Models of Learning in Various Games

Anton J. Kleywegt

School of Industrial and Systems Engineering
Georgia Institute of Technology
Atlanta, Georgia, USA

35th Conference on the Mathematics of Operations Research
Lunteren, The Netherlands

13 January 2010

Outline

Motivation

Basic questions for noncooperative games

- Under what conditions does a Nash equilibrium exist?
- Under what conditions does a unique Nash equilibrium exist?
- What is the usefulness of the Nash equilibrium concept?
  - Does knowledge of the Nash equilibrium give an advantage to players who know it over players who do not know it?
  - What is the predictive power of the Nash equilibrium?
    - If there are multiple Nash equilibria, which is the most likely outcome, if any?
    - If there is a unique Nash equilibrium, under what conditions is there reason to expect it to be an accurate prediction of the outcome of a real competitive situation?

Motivation

- Traditional game theory models assume players know all needed information, including their own objective function and how it depends on other players' decisions
- Players may attempt to learn
  - about other players
  - about other players' decisions
  - about correct structural form of model
  - about correct model parameters
- Learning dynamics: players make decisions, observe data, make decisions, etc.
- Players use observed data to learn about uncertain quantities

Learning and Games

- Discrete time $t = 0, 1, 2, \ldots$
- At each time t
  - each player i has some estimate $\widehat{H}_i^{t-1}$
  - each player i chooses a decision $x_i^t \in B_i(\widehat{H}_i^{t-1})$
  - each player i observes data $Y_i^t = F(x^t)$
  - each player i updates estimate $\widehat{H}_i^t = \phi(Y_i^0, \ldots, Y_i^t)$

Learning and Games

Questions

- Do the players' estimates $\widehat{H}_i^t$ converge, and if so, how does the limit compare with the correct parameter values (if the correct model is used)?
- Do the players' decisions x_i^t converge, and if so, how does the limit compare with the Nash equilibrium?

Motivation

- Most revenue management and dynamic pricing models make some restrictive assumptions:
  - Most models consider a single seller (monopolist), with demand that depends on prices of only the one seller
    - However, in many applications, demand depends on the prices of multiple sellers
  - In most models, demand at time t depends only on the prices at time t, and not on previous prices
    - If buyers forecast future prices to choose purchase time (strategic buyers), and forecasts depend on past prices, or if buyers form reference prices based on past prices, then demand depends not only on the current price, but also on past prices
- We want to investigate what happens if sellers use simple models, but the real process is more complicated

Bertrand Duopoly

- Two sellers, indexed -1 and 1
- Each seller i sells one product type, say product i
- Each seller i chooses the price p_i of product i
- Demand for product i is a function $d_i(p_i, p_{-i})$ of the prices chosen by both sellers
- Each seller i wants to maximize revenue $g_i(p_i, p_{-i}) := p_i d_i(p_i, p_{-i})$

Bertrand Duopoly

- Consider linear demand case: $d_i(p_i, p_{-i}) = \beta_{i,0} + \beta_{i,i} p_i + \beta_{i,-i} p_{-i}$ for $i = \pm 1$
- If $\beta_{i,i} < 0$, then given competitor's price p_{-i}, optimal price of seller i is

$$p_i = \arg\max_{p_i} g_i(p_i, p_{-i}) = -\frac{\beta_{i,0} + \beta_{i,-i} p_{-i}}{2\beta_{i,i}}$$

- If $4\beta_{-i,-i}\beta_{i,i} > \beta_{-i,i}\beta_{i,-i}$, then the unique Nash equilibrium is given by

$$p_i^* = \frac{\beta_{-i,0}\beta_{i,-i} - 2\beta_{i,0}\beta_{-i,-i}}{4\beta_{-i,-i}\beta_{i,i} - \beta_{-i,i}\beta_{i,-i}}$$

Bertrand Competition

- Case A: Seller i may not know which competitors potential buyers may choose from, or may not take all competitors into account in their models
  - Most revenue management models for use by seller i consider demand for product(s) i as an estimated function of prices of seller i only, and not as a function of prices of other sellers
  - Modeling error: demand model used by seller i does not have correct structural form — there do not exist parameter values that make model correct
- Case B: Seller i may use model with correct structural form, but may not know correct values of model parameters
- Case C: Seller i may use model with correct structural form, and may know correct values of model parameters, but may not know what prices other sellers will choose

