ORIE 4154: Pricing and Market Design **Fall 2026**

Lecture 1 — Pricing with Full Information

Lecturer: Sid Banerjee *Date: August 25, 2026*

1.1 Overview

A first instinct about pricing is that the difficult part is *information*: the seller does not know how much buyers value the product. In this lecture we temporarily remove that issue. We assume throughout that buyer values are known perfectly and ask:

If we know everyone's values, is choosing the right price easy?

We will see that even under complete information the answer depends on what a price is supposed to accomplish. In progressively richer examples, prices will:

(i) divide gains from trade;

(ii) ration scarce supply;

(iii) decentralize a welfare-maximizing allocation across heterogeneous goods; and

(iv) manage service capacity over time and regulate participation when transactions generate congestion externalities.

This gives a useful organizing question for the course:

When do simple prices coordinate a market, and when do we need richer rules?

1.2 One item and one buyer: price divides surplus

There is one seller, one buyer, and one indivisible item. The buyer's value is v and the seller's opportunity cost is c. If trade occurs at price p, then

$$\underbrace{v-c}_{\text{total gains from trade}} = \underbrace{v-p}_{\text{buyer surplus}} + \underbrace{p-c}_{\text{seller surplus}}.$$

Trade is efficient if and only if $v \geq c$. If the seller knows v and can make a take-it-or-leave-it offer, then (up to the usual tie-breaking convention) the seller can charge

$$p = v,$$

extracting all of the gains from trade whenever $v \geq c$.

Thus, in this simplest complete-information setting, pricing appears almost trivial: the price determines *how* the gains from trade are split, but not whether the efficient trade can be implemented.

Remark 1.1. *Already there are two possible objectives: maximize total surplus, or maximize seller revenue/profit. They happen to be easy to reconcile here, but this will cease to be true almost immediately.*

1.3 One item and many buyers: price also allocates

Suppose there is one item and m buyers with known values $v_1, \dots, v_m$. Write the order statistics as

$$v_{(1)} \geq v_{(2)} \geq \cdots \geq v_{(m)}.$$

The efficient allocation gives the item to a buyer with value $v_{(1)}$.

With a unique highest-valued buyer, an anonymous posted price $p = v_{(1)}$ implements this allocation (again assuming an indifferent buyer accepts). More generally, a posted price determines a *demand set*

$$D(p) = \{i : v_i \geq p\}.$$

If $|D(p)| > 1$, the price alone does not specify which buyer gets the unique item: a rationing or allocation rule is also needed.

This is our first hint that prices and allocation rules are intertwined.

1.4 k identical items: market-clearing versus revenue-maximizing prices

Now suppose the seller has k identical items and $m > k$ unit-demand buyers. The welfare-maximizing allocation gives items to the k highest-valued buyers.

Ignoring ties, a price p that induces exactly k buyers to demand the good satisfies

$$v_{(k+1)} < p \leq v_{(k)}.$$

Such a price *clears the market*: exactly the available supply is demanded.

But a revenue-maximizing uniform price need not clear the market. If the seller posts a single price $p = v_{(j)}$ with $j \le k$, revenue is

$$R_j = j v_{(j)}.$$

Hence a revenue-maximizing uniform price solves

$$\max_{1 \le j \le k} j v_{(j)}.$$

It can be optimal to leave capacity unused.

Example 1.2. *Let $k = 2$ and values be $(10, 4, 3)$. A market-clearing price near 4 sells two units and earns about 8. Charging 10 sells only one unit but earns 10. Thus the revenue-maximizing price does not clear the market.*

The phrase "optimal price" is therefore incomplete until we specify the objective:

$$\boxed{\text{market clearing? welfare? revenue?}}$$

1.5 Heterogeneous goods: prices as dual variables

The preceding examples involve a single homogeneous product. Now suppose there are three buyers and three distinct items A, B, C. Each buyer wants at most one item, and values are

$$V = (v_{ij}) = \begin{pmatrix} 10 & 8 & 3 \\ 9 & 5 & 7 \\ 4 & 9 & 8 \end{pmatrix}.$$

Rows correspond to buyers and columns to items.

