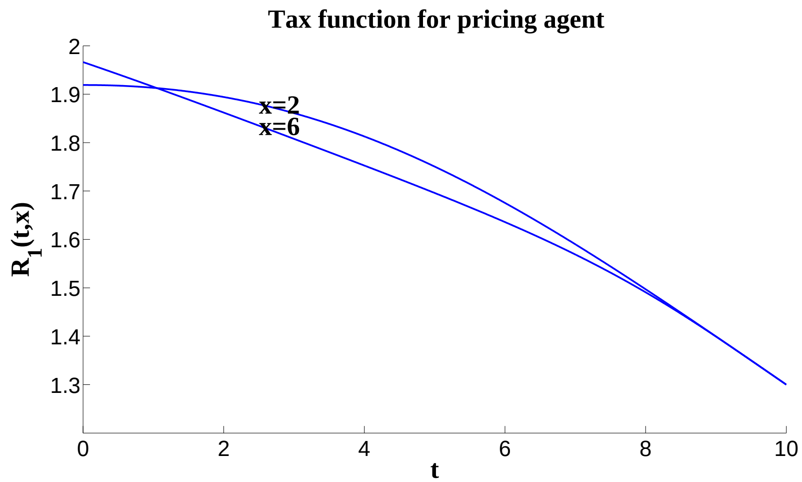
- service agent: costs $c(x, \mu) = kx + c(e^{d\mu} - 1)$ and $h(x) = sx$;
  assumes $\tilde{\lambda} \equiv 0$ (mis-specified);

$a = 10, b = 1, c = 1, d = 1, k = 0.1, s = 1.3, C = 10, T = 10$

Convergence check: $\|V - \{V_n^p - W_n\}\|_\infty$ as $n \to \infty$

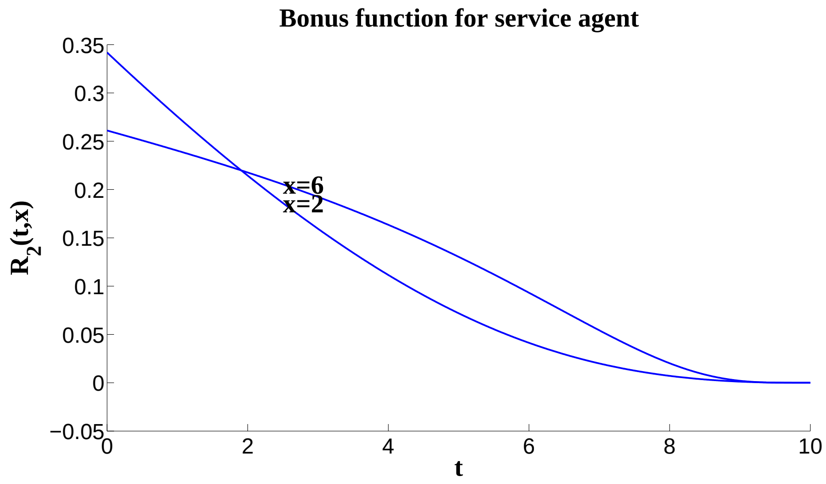


Optimal revenue sharing functions



When close to the terminal time $T = 10$ the pricing agent's "tax" converges to the cost of not serving the customer.

In this case, $h(x) = 1.3x$ (each unserved customer costs 1.3).



When close to terminal time $T = 10$, the service agent receives less reward for clearing a customer.

Intuition: The value of an extra space is of diminishing value as the selling time decreases.

Conclusion

- Coordination through transfer prices
  - TP's summarize what is important about other agents.
- Iterative algorithm for computing TP's that only require exchange of "valuations"
- Drawbacks:
  - Assumption of truth-telling when exchanging valuations
  - Each agent needs to solve a DP ($\Rightarrow$ approximations)
- Extensions:
  - heterogeneous risk attitudes
  - multi-agent portfolio choice problems
  - airline alliances