Learning about Other Player's Decisions

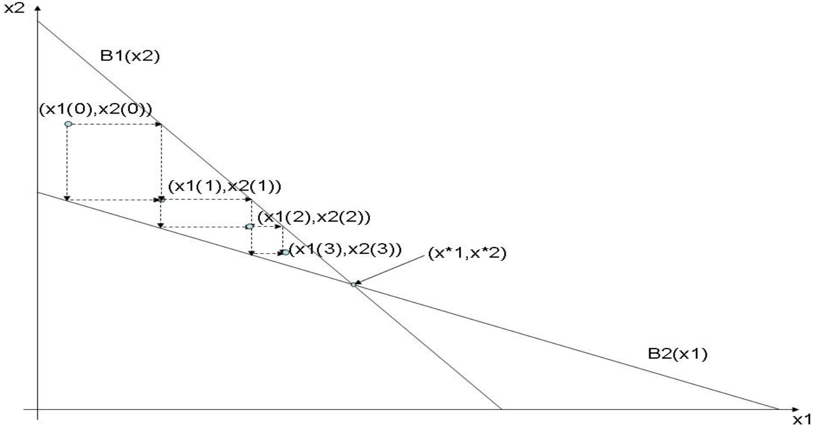
- Discrete time $t = 0, 1, 2, \ldots$
- At each time t
  - each player i has some estimate $\widehat{H}_i^{t-1}$ of the decisions of the other players
  - each player i chooses a decision $x_i^t \in B_i(\widehat{H}_i^{t-1})$
  - each player i observes decisions of other players $Y_i^t = x_{-i}^t$
  - each player i updates estimate $\widehat{H}_i^t = \phi(x_{-i}^0, \ldots, x_{-i}^t)$

Cournot Adjustment

- $\widehat{H}_i^t$ is a deterministic guess of the decisions of the other players
- Specifically, each player i observes data $Y_i^t = x_{-i}^t$
- Then each player i guesses that the other players will choose the same decision x_{-i}^t in period $t + 1$ as in period t: $\widehat{H}_i^t = \phi(x_{-i}^0, \ldots, x_{-i}^t) = x_{-i}^t$
- Each player i chooses a decision $x_i^{t+1} \in B_i(x_{-i}^t)$

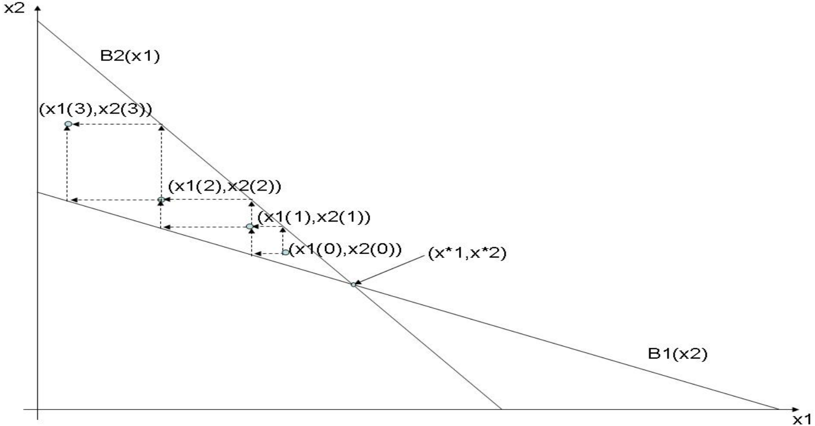
Cournot Adjustment

Example of convergent behavior:



Cournot Adjustment

Example of divergent behavior:



Cournot Adjustment

Example of cyclical behavior:

	A_3		B_3	
	A_2	B_2	A_2	B_2
A_1	1,1,-1	-1,-1,-1	1,-1,1	-1,1,1
B_1	-1,1,1	1,-1,1	-1,-1,-1	1,1,-1

- Process starts with $x^0 = (A_1, A_2, A_3)$
- Player 1,2 happy, player 3 changes, $x^1 = (A_1, A_2, B_3)$
- Player 1,3 happy, player 2 changes, $x^2 = (A_1, B_2, B_3)$
- Player 2,3 happy, player 1 changes, $x^3 = (B_1, B_2, B_3)$
- Player 1,2 happy, player 3 changes, $x^4 = (B_1, B_2, A_3)$
- Player 1,3 happy, player 2 changes, $x^5 = (B_1, A_2, A_3)$
- Player 2,3 happy, player 1 changes, $x^6 = (A_1, A_2, A_3)$

Bertrand Competition

Cournot Adjustment

- Consider linear demand case:
  $d_i(p_i, p_{-i}) = \beta_{i,0} + \beta_{i,i} p_i + \beta_{i,-i} p_{-i} + \varepsilon_i$ for $i = \pm 1$
- Each seller i uses the correct model with the correct parameter values $\beta_{i,0}$, $\beta_{i,i}$, $\beta_{i,-i}$
- At each time t, seller i chooses price

$$p_i^t = -\frac{\beta_{i,0} + \beta_{i,-i} p_{-i}^{t-1}}{2\beta_{i,i}}$$

- If $4\beta_{-i,-i}\beta_{i,i} > \beta_{-i,i}\beta_{i,-i}$, then $p_i^t \to p_i^*$ for each i, thus convergence to the unique Nash equilibrium

