

Contracting Problems for Outsourcing

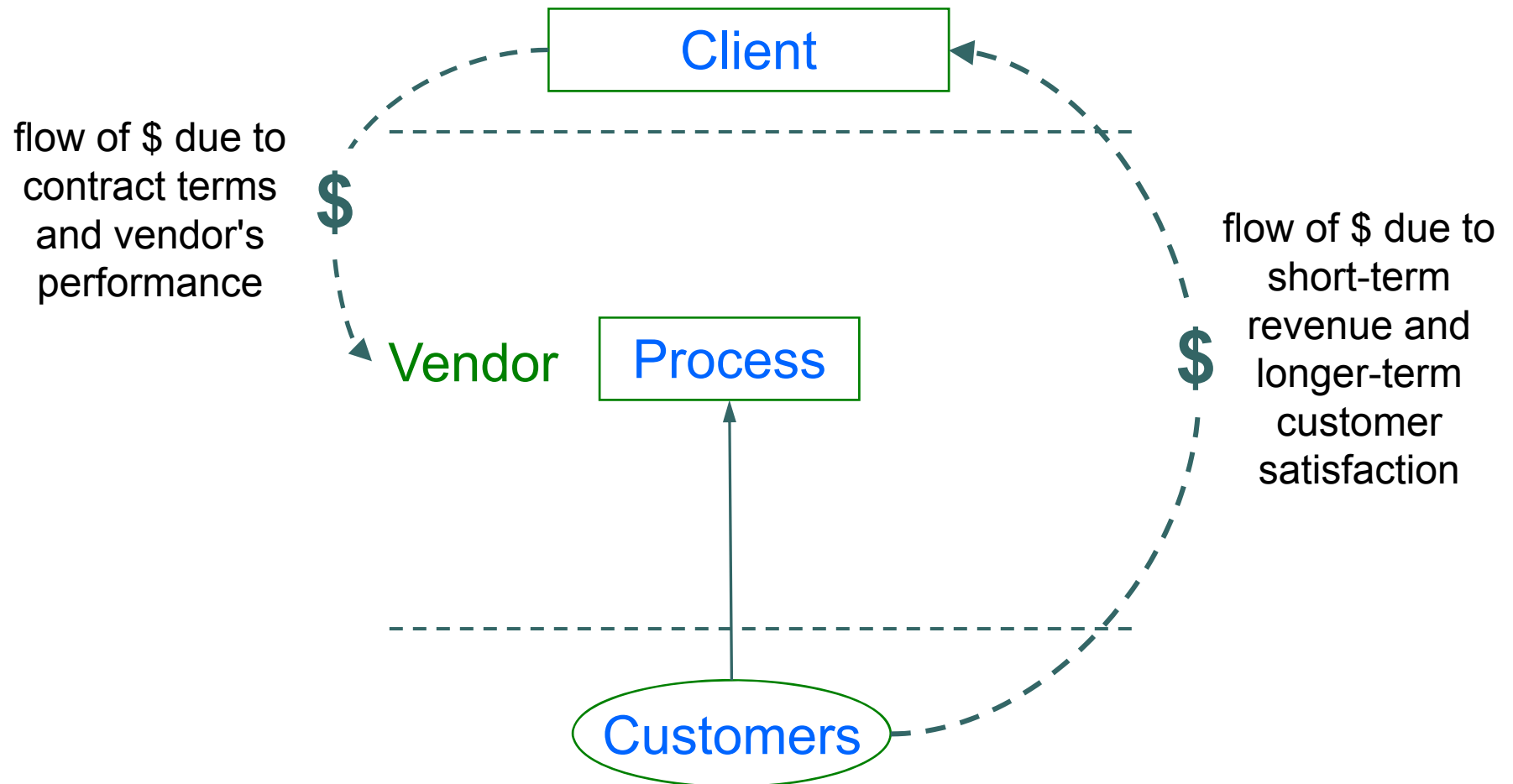
Sameer Hasija
School of Management, SUNY Binghamton

Edieal Pinker
Simon School, University of Rochester

Robert Shumsky
Tuck School, Dartmouth College

Wharton Contact Center Forum
February 29, 2008

A Service Supply Chain



Outsourcing Contracts for a Contact Center

Client	Service level agreement (SLA)	Financial Incentives
(A)	90% of calls answered within, 360 sec	Pay per call; SLA penalty; Monthly payment limit
(B)	70% of calls answered within 60 sec; 70% of calls resolved without escalation.	Pay per call; SLA penalty
(C)	80% of calls answered within 120 sec	Pay per min; SLA penalty; AHT penalty
(D)	80% of calls answered within 180 sec	Pay per min; SLA penalty; AHT penalty; Guaranteed base payment.
(E)	75% of calls answered within 20 sec	Pay per resolution; SLA penalty; Guaranteed base payment.
(F)	No more than 20% abandoned.	Pay per min; Abandonment probability penalty; AHT penalty

AHT: Average Handle Time SLA: Service Level Agreement

3

Customer Satisfaction (CSAT) monitored in all cases.