Remark 1.3 (Connection to Assignment 1). *The central economic conclusion is that one vector of item prices can make decentralized buyer choices reproduce a welfare-maximizing assignment. The primal–dual derivation below explains why this works; Assignment 1 develops the complementary-slackness argument in greater detail.*

Notice an immediate complication: the demand for item A cannot be determined from the values in column A alone. If B becomes cheaper, buyer 1 might switch from A to B. Prices interact through *substitution behavior*.

1.5.1 The centralized allocation problem

Let x_{ij} indicate whether buyer i receives item j. The welfare-maximizing assignment solves

$$\begin{aligned} \max_{x \geq 0} \quad & \sum_{i=1}^{3} \sum_{j \in \{A,B,C\}} v_{ij} x_{ij} & \\ \text{s.t.} \quad & \sum_{j} x_{ij} \leq 1 & \forall i, \\ & \sum_{i} x_{ij} \leq 1 & \forall j. \end{aligned}$$

The assignment polytope is integral, so we may solve the LP relaxation without explicitly imposing $x_{ij} \in \{0, 1\}$.

For our matrix, the optimal assignment is

$$1 \to A, \qquad 2 \to C, \qquad 3 \to B,$$

with total welfare

$$10 + 7 + 9 = 26.$$

The next-best cyclic assignment $1 \to B, 2 \to A, 3 \to C$ has value $8 + 9 + 8 = 25$.

1.5.2 The dual and market-clearing prices

Associate a dual variable $u_i \geq 0$ with each buyer constraint and a dual variable $p_j \geq 0$ with each item constraint. The dual LP is

$$\begin{aligned} \min_{u,p \geq 0} \quad & \sum_{i} u_i + \sum_{j} p_j & \\ \text{s.t.} \quad & u_i + p_j \geq v_{ij} & \forall i, j. \end{aligned}$$

The notation is suggestive. Interpret p_j as the price of item j and u_i as the utility earned by buyer i. Then the dual inequalities become

$$u_i \geq v_{ij} - p_j \qquad \forall j,$$

which says that u_i is at least the surplus buyer i could obtain from any item.

Consider the price vector

$$(p_A, p_B, p_C) = (5.8, 4, 3.7).$$

The buyers' net values are

	A	B	C	chosen item
1	$10 - 5.8 = 4.2$	$8 - 4 = 4$	$3 - 3.7 = -0.7$	A
2	$9 - 5.8 = 3.2$	$5 - 4 = 1$	$7 - 3.7 = 3.3$	C
3	$4 - 5.8 = -1.8$	$9 - 4 = 5$	$8 - 3.7 = 4.3$	B

Thus each buyer independently chooses exactly the welfare-maximizing assignment. Taking

$$(u_1, u_2, u_3) = (4.2, 3.3, 5)$$

gives dual objective

$$4.2 + 3.3 + 5 + 5.8 + 4 + 3.7 = 26,$$

equal to the primal optimum. Hence these are optimal dual variables.

Proposition 1 (Prices decentralize the assignment). *For the unit-demand assignment problem, an optimal primal-dual pair (x^*, u^*, p^*) has the property that whenever $x^*_{ij} > 0$,*

$$u_i^* = v_{ij} - p_j^* = \max_{\ell}\{v_{i\ell} - p_\ell^*\}.$$

Thus the dual item variables can be interpreted as market-clearing prices that support a welfare-maximizing allocation.

Proof (Proof sketch): Complementary slackness gives $u_i^* + p_j^* = v_{ij}$ on every allocated edge. Dual feasibility gives $u_i^* + p_\ell^* \geq v_{i\ell}$ for every alternative item ℓ. Rearranging yields the result. ■

This is an important OR/economics connection:

$$\boxed{\text{LP duality} \;\longleftrightarrow\; \text{market-clearing prices.}}$$

It also foreshadows later lectures on discrete choice, matching markets, and multi-item mechanism design.

Remark 1.4. *The supporting price vector generally need not be unique. Also, these prices are chosen to support an efficient allocation; they are not automatically the seller's revenue-maximizing prices.*

1.6 Queue Lab and a Naor-style queue

We now go back to a single homogeneous service. Surprisingly, pricing can again become subtle even though every customer's value is known and identical.

The classical Naor model studies strategic entry into a queue. Here we use two deliberately simple discrete-time variants. A deterministic calibration first shows why extracting as much as possible from the *current* customer can be a poor long-run policy. We then turn to stochastic arrivals to isolate the congestion externality.

