Dynamic Programming Approach to Behavioral Operations

Park Sinchaisri

OIDD 934 Spring 2018

What is *Behavioral Operations*?

Operations Management	+	Human Behavior /Emotions
"People are fully rational"		Biases, limited cognitive ability, ...

What is *Behavioral Operations*?

Operations Management + Human Behavior /Emotions

Supply Chain

Capacity Allocation

Adelman & Mersereau
Management Science '13

Memory

Revenue Management

Dynamic Pricing

Nasiry & Popescu
Operations Research '11

Loss Aversion

Dynamic Capacity Allocation to Customers Who Remember Past Service

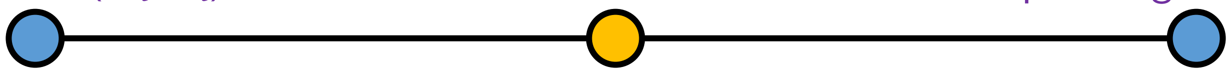
Adelman & Mersereau, Management Science '13

- Supplier sells a single good to a finite set of customers: 1, 2, ..., n
- Capacity constraint of Xbar > 0 units per period
- **Goal:** Maximize supplier's long-run average profit
- Model **"goodwill" toward the supplier / customer's memory**
- Trade-off **margin, demand volatility,** and **memory**
- Show that **ADP policy** outperforms Greedy

MDP Formulation

Customer places an order $y^i(G^i_t, D^i_t)$

Receiving x^i_t, customer updates goodwill G^i_{t+1}



Supplier chooses vector of shipments $\mathbf{x_t}$
makes profit $r^i > 0$ for each unit of x^i_t

- **Goodwill**: based on quality of past service $\boldsymbol{G^i_t}$ entering period t, modeled as exponential smoothing of utilities from past fill rates.

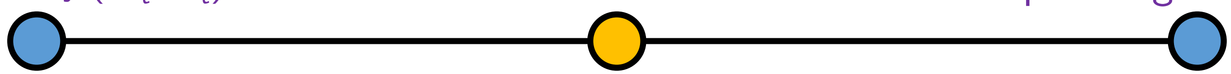
$$G^i_{t+1} = \begin{cases} \beta^i G^i_t + u^i(x^i_t / y^i(G^i_t, D^i_t)) & \text{if } y^i(G^i_t, D^i_t) > 0, \\ \beta^i G^i_t & \text{if } y^i(G^i_t, D^i_t) = 0, \end{cases}$$

B^i: (0,1): customer-specific memory parameter
u^i: [0, 1] -> R+: nondecreasing c-specific utility function
D_t: demands follow stationary dist, dependent across i, but independent across t

MDP Formulation

Customer places an order $y^i(G^i_t, D^i_t)$

Receiving x^i_t, customer updates goodwill G^i_{t+1}



Supplier chooses vector of shipments $\mathbf{x_t}$
makes profit $r^i > 0$ for each unit of x^i_t

- **Goodwill**: based on quality of past service $\boldsymbol{G^i_t}$ entering period t, modeled as exponential smoothing of utilities from past fill rates.

$$G^i_{t+1} = \begin{cases} \beta^i G^i_t + u^i(x^i_t / y^i(G^i_t, D^i_t)) & \text{if } y^i(G^i_t, D^i_t) > 0, \\ \beta^i G^i_t & \text{if } y^i(G^i_t, D^i_t) = 0, \end{cases}$$

  B^i: (0,1): customer-specific memory parameter
  u^i: [0, 1] -> R+: nondecreasing c-specific utility function
  D_t: demands follow stationary dist, dependent across i, but independent across t

- **State space:** $\mathcal{G} = \{\mathbf{G} \in \mathbb{R}^n_+ : u^i(0)/\ (1-\beta^i) \le G^i \le u^i(1)/(1-\beta^i)\, \forall\, i \in \mathcal{I}\}$
- **Action space:** allocations provided under each demand scenario

$$\mathcal{X}(\mathbf{G}, \mathbf{D}) = \left\{ \mathbf{x} \in \mathbb{R}^n_+ : \sum_{i \in \mathcal{I}} x^i \le \bar{X}, 0 \le x^i \le y^i(G^i, D^i)\ \forall i \in \mathcal{I} \right\}$$

Challenges from MDP

Let $\boldsymbol{\beta}\mathbf{G} + \mathbf{u}(\mathbf{x}/\mathbf{y}(\mathbf{G}, \mathbf{D}))$ be a vectorized representation of the expressions $\beta^i G^i + u^i(x^i/y^i(G^i, D^i))$ for all i, which denote the next state. We can write the unichain optimality equations in two equivalent ways:

$$h(\mathbf{G}) = \mathsf{E}_{\mathbf{D}}\left[\max_{\mathbf{x}\in\mathscr{X}(\mathbf{G},\mathbf{D})}\left\{\sum_i r^i x^i - \rho + h\left(\boldsymbol{\beta}\mathbf{G} + \mathbf{u}\left(\frac{\mathbf{x}}{\mathbf{y}(\mathbf{G},\mathbf{D})}\right)\right)\right\}\right] \quad \forall \mathbf{G} \in \mathscr{G}, \qquad (1)$$

$$h(\mathbf{G}) = \max_{\mathbf{x}\in\mathscr{X}(\mathbf{G})} \mathsf{E}_{\mathbf{D}}\left[\sum_i r^i x^i(\mathbf{D}) - \rho + h\left(\boldsymbol{\beta}\mathbf{G} + \mathbf{u}\left(\frac{\mathbf{x}(\mathbf{D})}{\mathbf{y}(\mathbf{G},\mathbf{D})}\right)\right)\right] \quad \forall \mathbf{G} \in \mathscr{G}, \qquad (2)$$

where $h(\cdot) \in \mathbb{B}(\mathscr{G})$ is the bias function

Challenges from MDP

Let $\boldsymbol{\beta}\mathbf{G} + \mathbf{u}(\mathbf{x}/\mathbf{y}(\mathbf{G}, \mathbf{D}))$ be a vectorized representation of the expressions $\beta^i G^i + u^i(x^i/y^i(G^i, D^i))$ for all i, which denote the next state. We can write the unichain optimality equations in two equivalent ways:

$$h(\mathbf{G}) = \mathsf{E}_{\mathbf{D}}\left[\max_{\mathbf{x}\in\mathscr{X}(\mathbf{G},\mathbf{D})}\left\{\sum_i r^i x^i - \rho + h\left(\boldsymbol{\beta}\mathbf{G} + \mathbf{u}\left(\frac{\mathbf{x}}{\mathbf{y}(\mathbf{G},\mathbf{D})}\right)\right)\right\}\right] \quad \forall \mathbf{G} \in \mathscr{G}, \qquad (1)$$

$$h(\mathbf{G}) = \max_{\mathbf{x}\in\mathscr{X}(\mathbf{G})} \mathsf{E}_{\mathbf{D}}\left[\sum_i r^i x^i(\mathbf{D}) - \rho + h\left(\boldsymbol{\beta}\mathbf{G} + \mathbf{u}\left(\frac{\mathbf{x}(\mathbf{D})}{\mathbf{y}(\mathbf{G},\mathbf{D})}\right)\right)\right] \quad \forall \mathbf{G} \in \mathscr{G}, \qquad (2)$$

where $h(\cdot) \in \mathbb{B}(\mathscr{G})$ is the bias function

However,

- Capacity constraint makes it imp. to decompose problem by customers
- Stochastic demands lead to stochastic dynamics -> steady states needs to be characterized by complex joint probability distribution
- Using fill rate to measure service quality induces nonlinear state dynamics