Research Overview

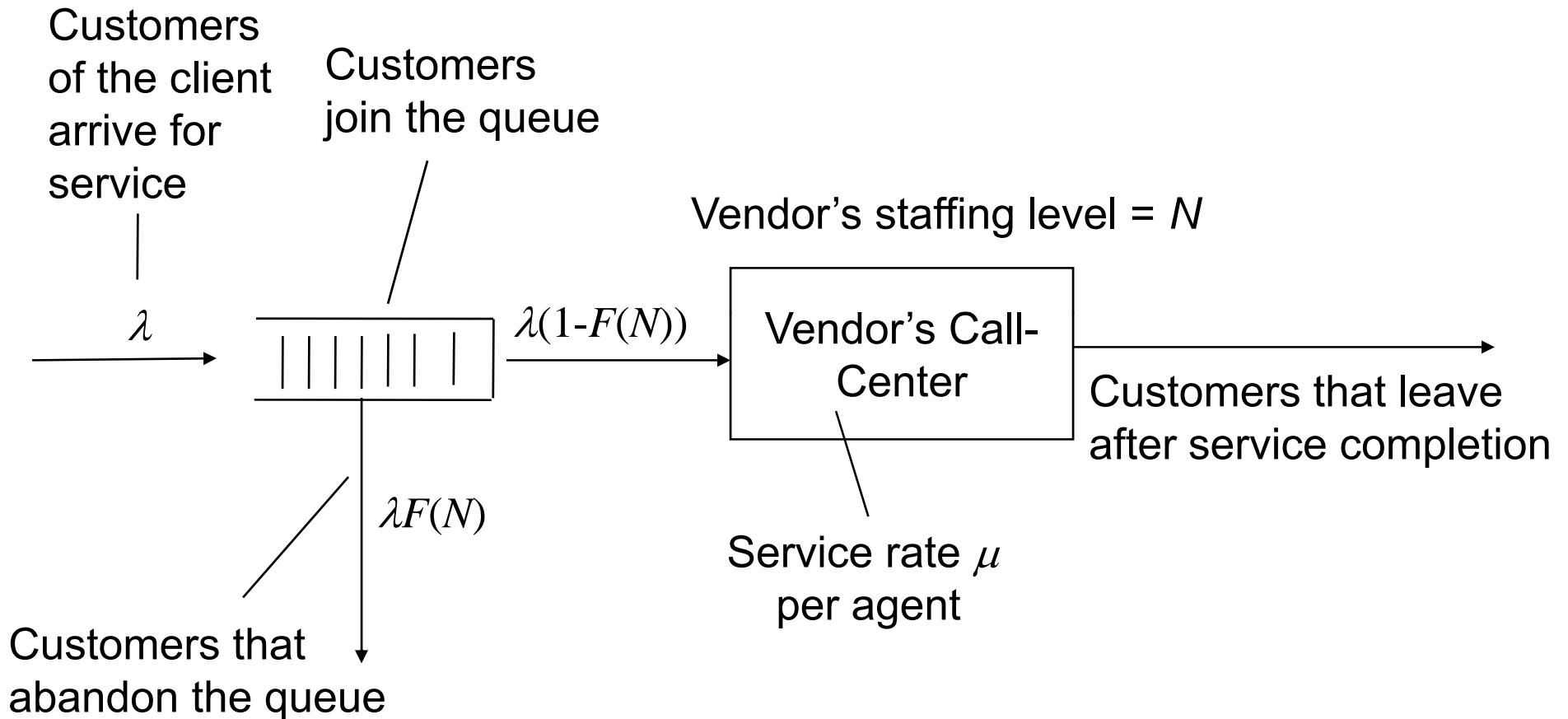
We assume that the client designs and offers the contracts.

- Which contracts...
 - *Coordinate* the chain (vendor chooses system-optimal capacity)?
 - Maximize client profits?
- How do the contracts perform when the client is uncertain about vendor productivity?

Closely Related Literature on Contact Center Outsourcing

- Aksin, de Vericourt and Karaesmen (2006): outsource the base or the peak?
- Allon and Federgruen (2006): when will multiple clients outsource to a common vendor?
- Ren and Zhou (2006): how do contracts influence vendor capacity decisions and resolution rates?
- Ren and Zhang (2008): contracting when the client is uncertain about the vendor's costs.

The Queueing Model for Our Contact Center



$F(N) = \text{probability of abandonment}$

We assume an $M/M/N+M$ queueing system.