Remark 1.5 (Companion activity: Queue Lab). *The companion Queue Lab game separates two objectives that should not be conflated.*

(1) In Calibration Day*, price the first few scheduled robots manually and then program a workload-dependent controller. The score is operator credits per cycle.*

(2) In Quiet Shift *and* Rush Hour*, replay a fixed dispatch under different controllers. The score is city surplus per request.*

(3) Keep both operator credits and city surplus visible. A controller that performs well for one objective need not perform well for the other.

(4) The optional Mixed Fleet *mode varies values, waiting costs, and repair requirements; it asks whether workload alone remains a sufficient state for pricing.*

The calculations below explain the deterministic calibration and the two homogeneous stochastic shifts.

1.6.1 Calibration Day: current extraction versus long-run revenue

In Calibration Day, one robot arrives every two cycles. Each repair requires $s = 3$ cycles, every robot has value $V = 5$, and waiting costs $c = 1$ per cycle. If W_n is the workload seen by the nth scheduled robot and J_n indicates whether it docks, then

$$W_{n+1} = \left[W_n + 3J_n - 2\right]^+.$$

A workload-dependent access price $\tau(w)$ gives the arriving robot utility

$$U(w) = 5 - \tau(w) - w.$$

Suppose the operator behaves myopically and extracts all of the current robot's residual willingness to pay whenever possible:

$$\tau_{\text{myopic}}(w) = [5 - w]^+.$$

Starting from $W_1 = 0$, the first six robots dock at workloads $0, 1, \ldots, 5$ and pay $5, 4, \ldots, 0$. The workload then cycles through $4, 5, 6$: two robots dock, one reroutes, and the operator collects only one credit every six cycles.

By contrast, the fixed price $\tau = 4$ produces the workload cycle $0, 1, 2$: two robots dock at price 4, then one reroutes while the bay drains. It collects eight credits every six cycles. Thus

$$\underbrace{\text{extract the most from the next customer}}_{\text{myopic rule}} \quad \neq \quad \underbrace{\text{maximize revenue over time}}_{\text{dynamic objective}}.$$

The point is not that a fixed price is always optimal. It is that admitting work changes the state faced by future customers, so a pricing policy must account for an opportunity cost that a one-customer calculation misses. This is an early preview of dynamic pricing and revenue management.

1.6.2 The stochastic workload model

Time is slotted, $t = 0, 1, 2, \ldots$. There is one server. Let

$$W_t \in \mathbb{Z}_+$$

be the unfinished workload (in service slots) immediately before the arrival decision in period t.

In every period a potential customer arrives independently with probability λ. Every admitted customer requires exactly

$$s = 4$$

service slots. The value of service is

$$V = 10,$$

and waiting costs one unit per slot. If the platform charges an admission toll $\tau(w)$, a customer arriving to workload $W_t = w$ gets utility

$$U(w) = 10 - \tau(w) - w.$$

Hence a self-interested customer joins exactly when

$$w \leq 10 - \tau(w).$$

If $J_t \in \{0, 1\}$ indicates that an arrival is admitted, workload evolves as

$$W_{t+1} = \big[W_t + sJ_t - 1\big]^+.$$

The toll therefore acts as an admission-control instrument. A fixed toll τ creates the workload threshold $w \leq 10 - \tau$; a programmed controller $\tau(w)$ permits richer state-dependent admission rules.

The game reports normalized versions of two different accounting measures. Operator credits are the toll payments,

$$\sum_t \tau(W_t) J_t.$$

City surplus counts the value of each admitted repair minus that robot's waiting cost,

$$\sum_t (10 - W_t) J_t.$$

The toll is a transfer from the robot to the operator, so it is not subtracted from city surplus. More importantly, neither expression by itself displays the future delay that today's admission creates; that effect appears through later workloads.

1.6.3 The load parameter

First imagine that all potential arrivals are admitted. When the server is busy,

$$\mathbb{E}[W_{t+1} - W_t] \approx s\lambda - 1.$$

This identifies the familiar load parameter

$$\rho = s\lambda.$$

The sign of $\rho - 1$ already tells us something important.

1.6.4 Example 1: low congestion

Take

$$\lambda = 0.05, \qquad s = 4,$$

so

$$\rho = s\lambda = 0.2 \ll 1.$$

When the server is busy, expected workload drift under unrestricted admission is

$$0.2 - 1 = -0.8.$$

Thus workload tends to drain rapidly.