Challenges from MDP

Let $\boldsymbol{\beta}\mathbf{G}+\mathbf{u}(\mathbf{x}/\mathbf{y}(\mathbf{G},\mathbf{D}))$ be a vectorized representation of the expressions $\beta^i G^i + u^i(x^i/y^i(G^i, D^i))$ for all i, which denote the next state. We can write the unichain optimality equations in two equivalent ways:

$$h(\mathbf{G}) = \mathsf{E}_{\mathbf{D}}\left[\max_{\mathbf{x}\in\mathscr{X}(\mathbf{G},\mathbf{D})}\left\{\sum_i r^i x^i - \rho + h\left(\boldsymbol{\beta}\mathbf{G}+\mathbf{u}\left(\frac{\mathbf{x}}{\mathbf{y}(\mathbf{G},\mathbf{D})}\right)\right)\right\}\right] \quad \forall \mathbf{G}\in\mathscr{G}, \tag{1}$$

$$h(\mathbf{G}) = \max_{\mathbf{x}\in\mathscr{X}(\mathbf{G})}\mathsf{E}_{\mathbf{D}}\left[\sum_i r^i x^i(\mathbf{D}) - \rho + h\left(\boldsymbol{\beta}\mathbf{G}+\mathbf{u}\left(\frac{\mathbf{x}(\mathbf{D})}{\mathbf{y}(\mathbf{G},\mathbf{D})}\right)\right)\right] \quad \forall \mathbf{G}\in\mathscr{G}, \tag{2}$$

where $h(\cdot)\in\mathbb{B}(\mathscr{G})$ is the bias function

However,

- Capacity constraint makes it imp. to decompose problem by customers
- Stochastic demands lead to stochastic dynamics -> steady states needs to be characterized by complex joint probability distribution
- Using fill rate to measure service quality induces nonlinear state dynamics

- **Solution**: Switch = in (2) into >= and get infinite-d LP

$$(\mathrm{D}_0) \quad \inf_{\rho,h} \rho \tag{3a}$$

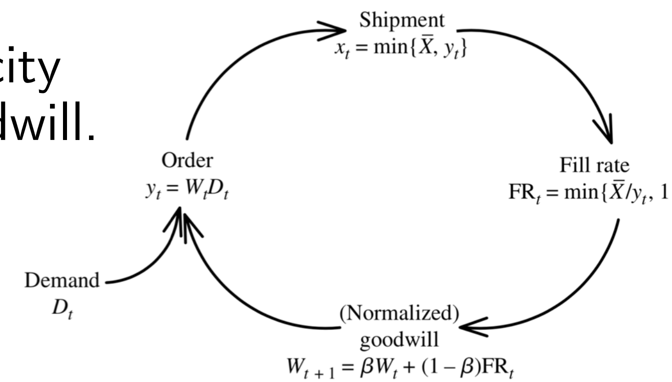
$$h(\mathbf{G}) \geq \mathsf{E}_{\mathbf{D}}\left[\sum_{i\in\mathcal{I}} r^i x^i(\mathbf{D}) - \rho + h\left(\boldsymbol{\beta}\mathbf{G}+\mathbf{u}\left(\frac{\mathbf{x}(\mathbf{D})}{\mathbf{y}(\mathbf{G},\mathbf{D})}\right)\right)\right] \quad \forall (\mathbf{G},\mathbf{x})\in\mathscr{M}, \tag{3b}$$

$$h\in\mathbb{B}(\mathscr{G}), \rho\in\mathbb{R}. \tag{3c}$$

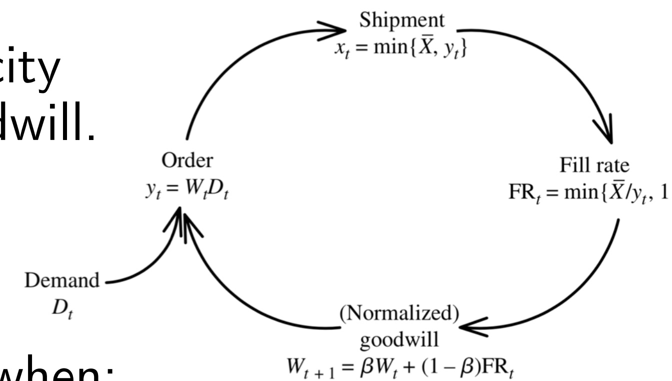
Greedy Policy Can Be Bad

Myopic margin-greedy policy = allocates capacity based on margins alone without regard to goodwill.

$$\max \sum_i r^i x_t^i \quad \text{s.t.} \ \sum_i x_t^i \le \bar{X}, \quad 0 \le x_t^i \le y_t^i \quad \forall i \in \mathcal{I}.$$



Greedy Policy Can Be Bad



Myopic margin-greedy policy = allocates capacity based on margins alone without regard to goodwill.

$$\max \sum_i r^i x_t^i \quad \text{s.t.} \sum_i x_t^i \le \bar{X}, \quad 0 \le x_t^i \le y_t^i \quad \forall i \in \mathcal{I}.$$

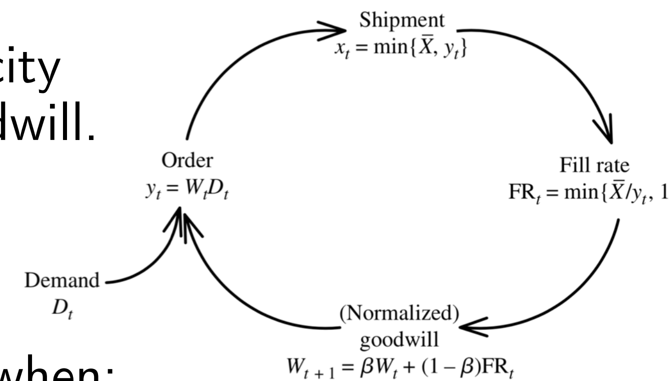
Generally, greedy policy performs well, except when:

- Customers have high expectations
- Little tolerance for lapses in service
- Long memory of past service
- Goodwills are not maintained to be similar across customers

Greedy Policy Can Be Bad

Myopic margin-greedy policy = allocates capacity based on margins alone without regard to goodwill.

$$\max \sum_i r^i x_t^i \quad \text{s.t.} \ \sum_i x_t^i \leq \bar{X}, \quad 0 \leq x_t^i \leq y_t^i \quad \forall i \in \mathcal{I}.$$



Generally, greedy policy performs well, except when:

- Customers have high expectations
- Little tolerance for lapses in service
- Long memory of past service
- Goodwills are not maintained to be similar across customers

Better policies strategically withhold capacity to keep customer goodwills within a manageable range, keep multiple customers ordering actively, and pooling demand variances across customers.