The Client's Objective Function

revenue from
non-abandoning calls

cost of customers
waiting in queue

payment
to the vendor

$$\pi_c(N) = R \lambda (1 - F(N)) - P \frac{\lambda}{\theta} F(N) - T$$

N = Vendor's staffing level

R = Client revenue per completed service request

λ = Arrival rate

μ = Service rate

θ = Customer queue abandonment rate

P = Customer waiting cost per unit time

T = Transfer payment from client to vendor

Vendor and Chain Profit Functions

Vendor's profit:

$$\pi_v(N) = T - cN$$

payment to the vendor

staffing cost per unit time

Service Supply Chain Profit:

$$\pi_s(N) = R \lambda (1 - F(N)) - P \frac{\lambda}{\theta} F(N) - cN$$

revenue from non-abandoning calls

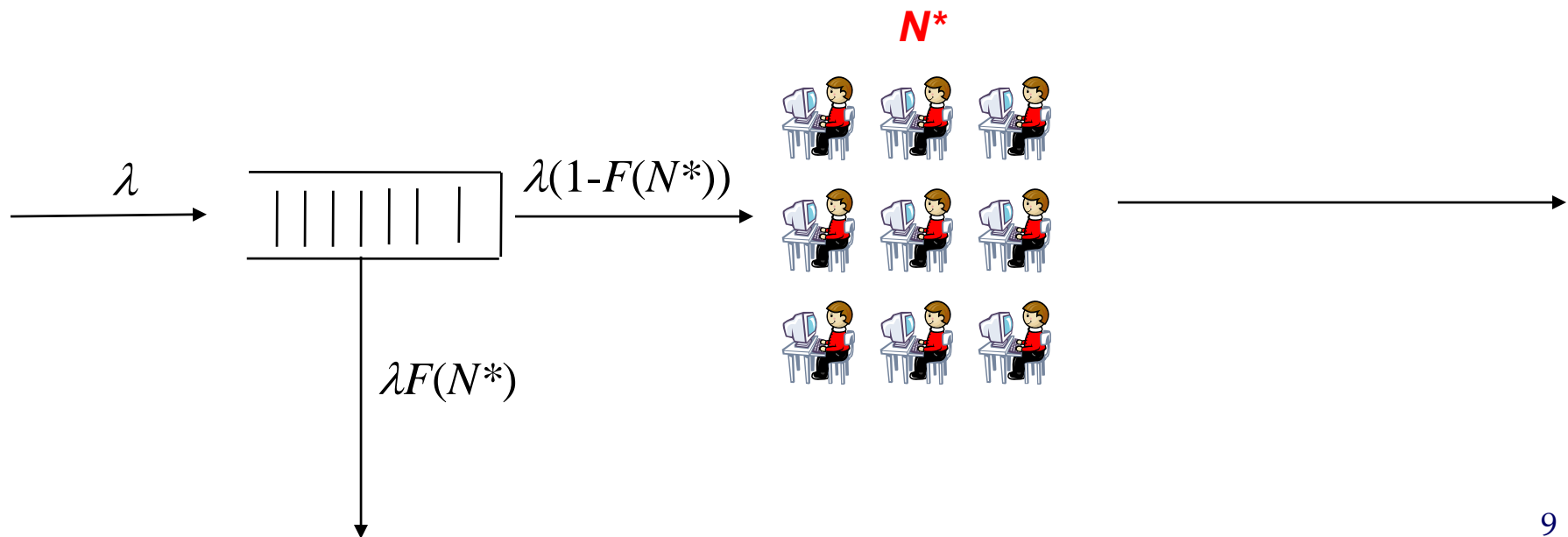
cost of customers waiting in queue

staffing cost per unit time

The 'First-Best' Capacity

N^* is the staffing level that optimally balances the service level (waiting time and abandonments) with the cost of agents.

N^* is the optimum over the *entire service supply chain*.