We can also make a simple local externality calculation. Suppose the system is initially empty and a customer enters. A later arrival r slots into this customer's service is delayed by roughly $(s-r)^+$ slots solely because of this customer. Ignoring downstream interactions, the expected direct waiting cost imposed on arrivals during these four service slots is

$$\lambda \sum_{r=1}^{s-1} (s-r) = 0.05(3+2+1) = 0.3.$$

Relative to the service value $V = 10$, the direct congestion externality is small in this low-load example.

1.6.5 Example 2: congestion

Now take

$$\lambda = 0.4, \qquad s = 4,$$

so

$$\rho = s\lambda = 1.6 > 1.$$

If everybody were admitted, then while the system is busy,

$$\mathbb{E}[W_{t+1} - W_t] \approx 1.6 - 1 = 0.6 > 0,$$

so workload would grow rather than drain. Some form of admission control is essential for a stable, usable system.

The same local externality calculation gives

$$\lambda \sum_{r=1}^{s-1} (s - r) = 0.4(3 + 2 + 1) = 2.4,$$

even before accounting for downstream congestion effects. Yet an entering customer considers only their own waiting cost w, not the additional delay imposed on other customers.

At zero toll, the private entry rule is

$$w \leq 10.$$

A positive fixed toll shifts this threshold downward. For example, $\tau = 4$ produces the private rule

$$w \leq 6.$$

We are *not* claiming that 6 is the socially optimal threshold in this example; the point is that any desired threshold $h \leq 10$ can be decentralized by the toll

$$\tau = 10 - h.$$

Thus, if a planner accounts for congestion and chooses a social admission threshold h^* below the laissez-faire threshold, a price can implement it.

Remark 1.6 (The Naor lesson). *The customer internalizes their own waiting cost but not the waiting cost their entry creates for others. Thus private entry can exceed socially desirable entry. Here price is not merely dividing surplus or rationing a fixed stock of goods: it is an instrument for internalizing a congestion externality.*

The two parts of Queue Lab therefore expose two distinct reasons to condition current admission on future consequences. Under a revenue objective, accepting a low-paying robot may crowd out more valuable future access. Under a welfare objective, an entering robot may impose delay on others. In both cases the state variable W_t compresses the relevant history, but the correct controller depends on the objective.

1.7 What did "price" mean in the five examples?

The examples can be summarized as follows.

Setting	**Role of price**
One buyer, one item	Divide gains from trade; extract surplus.
One item, many buyers	Determine demand and help allocate a scarce item.
k identical items	Clear the market—or, under a revenue objective, deliberately leave supply unsold.
Heterogeneous items	Coordinate substitution choices and decentralize a welfare-maximizing assignment; LP dual variables become prices.
Congested service	Manage future opportunity cost under a revenue objective; alter participation to internalize a delay externality under a welfare objective.

The overarching lesson is that there is no unique notion of the "right" price without specifying both the objective and the institutional environment.

1.8 Looking ahead

A recurring question: what changes with scale?

Large systems can be harder because their state and action spaces defeat exact optimization. Yet scale can also create useful regularity: aggregate quantities stabilize, competition and matching opportunities thicken, and individual participants may become less pivotal. The queue's load parameter $\rho = s\lambda$ is our first glimpse of this compression—many arrival and service events summarized by one first-order quantity. We will keep asking which benefit or cost of scale is operating.

Everything in this lecture assumed that values were known. In most real markets they are not. The next step is therefore to replace the known numbers v_i by a model of heterogeneous values,

$$V \sim F,$$

and understand how a distribution of values generates a demand curve

$$D(p) = N \Pr(V \geq p) = N(1 - F(p)).$$

This leads to the next question: *where does demand come from, and what can we infer about it from data?*

References and pointers

- R. Vohra and L. Krishnamurthi, *Principles of Pricing: An Analytical Approach*, for the basic surplus/pricing perspective.

- K. Talluri and G. van Ryzin, *The Theory and Practice of Revenue Management*, for pricing, capacity, and opportunity-cost viewpoints developed later in the course.
- P. Naor, "The Regulation of Queue Size by Levying Tolls," *Econometrica*, 1969, for the classical queue-entry model motivating the final example. The discrete-time model above is a pedagogical variant, not Naor's original model.