Approximate DP Approach

Based on a polynomial approximation of the optimal bias function *h* of D_0

$$(\mathrm{D}_0) \quad \inf_{\rho, h} \rho \tag{3a}$$

$$h(\mathbf{G}) \geq \mathsf{E}_{\mathbf{D}}\left[\sum_{i \in \mathcal{I}} r^i x^i(\mathbf{D}) - \rho + h\left(\boldsymbol{\beta}\mathbf{G} + \mathbf{u}\left(\frac{\mathbf{x}(\mathbf{D})}{\mathbf{y}(\mathbf{G}, \mathbf{D})}\right)\right)\right]$$

$$\forall (\mathbf{G}, \mathbf{x}) \in \mathcal{M}, \tag{3b}$$

$$h \in \mathbb{B}(\mathcal{G}), \rho \in \mathbb{R}. \tag{3c}$$

Now suppose we are given a collection $\mathcal{K} = \{1, 2, \ldots, K\}$ of basis functions $\{\phi_k(\cdot)\}_{k \in \mathcal{K}}$, where $\phi_k \in \mathbb{B}(\mathcal{G})$. For a weight vector $\mathbf{w} = (w_1, w_2, \ldots, w_K) \in \mathbb{R}^K$, we can approximate the bias function as

$$h(\mathbf{G}) \approx \sum_{k \in \mathcal{K}} w_k \phi_k(\mathbf{G}) \quad = \sum_{i \in \mathcal{I}} \sum_{j=1}^{N} w_{ij} \cdot (G^i)^j$$

Approximate DP Approach

Based on a polynomial approximation of the optimal bias function h of D_0

$$(\mathrm{D}_0) \quad \inf_{\rho, h} \rho \tag{3a}$$

$$h(\mathbf{G}) \geq \mathsf{E}_{\mathbf{D}}\left[\sum_{i \in \mathcal{I}} r^i x^i(\mathbf{D}) - \rho + h\left(\boldsymbol{\beta}\mathbf{G} + \mathbf{u}\left(\frac{\mathbf{x}(\mathbf{D})}{\mathbf{y}(\mathbf{G}, \mathbf{D})}\right)\right)\right]$$

$$\forall (\mathbf{G}, \mathbf{x}) \in \mathcal{M}, \tag{3b}$$

$$h \in \mathbb{B}(\mathcal{G}), \rho \in \mathbb{R}. \tag{3c}$$

$$(\mathrm{D}_\phi) \quad \inf_{\rho, \mathbf{w}} \rho \tag{10a}$$

$$\sum_{k \in \mathcal{K}} w_k \phi_k(\mathbf{G}) \geq \mathsf{E}_{\mathbf{D}}\left[\sum_{i \in \mathcal{I}} r^i x^i(\mathbf{D}) - \rho + \sum_{k \in \mathcal{K}} w_k \phi_k\left(\boldsymbol{\beta}\mathbf{G} + \mathbf{u}\left(\frac{\mathbf{x}(\mathbf{D})}{\mathbf{y}(\mathbf{G}, \mathbf{D})}\right)\right)\right]$$

$$\forall (\mathbf{G}, \mathbf{x}) \in \mathcal{M}, \tag{10b}$$

Upperbounds average reward of any policy

$$\mathbf{w} \in \mathbb{R}^K, \rho \in \mathbb{R}. \tag{10c}$$

Now suppose we are given a collection $\mathcal{K} = \{1, 2, \ldots, K\}$ of basis functions $\{\phi_k(\cdot)\}_{k \in \mathcal{K}}$, where $\phi_k \in \mathbb{B}(\mathcal{G})$. For a weight vector $\mathbf{w} = (w_1, w_2, \ldots, w_K) \in \mathbb{R}^K$, we can approximate the bias function as

$$h(\mathbf{G}) \approx \sum_{k \in \mathcal{K}} w_k \phi_k(\mathbf{G}) \quad = \sum_{i \in \mathcal{I}} \sum_{j=1}^{N} w_{ij} \cdot (G^i)^j$$

Approximate DP Approach

Based on a polynomial approximation of the optimal bias function h of D_0

$$(\mathrm{D}_0) \quad \inf_{\rho, h} \rho \tag{3a}$$

$$h(\mathbf{G}) \geq \mathsf{E}_{\mathbf{D}}\left[\sum_{i \in \mathcal{I}} r^i x^i(\mathbf{D}) - \rho + h\left(\boldsymbol{\beta}\mathbf{G} + \mathbf{u}\left(\frac{\mathbf{x}(\mathbf{D})}{\mathbf{y}(\mathbf{G}, \mathbf{D})}\right)\right)\right] \quad \forall (\mathbf{G}, \mathbf{x}) \in \mathcal{M}, \tag{3b}$$

$$h \in \mathbb{B}(\mathcal{G}), \rho \in \mathbb{R}. \tag{3c}$$

Now suppose we are given a collection $\mathcal{K} = \{1, 2, \ldots, K\}$ of basis functions $\{\phi_k(\cdot)\}_{k \in \mathcal{K}}$, where $\phi_k \in \mathbb{B}(\mathcal{G})$. For a weight vector $\mathbf{w} = (w_1, w_2, \ldots, w_K) \in \mathbb{R}^K$, we can approximate the bias function as

$$h(\mathbf{G}) \approx \sum_{k \in \mathcal{K}} w_k \phi_k(\mathbf{G}) \quad = \sum_{i \in \mathcal{I}} \sum_{j=1}^{N} w_{ij} \cdot (G^i)^j$$

$$(\mathrm{D}_\phi) \quad \inf_{\rho, \mathbf{w}} \rho \tag{10a}$$

$$\sum_{k \in \mathcal{K}} w_k \phi_k(\mathbf{G}) \geq \mathsf{E}_{\mathbf{D}}\left[\sum_{i \in \mathcal{I}} r^i x^i(\mathbf{D}) - \rho + \sum_{k \in \mathcal{K}} w_k \phi_k\left(\boldsymbol{\beta}\mathbf{G} + \mathbf{u}\left(\frac{\mathbf{x}(\mathbf{D})}{\mathbf{y}(\mathbf{G}, \mathbf{D})}\right)\right)\right] \quad \forall (\mathbf{G}, \mathbf{x}) \in \mathcal{M}, \tag{10b}$$