Contracts that Maximize Client Profit

The client would like to offer a contract under which,

1. The vendor *chooses* to staff $N_v^* = N^*$
(a coordinating contract).

2. The contract pays V to the vendor and leaves

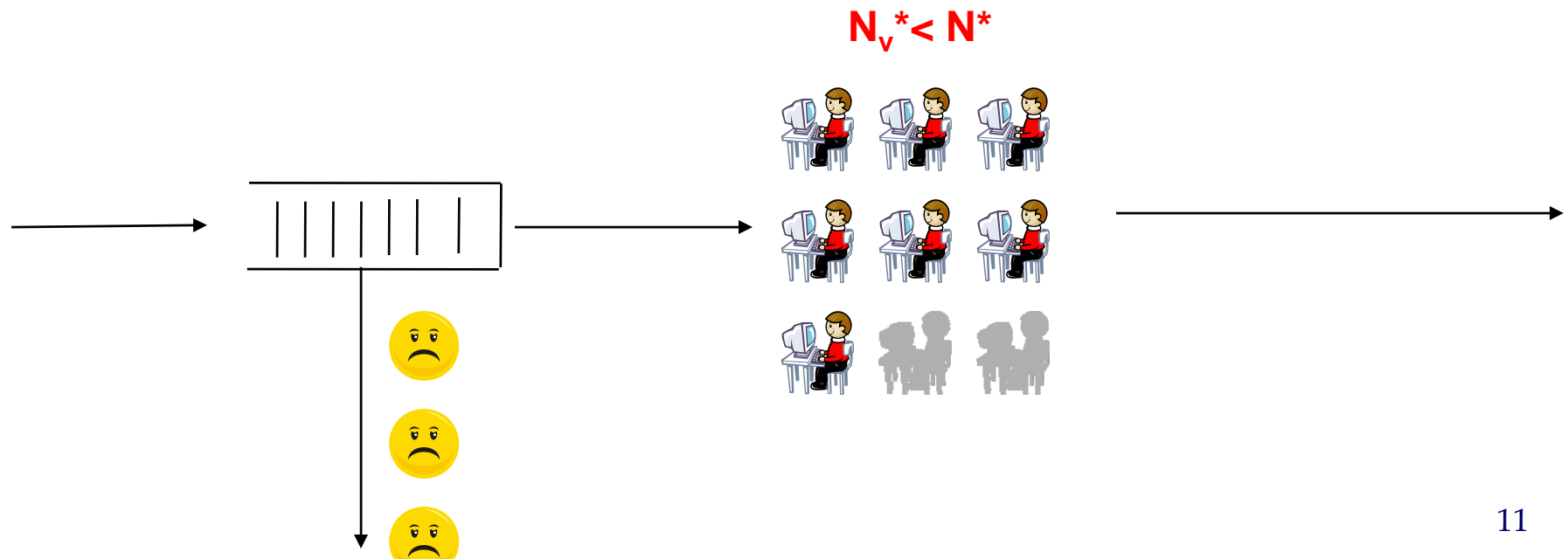
$$\pi_c^* = \pi_s(N^*) - V$$

for the client, where V = the minimum acceptable payment to the vendor (reservation value).

Why not just pay per call (PPC)?

Too expensive

- Because the system is stochastic, N^* must include a capacity buffer.
- To get the vendor to build the capacity buffer, the payment-per-call must be extremely high.
- Client loses money *or* the system is understaffed....



A Contract To Maximize Client Profits: PPC with a Service Level Agreement (SLA)

1. A large penalty for not meeting the SLA, e.g., $p \geq r\lambda$.
2. Set SLA parameters so that the constraint is tight at N^* ,
i.e., choose t and α such that $Pr\{wait < t\} = \alpha$ at N^* .
3. Pay-per-call sufficient to meet vendor's participation

constraint,

$$r = \frac{V + cN^*}{\lambda(1 - F(N^*))}.$$

Other Client-Optimal Contracts

- Use Pay-per-time (PPT) instead of PPC
- Can replace SLA, $Pr\{\text{wait} < t\} \geq \alpha$,
with constraints on the average waiting time or the abandonment rate.
- Average Handle Time (AHT) terms are not needed to make these contracts 'work'.
(But AHT's will be needed later...)

Another Complication: Information Asymmetry

- When outsourcing, the client can lose information about,
 - How the work is done (process design)
 - Knowledge of labor market (available skills)
 - Training methods
 - How hard is the work (slope of the learning curve).

The client is unsure about the vendor's productivity.
- Therefore, in our model, the client is uncertain about the service rate μ .
- Now which contracts work well for the vendor?

Two Types of Vendors, Two PPC Contracts?

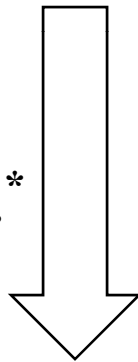


'slow' vendor, μ_L



'fast' vendor, $\mu_H > \mu_L$

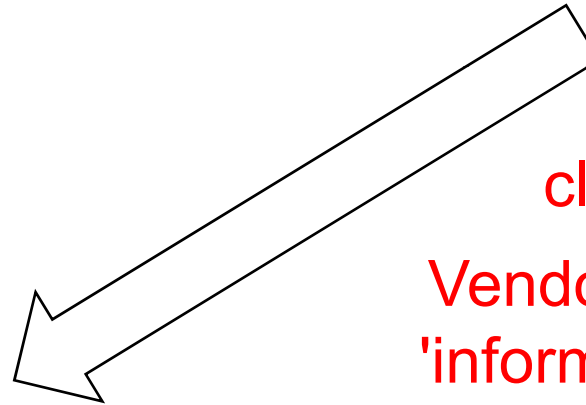
$N^*_V = N^*_L$
client profit = π_c^*



PPC+SLA contract tailored
for slow vendor



$N^*_V \neq N^*_H$
client profit $< \pi_c^*$
Vendor captures
'information rents'