Upperbounds average reward of any policy

$$\mathbf{w} \in \mathbb{R}^K, \rho \in \mathbb{R}. \tag{10c}$$

Dual

$$(\mathrm{P}_\phi) \quad \sup_{\mu} \sum_{(\mathbf{G}, \mathbf{x}) \in M} \mathsf{E}_{\mathbf{D}}\left[\sum_{i \in \mathcal{I}} r^i x^i(\mathbf{D})\right] \mu(\mathbf{G}, \mathbf{x}), \tag{11a}$$

$$\sum_{(\mathbf{G}, \mathbf{x}) \in M} \left[\phi_k(\mathbf{G}) - \mathsf{E}_{\mathbf{D}}\left[\phi_k\left(\boldsymbol{\beta}\mathbf{G} + \mathbf{u}\left(\frac{\mathbf{x}(\mathbf{D})}{\mathbf{y}(\mathbf{G}, \mathbf{D})}\right)\right)\right]\right] \cdot \mu(\mathbf{G}, \mathbf{x}) = 0 \quad \forall k \in \mathcal{K}, \tag{11b}$$

$$\sum_{(\mathbf{G}, \mathbf{x}) \in M} \mu(\mathbf{G}, \mathbf{x}) = 1, \tag{11c}$$

$$\mu \geq 0, M = \mathrm{supp}(\mu) \subset \mathcal{M}, |M| < \infty, \tag{11d}$$

where $\mathrm{supp}(\cdot)$ denotes the support set.

Approximate DP Approach

Based on a polynomial approximation of the optimal bias function *h* of D_0

$$(\mathrm{D}_0) \quad \inf_{\rho, h} \rho \tag{3a}$$

$$h(\mathbf{G}) \geq \mathsf{E}_{\mathbf{D}}\left[\sum_{i \in \mathcal{I}} r^i x^i(\mathbf{D}) - \rho + h\left(\boldsymbol{\beta}\mathbf{G} + \mathbf{u}\left(\frac{\mathbf{x}(\mathbf{D})}{\mathbf{y}(\mathbf{G}, \mathbf{D})}\right)\right)\right] \quad \forall (\mathbf{G}, \mathbf{x}) \in \mathcal{M}, \tag{3b}$$

$$h \in \mathbb{B}(\mathcal{G}), \rho \in \mathbb{R}. \tag{3c}$$

Now suppose we are given a collection $\mathcal{K} = \{1, 2, \ldots, K\}$ of basis functions $\{\phi_k(\cdot)\}_{k \in \mathcal{K}}$, where $\phi_k \in \mathbb{B}(\mathcal{G})$. For a weight vector $\mathbf{w} = (w_1, w_2, \ldots, w_K) \in \mathbb{R}^K$, we can approximate the bias function as

$$h(\mathbf{G}) \approx \sum_{k \in \mathcal{K}} w_k \phi_k(\mathbf{G}) \quad = \sum_{i \in \mathcal{I}} \sum_{j=1}^{N} w_{ij} \cdot (G^i)^j$$

$$(\mathrm{D}_\phi) \quad \inf_{\rho, \mathbf{w}} \rho \tag{10a}$$

$$\sum_{k \in \mathcal{K}} w_k \phi_k(\mathbf{G}) \geq \mathsf{E}_{\mathbf{D}}\left[\sum_{i \in \mathcal{I}} r^i x^i(\mathbf{D}) - \rho + \sum_{k \in \mathcal{K}} w_k \phi_k\left(\boldsymbol{\beta}\mathbf{G} + \mathbf{u}\left(\frac{\mathbf{x}(\mathbf{D})}{\mathbf{y}(\mathbf{G}, \mathbf{D})}\right)\right)\right] \quad \forall (\mathbf{G}, \mathbf{x}) \in \mathcal{M}, \tag{10b}$$

Upperbounds average reward of any policy

$$\mathbf{w} \in \mathbb{R}^K, \rho \in \mathbb{R}. \tag{10c}$$

Dual

$$(\mathrm{P}_\phi) \quad \sup_{\mu} \sum_{(\mathbf{G}, \mathbf{x}) \in M} \mathsf{E}_{\mathbf{D}}\left[\sum_{i \in \mathcal{I}} r^i x^i(\mathbf{D})\right] \mu(\mathbf{G}, \mathbf{x}), \tag{11a}$$

$$\sum_{(\mathbf{G}, \mathbf{x}) \in M} \left[\phi_k(\mathbf{G}) - \mathsf{E}_{\mathbf{D}}\left[\phi_k\left(\boldsymbol{\beta}\mathbf{G} + \mathbf{u}\left(\frac{\mathbf{x}(\mathbf{D})}{\mathbf{y}(\mathbf{G}, \mathbf{D})}\right)\right)\right]\right] \cdot \mu(\mathbf{G}, \mathbf{x}) = 0 \quad \forall k \in \mathcal{K}, \tag{11b}$$

$$\sum_{(\mathbf{G}, \mathbf{x}) \in M} \mu(\mathbf{G}, \mathbf{x}) = 1, \tag{11c}$$

$$\mu \geq 0, M = \mathrm{supp}(\mu) \subset \mathcal{M}, |M| < \infty, \tag{11d}$$

where $\mathrm{supp}(\cdot)$ denotes the support set.

Solved using column generation for a given dual price vector and price

THEOREM 2 (DUALITY). *Suppose $\{\phi_k\}_{k \in \mathcal{K}}$ are continuous functions on $\mathcal{G}$. Then (P_ϕ) is solvable, and* $\sup (P_\phi) = \inf (D_\phi)$.

Optimal ADP Policy

- Given demand realization **D**, customers are in goodwill state **G**, optimal dual variables **w*** from solving (P_{phi}), the supplier solves

$$\begin{aligned} \max_{\mathbf{x}} \quad & \sum_{i \in \mathcal{I}} \left(r^i x^i + \sum_{k \in \mathcal{K}} w_k^* \phi_k \left(\boldsymbol{\beta} \mathbf{G} + \mathbf{u} \left(\frac{\mathbf{x}}{\mathbf{y}(\mathbf{G}, \mathbf{D})} \right) \right) \right) \\ \text{s.t.} \quad & \sum_{i \in \mathcal{I}} x^i \le \bar{X}, \quad 0 \le x^i \le y^i(G^i, D^i) \quad \forall i \in \mathcal{I}, \end{aligned}$$

or possibly a multiperiod lookahead version of this problem

Optimal ADP Policy

- Given demand realization **D**, customers are in goodwill state **G**, optimal dual variables **w*** from solving (P_{phi}), the supplier solves

$$\max_{\mathbf{x}} \sum_{i\in\mathcal{I}} \left(r^i x^i + \sum_{k\in\mathcal{K}} w_k^* \phi_k \left(\boldsymbol{\beta}\mathbf{G} + \mathbf{u}\left(\frac{\mathbf{x}}{\mathbf{y}(\mathbf{G},\mathbf{D})} \right) \right) \right)$$
$$\text{s.t.} \sum_{i\in\mathcal{I}} x^i \le \bar{X}, \quad 0 \le x^i \le y^i(G^i, D^i) \quad \forall i \in \mathcal{I},$$