PPC+SLA contract tailored
for fast vendor



**Contracts fail
to 'screen' vendors**

Identifying Screening Contracts: An Observation about Queues

- Queue L: Service rate μ_L , staffing level N_L
- Queue H: Service rate μ_H , staffing level N_H
- Assume $\mu_H > \mu_L$ but $\mu_H N_H = \mu_L N_L$
- Then given the same customer arrival rate,

customers in queue L will experience lower queueing times and abandonment rates.

This property (which follows from a heavy-traffic queueing approximation) allows us to show that the following contracts successfully screen.

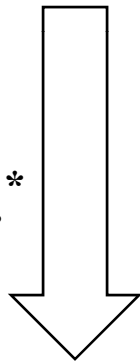
Screening with Pay-Per-Time and Pay-Per-Call



'slow' vendor, μ_L

$$N^*_V = N^*_L$$

client profit = π_c^*



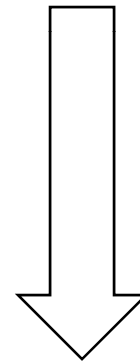
PPT+SLA contract tailored
for slow vendor



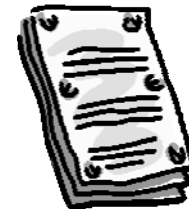
'fast' vendor, μ_H

$$N^*_V = N^*_H$$

client profit = π_c^*



PPC+SLA contract tailored
for fast vendor



Screening with Two Pay-Per-Time Contracts



'slow' vendor, μ_L

$N^*_V = N^*_L$
client profit = π_c^*

PPT+SLA contract tailored
for slow vendor



'fast' vendor, μ_H

$N^*_V = N^*_H$
client profit = π_c^*

PPT+SLA+**AHT** contract tailored
for fast vendor

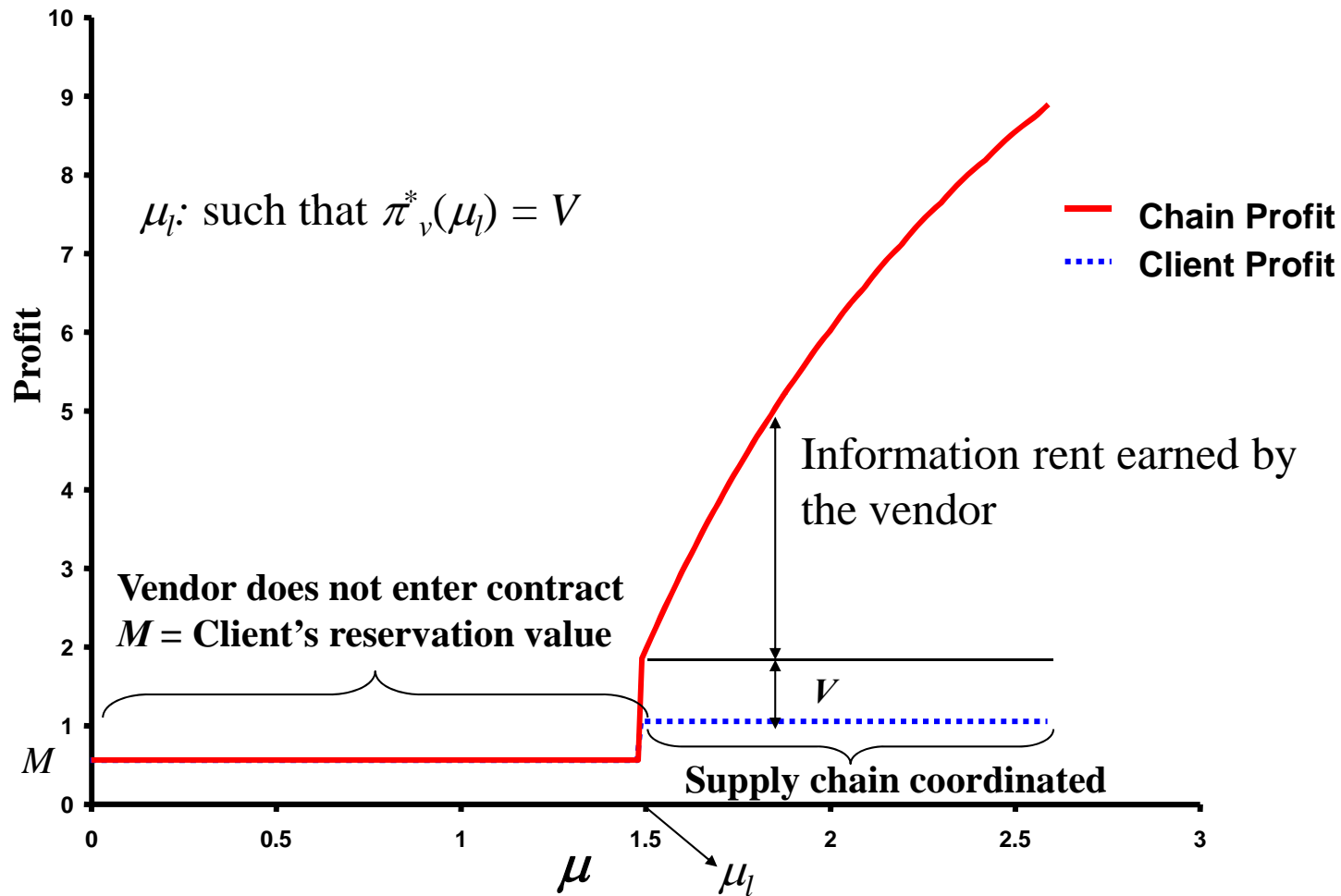


More Vendor Types: Arbitrary Productivity

- Vendor's true service rate = μ , known only to the vendor.
- Client has a prior probability distribution of μ .
- Vendor can choose to operate at

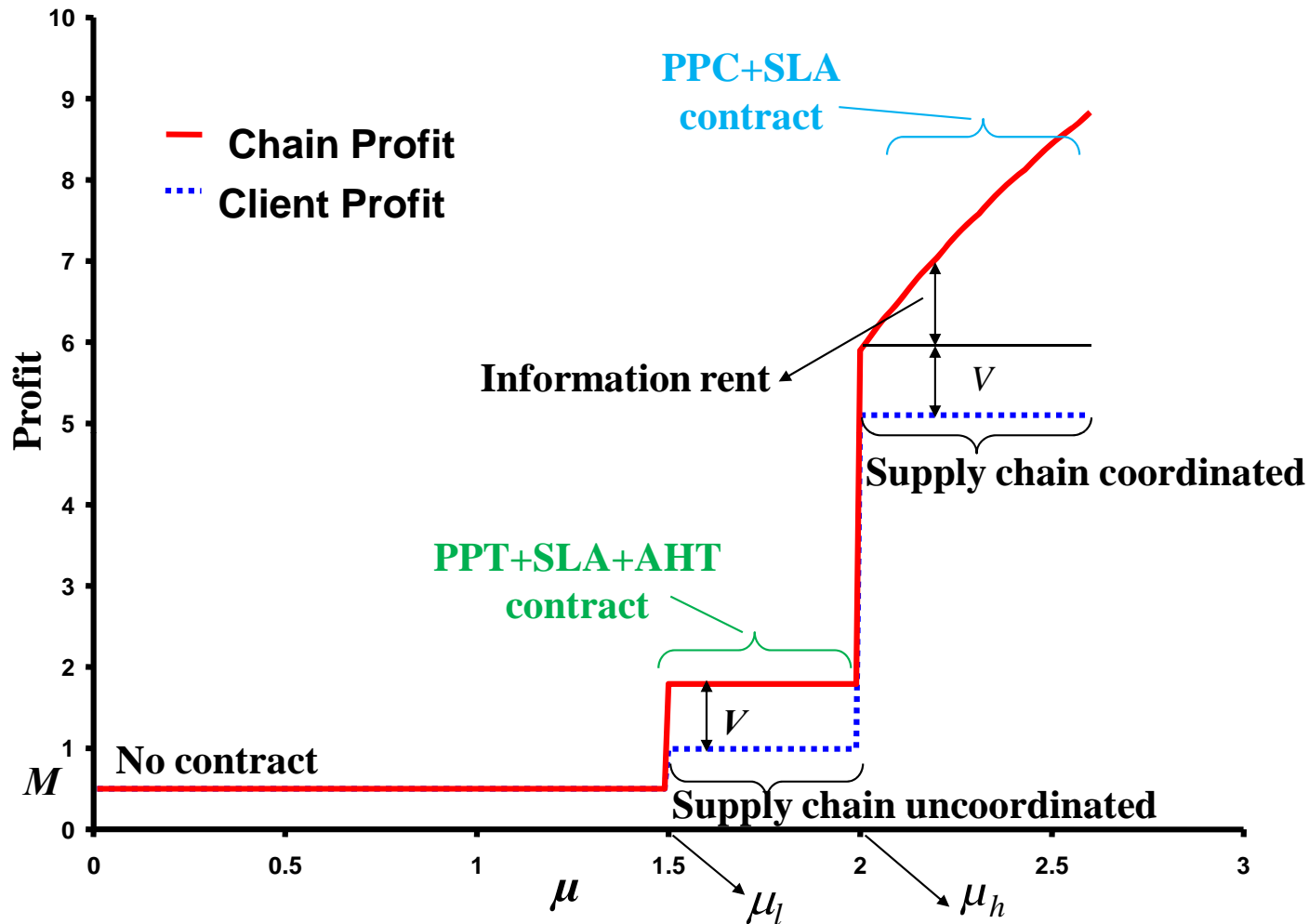
$$\mu_v \in (0, \mu].$$

A Single Pay-Per-Call (+SLA) Contract



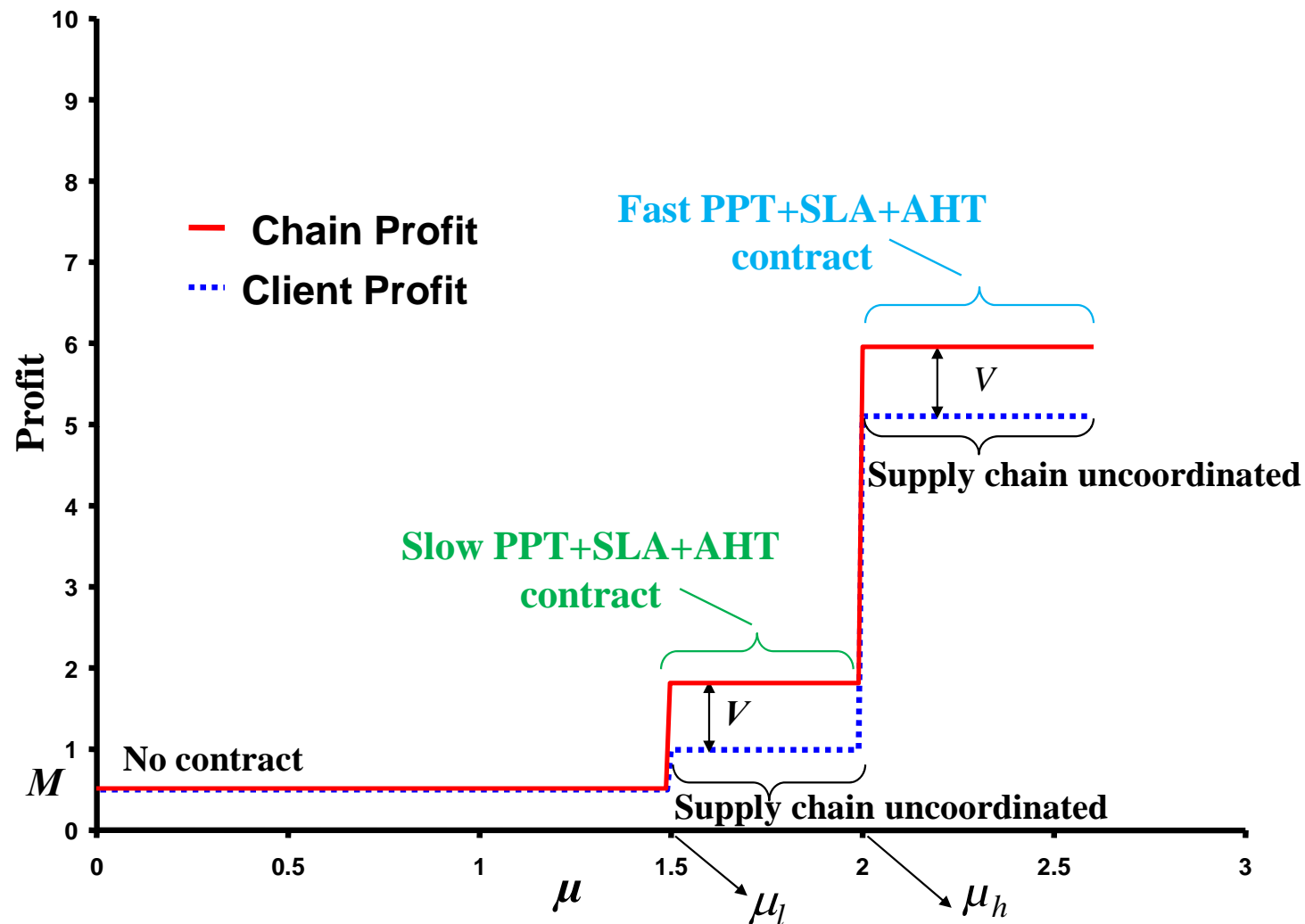
- $N^*_V = N^*$ (coordinates chain) for *participating* vendors
- Large information rents paid to fastest vendors.