or possibly a multiperiod lookahead version of this problem

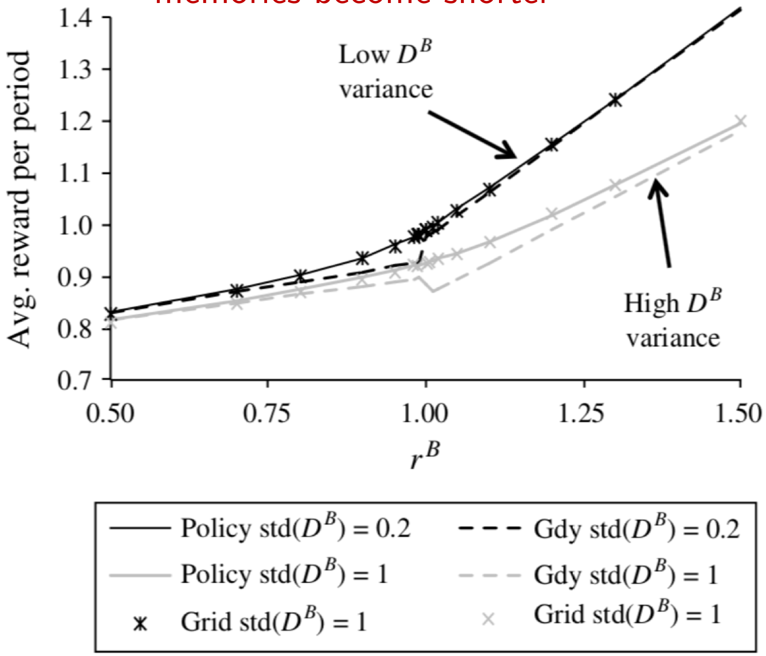
- Linear Approx. $u^i(f) = f$ and linear basis function $phi_i(G^i) = G^i$ Optimization becomes simple LP with an index solution: allocate capacity to customers in decreasing order of the customer indices

$$r^i + \frac{w_i^*}{y^i(G^i, D^i)}$$

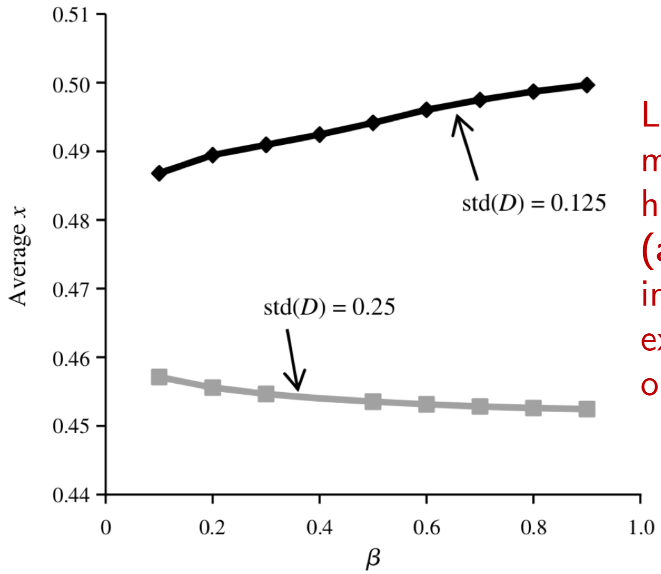
until capacity is exhausted.

Conclusions

Goodwills are important when
D variance increases and
memories become shorter

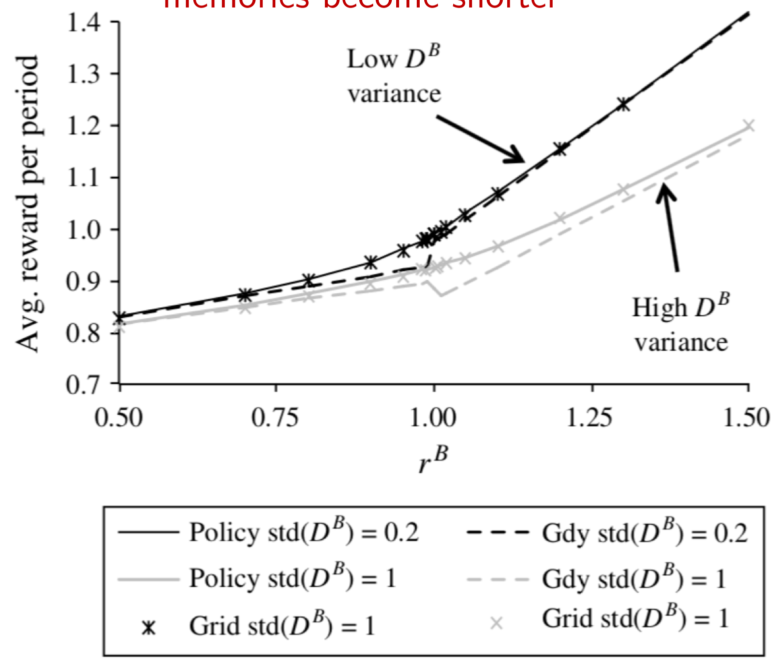


Conclusions



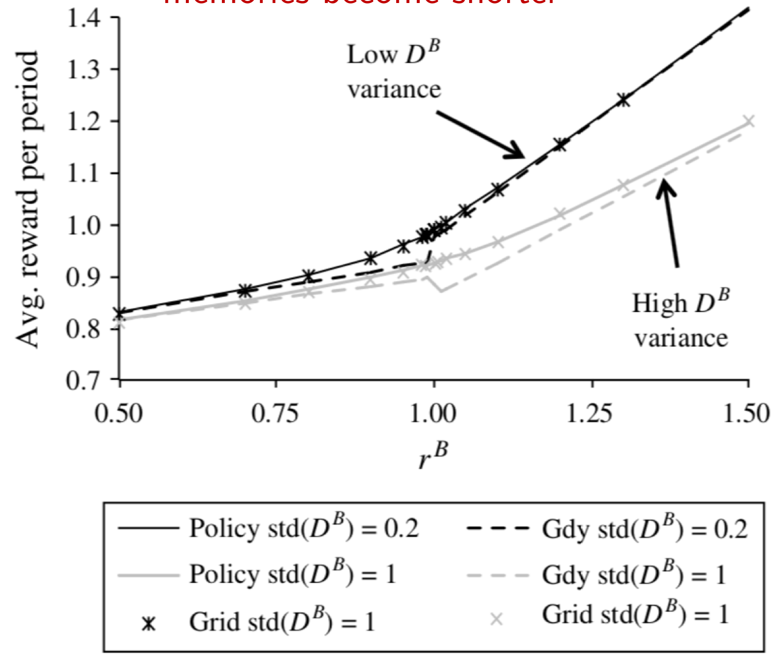
Longer memory can have + or − **(ambiguous)** impact on expected orders

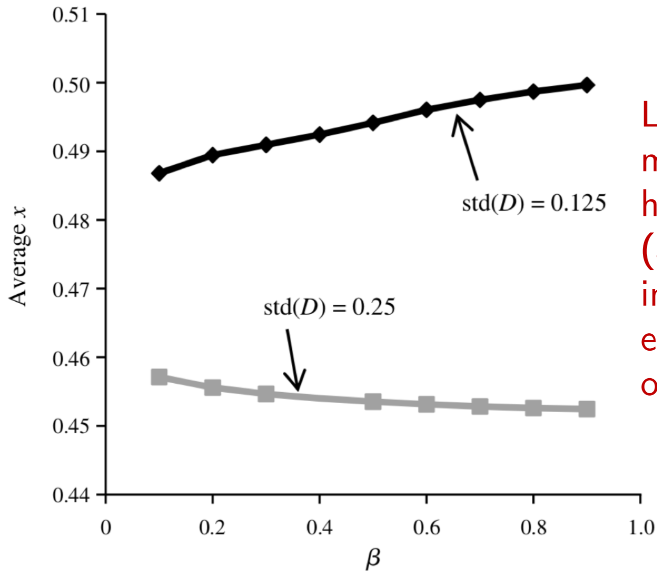
Goodwills are important when D variance increases and memories become shorter



Conclusions

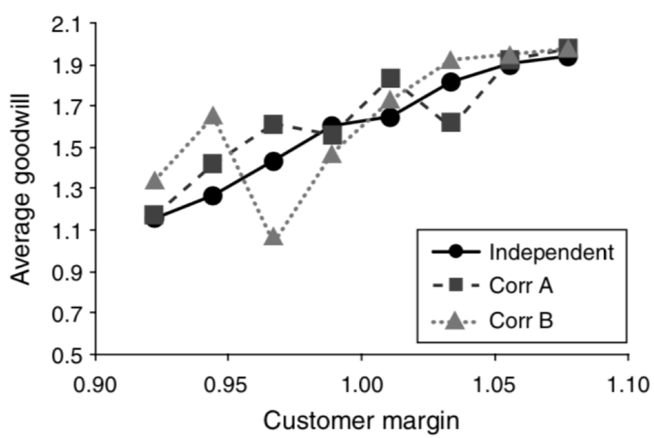
Goodwills are important when
D variance increases and
memories become shorter

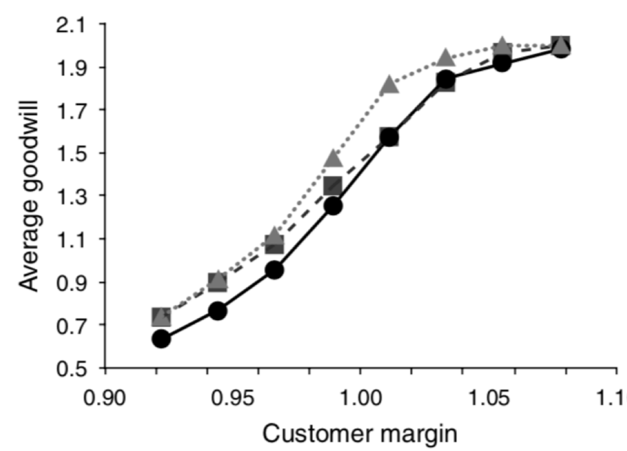




Longer memory can have + or − **(ambiguous)** impact on expected orders

ADP policy maintains more balanced goodwills
Goodwill-sensitive: prioritize customers by "adjusted margins"





▲	Name		Artist	Price	
1	Boom Boom Pow		Black Eyed Peas	$1.29	BUY SONG
2	Poker Face		Lady GaGa	$1.29	BUY SONG
3	Right Round		Flo Rida	$1.29	BUY SONG
4	The Climb		Miley Cyrus	$0.99	BUY SONG
5	Day 'n' Nite	EXPLICIT	Kid Cudi	$0.99	BUY SONG
6	Kiss Me Thru the Phone (feat. ...		Soulja Boy Tell 'Em	$1.29	BUY SONG
7	Gives You Hell		The All-America...	$1.29	BUY SONG
8	You Found Me		The Fray	$0.99	BUY SONG
9	Blame It (feat. T-Pain)	EXPLICIT	Jamie Foxx	$0.99	BUY SONG
10	Love Sex Magic (feat. Justin Ti...		Ciara	$1.29	BUY SONG
11	Just Dance		Lady GaGa & Col...	$1.29	BUY SONG
12	I Love College	EXPLICIT	Asher Roth	$0.99	BUY SONG
13	Halo		Beyoncé	$0.99	BUY SONG
14	Sugar (feat. Wynter)		Flo Rida	$1.29	BUY SONG
15	If U Seek Amy		Britney Spears	$0.99	BUY SONG
16	My Life Would Suck Without You		Kelly Clarkson	$1.29	BUY SONG
17	Dead and Gone (feat. J...	EXPLICIT	T.I.	$0.99	BUY SONG
18	Don't Trust Me	EXPLICIT	3OH!3	$0.99	BUY SONG
19	Jai Ho! (You Are My Destiny) [f...		A. R. Rahman & ...	$1.29	BUY SONG
20	Second Chance		Shinedown	$0.99	BUY SONG

Dynamic Pricing with Loss-Averse Consumers and Peak-End Anchoring

Nasiry & Popescu, Operations Research '11

▲	Name	Artist	Price	
1	Boom Boom Pow	Black Eyed Peas	$1.29	BUY SONG
2	Poker Face	Lady GaGa	$1.29	BUY SONG
3	Right Round	Flo Rida	$1.29	BUY SONG
4	The Climb	Miley Cyrus	$0.99	BUY SONG
5	Day 'n' Nite EXPLICIT	Kid Cudi	$0.99	BUY SONG
6	Kiss Me Thru the Phone (feat. ...	Soulja Boy Tell 'Em	$1.29	BUY SONG
7	Gives You Hell	The All-America...	$1.29	BUY SONG
8	You Found Me	The Fray	$0.99	BUY SONG
9	Blame It (feat. T-Pain) EXPLICIT	Jamie Foxx	$0.99	BUY SONG
10	Love Sex Magic (feat. Justin Ti...	Ciara	$1.29	BUY SONG
11	Just Dance	Lady GaGa & Col...	$1.29	BUY SONG
12	I Love College EXPLICIT	Asher Roth	$0.99	BUY SONG
13	Halo	Beyoncé	$0.99	BUY SONG
14	Sugar (feat. Wynter)	Flo Rida	$1.29	BUY SONG
15	If U Seek Amy	Britney Spears	$0.99	BUY SONG
16	My Life Would Suck Without You	Kelly Clarkson	$1.29	BUY SONG
17	Dead and Gone (feat. J... EXPLICIT	T.I.	$0.99	BUY SONG
18	Don't Trust Me EXPLICIT	3OH!3	$0.99	BUY SONG
19	Jai Ho! (You Are My Destiny) [f...	A. R. Rahman & ...	$1.29	BUY SONG
20	Second Chance	Shinedown	$0.99	BUY SONG

- In repeated purchases, consumers form price expectations.
- Operationalize RF in a dynamic pricing context and study how memory and anchoring process influences pricing strategies.
- Solves DP with nonsmooth reward and transition functions.