Screening with Pay-Per-Time and Pay-Per-Call



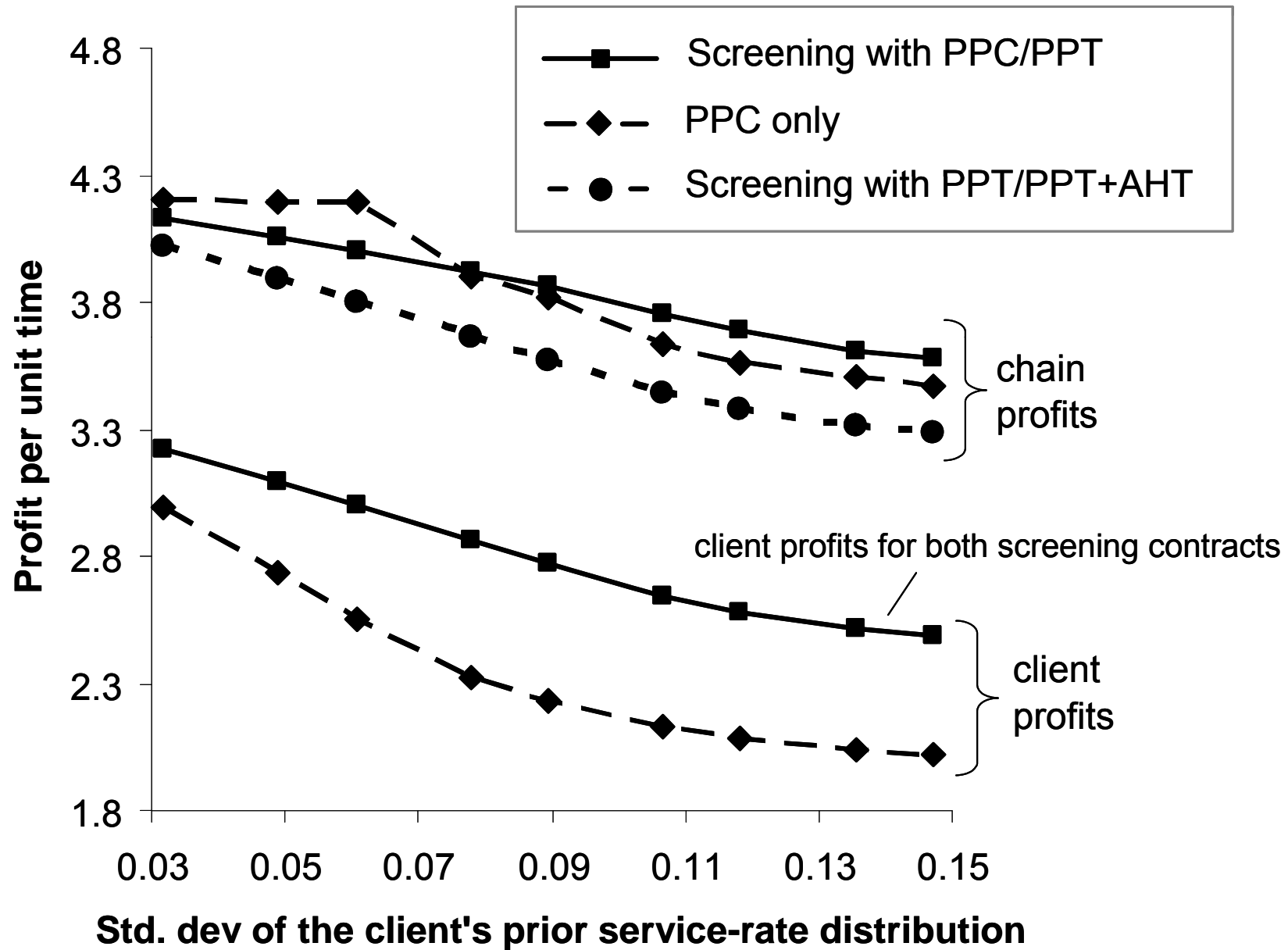
- Reduces information rents paid to fastest vendors
- Can capture a wider range of vendors
- $N^*_v \neq N^*$ in certain regions (chain uncoordinated)

Screening using two PPT contracts



- Chain not coordinated when vendors are fast (they make more money by slowing down).

Performance As Client Uncertainty Increases



Conclusions

- PPC or PPT must be supplemented (e.g., with an SLA) to ensure optimal staffing.
- When clients are uncertain about vendor productivity,
 - A choice between Pay-Per-Time and Pay-Per-Call-based contracts can *screen* one fast, one slow vendor.
 - A choice of Pay-Per-Time contracts also screens.
 - The PPT/PPC choice offers superior performance when there is a continuum of vendors.