Modeling Reference Prices

- Utility = acquisition + transaction utilities **(mental accounting)**
- D increases in the reference gap $r - p$ **(prospect theory)**

$$d(p,r) = d_0(p) - \lambda(p-r)^+ + \gamma(r-p)^+$$
$$= \begin{cases} d_0(p) + \lambda(r-p), & \text{if } p \geqslant r; \\ d_0(p) + \gamma(r-p), & \text{if } p \leqslant r. \end{cases}$$

Loss aversion is captured by $\lambda \geqslant \gamma > 0$. We assume that the base demand, $d_0(p)$, is nonnegative, bounded, continuously differentiable, and decreasing in price, and the base profit $\pi_0(p) = p\, d_0(p)$ is nonmonotone and strictly concave. The firm's short-term profit is denoted $\pi(p,r) = p\, d(p,r)$. All our results extend for a nonzero marginal cost c.

Modeling Reference Prices

- Utility = acquisition + transaction utilities **(mental accounting)**
- D increases in the reference gap $r - p$ **(prospect theory)**

$$d(p,r) = d_0(p) - \lambda(p-r)^+ + \gamma(r-p)^+$$
$$= \begin{cases} d_0(p) + \lambda(r-p), & \text{if } p \geqslant r; \\ d_0(p) + \gamma(r-p), & \text{if } p \leqslant r. \end{cases}$$

Loss aversion is captured by $\lambda \geqslant \gamma > 0$. We assume that the base demand, $d_0(p)$, is nonnegative, bounded, continuously differentiable, and decreasing in price, and the base profit $\pi_0(p) = p\, d_0(p)$ is nonmonotone and strictly concave. The firm's short-term profit is denoted $\pi(p,r) = p\, d(p,r)$. All our results extend for a nonzero marginal cost c.

- **Peak-end rule**: Consumers anchor on a reference price = weighted average of lowest + most recent prices

Reference price

(0,1]: how much relying on low prices

$$r_t = \theta m_{t-1} + (1-\theta) p_{t-1}, \quad m_{t-1} = \min(m_{t-2}, p_{t-1})$$

Minimum price

Most recent price

The Model

- Firm maximizes infinite horizon B-discounted revenues:

$$J(m_0, p_0) = \max_{p_t \in \mathbf{P}} \sum_{t=1}^{\infty} \beta^{t-1} \pi(p_t, \theta m_{t-1} + (1-\theta) p_{t-1}),$$

$$\text{where } m_t = \min(m_{t-1}, p_t),$$

- Prices are bounded to [0, pbar] and at price pbar there is no demand.
- Implicitly assume that lowest prices can be remembered indefinitely. Insights remain valid even when consumers forget/update.
- **Bellman equation:**

$$J(m_{t-1}, p_{t-1}) = \max_{p_t \in \mathbf{P}} \Big\{ \pi(p_t, \theta m_{t-1} + (1-\theta) p_{t-1}) + \beta J(\min(p_t, m_{t-1}), p_t) \Big\}$$

Thanks to Loss Aversion…

- J(m,p) is increasing in both arguments.
  - Transitions and profit per stage are increasing. Also, intuitive.
- Loss aversion allows us to rewrite kinked profit pi(p,r) as the minimum of two smooth profit functions.

$$\pi_k(p, r) = [d_0(p) + k(r - p)]p = \pi_0(p) + k(r - p)p, \quad \text{for } k \in \{\lambda, \gamma\}$$

LEMMA 1. *The short-term profit is*

$$\pi(p, r) = \min(\pi_\lambda(p, r), \pi_\gamma(p, r))$$

and it is supermodular in (p, r).

Thanks to Loss Aversion…

- J(m,p) is increasing in both arguments.
  - Transitions and profit per stage are increasing. Also, intuitive.
- Loss aversion allows us to rewrite kinked profit pi(p,r) as the minimum of two smooth profit functions.

$$\pi_k(p, r) = [d_0(p) + k(r - p)]p = \pi_0(p) + k(r - p)p, \text{ for } k \in \{\lambda, \gamma\}$$

LEMMA 1. *The short-term profit is*

$$\pi(p, r) = \min(\pi_\lambda(p, r), \pi_\gamma(p, r))$$

and it is supermodular in (p, r).

- We can re-write the Bellman equation as:

$$J(m_{t-1}, p_{t-1}) = \max_{p_t \in \mathbf{P}} \{\min(\pi_\lambda, \pi_\gamma)(p_t, r_t) + \beta J(\min(p_t, m_{t-1}), p_t)\};$$
$$r_t = \theta m_{t-1} + (1 - \theta) p_{t-1}$$

  - Optimal pricing policy $p^*(m_{t-1},p_{t-1})$ solves the above equation.
  - (m,p) is a steady-state if p*(m,p) = p / optimal state path is constant.

Solution Approach

1. Characterize the **steady states**
   - Identify smooth upper-bound problems & characterize their steady states

$$J_m^\nu(p_{t-1}) = \max_{p_t \in \mathbf{P}} \{(1-\nu)\pi_\lambda(p_t, \theta m + (1-\theta)p_{t-1}) + \nu\pi_\gamma(p_t, p_{t-1}) + \beta J_m^\nu(p_t)\} \geq J(m,p)$$

   - Match steady states of original problem with those of these smooth relaxations and show that no other steady states exists

2. Characterize the **transient properties** of the optimal policy
   - Starting at any initial state (m_0,p_0), price paths remain in the same region as initial m_0
   - Within each region, the optimal price paths are monotone, so the prices converge monotonically to a steady state in the corresponding region.

Optimal Policy

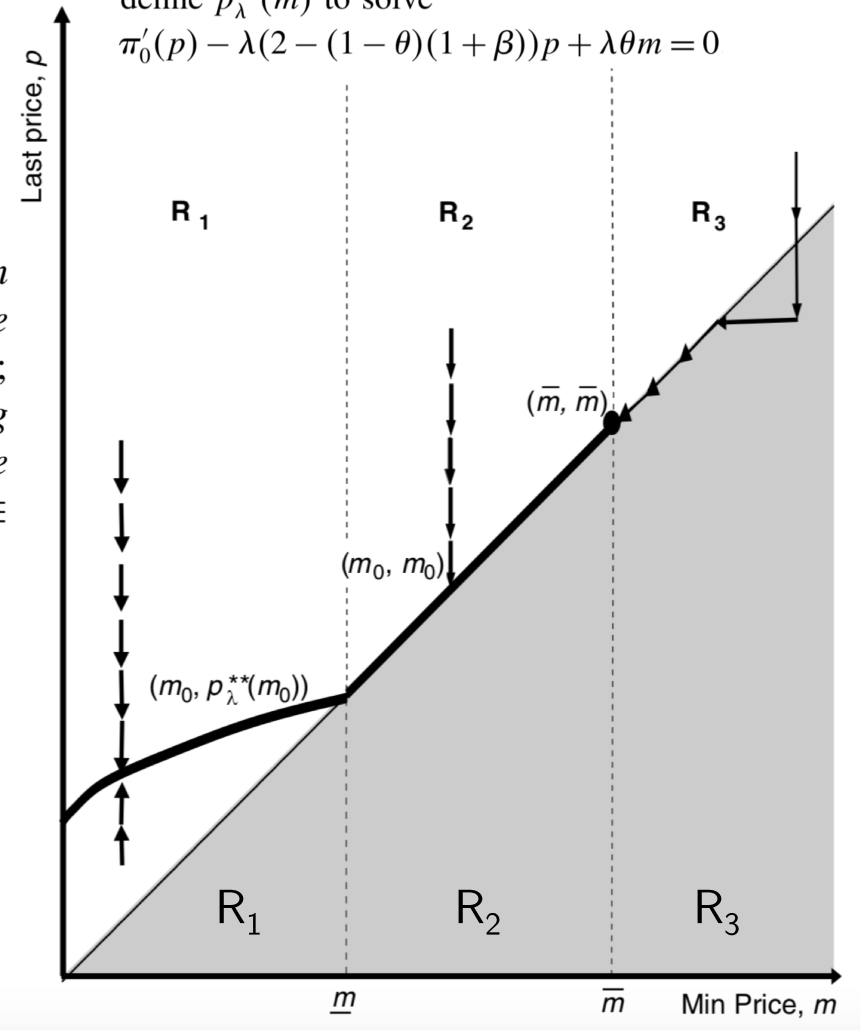
Optimal price path converges monotonically to a steady-state price (or is constant).

PROPOSITION 1. *Given any* (m_0, p_0), *the optimal price path of Problem* (3) *converges monotonically to a steady-state price, which is* (a) $p_\lambda^{**}(m_0)$, *if* $m_0 \in \mathbf{R}_1$; (b) m_0, *if* $m_0 \in \mathbf{R}_2$; *and* (c) $\overline{m}$, *if* $m_0 \in \mathbf{R}_3$. *In particular, the optimal pricing policy,* $p^*(m, p)$, *is increasing in both* m *and* p, *and the set of steady states of Problem* (3) *is* $\{(m, p_\lambda^{**}(m)) \mid m \in \mathbf{R}_1\} \cup \{(m, m) \mid m \in \mathbf{R}_2\}$.

Figure 1. Steady states (bold line) and generic optimal price paths (arrows) for each region of Problem (3).

define $p_\lambda^{**}(m)$ to solve

$\pi_0'(p) - \lambda(2-(1-\theta)(1+\beta))p + \lambda\theta m = 0$



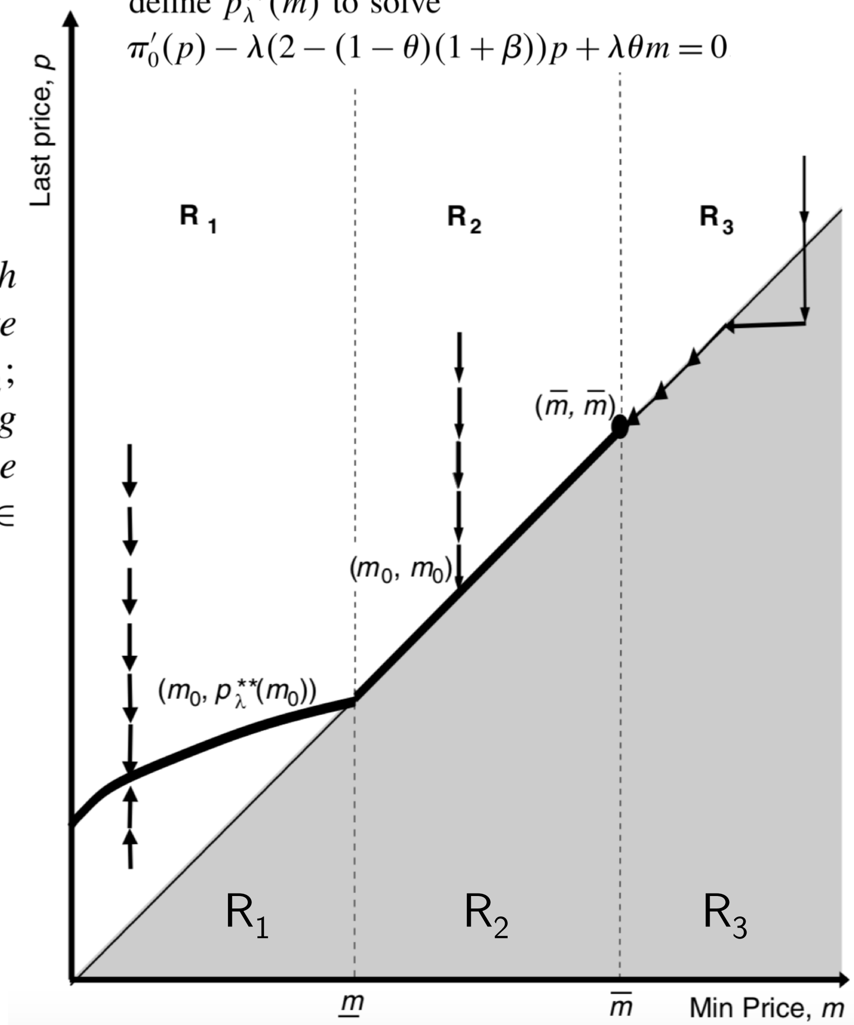
Optimal Policy

Optimal price path converges monotonically to a steady-state price (or is constant).

Proposition 1. *Given any* (m_0, p_0), *the optimal price path of Problem* (3) *converges monotonically to a steady-state price, which is* (a) $p_\lambda^{**}(m_0)$, *if* $m_0 \in \mathbf{R}_1$; (b) m_0, *if* $m_0 \in \mathbf{R}_2$; *and* (c) $\overline{m}$, *if* $m_0 \in \mathbf{R}_3$. *In particular, the optimal pricing policy,* $p^*(m, p)$, *is increasing in both* m *and* p, *and the set of steady states of Problem* (3) *is* $\{(m, p_\lambda^{**}(m)) \mid m \in \mathbf{R}_1\} \cup \{(m, m) \mid m \in \mathbf{R}_2\}$.

- Oscillating prices not LR optimal
- Steady-state price depends only on m_0
- Firm ignoring reference effects forgoes profit by overcharging customers.

Figure 1. Steady states (bold line) and generic optimal price paths (arrows) for each region of Problem (3).



Summary

Supply Chain

Capacity Allocation

Adelman & Mersereau
Management Science '13

Memory

- Memories matter particularly when # customers is small-moderate, high demand volatility, short memories.
- ADP policy dominates greedy policy by managing expectations, pooling demand volatilities
- When goodwill matters, mechanisms that link goodwill to profits.

Revenue Management

Dynamic Pricing

Nasiry & Popescu
Operations Research '11

Loss Aversion

- Operationalize peak-end rule in a dynamic pricing context.
- Constant pricing policy is optimal.
  - Similar to Su 2007, Gallego & van Ryzin 1994, and Celik et al 2009
- Behavioral regularities ensure prices will converge to steady-state (initial low expectation affects SS price)