Fictitious Play

- $\widehat{H}_i^t$ is the empirical distribution of the decisions of the other players
- Specifically, each player i observes data $Y_i^t = x_{-i}^t$
- Then each player i guesses that the other players randomize their decisions x_{-i}^{t+1} in period $t+1$ according to the empirical distribution $\widehat{H}_i^t$
- Each player i chooses a decision

$$x_i^{t+1} \quad \in \quad B_i(\widehat{H}_i^t) \quad := \quad \arg\max_{x_i} \frac{1}{t+1} \sum_{\tau=0}^{t} g_i(x_i, x_{-i}^{\tau})$$

Bertrand Competition

Fictitious Play

- Consider linear demand case:
  $d_i(p_i, p_{-i}) = \beta_{i,0} + \beta_{i,i} p_i + \beta_{i,-i} p_{-i} + \varepsilon_i$ for $i = \pm 1$
- Each seller i uses the correct model with the correct parameter values $\beta_{i,0}$, $\beta_{i,i}$, $\beta_{i,-i}$
- At each time t, seller i chooses price

$$p_i^t = -\frac{\beta_{i,0} + \beta_{i,-i} \frac{1}{t} \sum_{\tau=0}^{t-1} p_{-i}^{\tau}}{2\beta_{i,i}}$$

- If $4\beta_{-i,-i}\beta_{i,i} > \beta_{-i,i}\beta_{i,-i}$, then $p_i^t \to p_i^*$ for each i, thus convergence to the unique Nash equilibrium

Learning about Model Parameter Values

- Discrete time $t = 0, 1, 2, \ldots$
- At each time t
  - each player i has some estimate $\widehat{H}_i^{t-1}$ of the parameters of the model of player i and of the decisions of the other players
  - each player i chooses a decision $x_i^t \in B_i(\widehat{H}_i^{t-1})$
  - each player i observes data, including decisions of other players $Y_i^t = F(x^t)$
  - each player i updates estimate $\widehat{H}_i^t = \phi(Y_i^0, \ldots, Y_i^t)$

Bertrand Competition

- Consider linear demand case:
  $d_i(p_i, p_{-i}) = \beta_{i,0} + \beta_{i,i}p_i + \beta_{i,-i}p_{-i} + \varepsilon_i$ for $i = \pm 1$
- Each seller i uses the correct model, but does not know the correct parameter values $\beta_{i,0}$, $\beta_{i,i}$, $\beta_{i,-i}$
- $\widehat{H}_i^{t-1}$ represents OLS estimates $\widehat{\beta}_{i,0}^{t-1}$, $\widehat{\beta}_{i,i}^{t-1}$, $\widehat{\beta}_{i,-i}^{t-1}$ of $\beta_{i,0}$, $\beta_{i,i}$, $\beta_{i,-i}$ and Cournot adjustment guess p_{-i}^{t-1} of p_{-i}^t
- At each time t, seller i chooses price

$$p_i^t = -\frac{\widehat{\beta}_{i,0}^{t-1} + \widehat{\beta}_{i,-i}^{t-1} p_{-i}^{t-1}}{2\widehat{\beta}_{i,i}^{t-1}}$$

- Questions:
  - Does $\widehat{\beta}^t$ converge, and if so, to β?
  - Does p_i^t converge, and if so, to p_i^*?

Simpler Case: Monopoly

- Consider linear demand case: $d(p) = \beta_0 + \beta_1 p + \varepsilon$
- Seller uses the correct model, but does not know the correct parameter values β_0, β_1
- $\widehat{H}^{t-1}$ represents OLS estimates $\widehat{\beta}_0^{t-1}$, $\widehat{\beta}_1^{t-1}$ of β_0, β_1
- At each time t, seller chooses price

$$p^t \quad = \quad -\frac{\widehat{\beta}_0^{t-1}}{2\widehat{\beta}_1^{t-1}}$$

- Questions:
  - Does $\widehat{\beta}^t$ converge, and if so, to β?
  - Does p^t converge, and if so, to $p^* := -\frac{\beta_0}{2\beta_1}$?

Simpler Case: Monopoly

- Taylor (1974) showed that if $\beta_0 \neq 0$ is known, and $\widehat{\beta}_1^{t-1}$ is the OLS estimate of β_1, and $p^0 \neq 0$, then w.p.1 $\widehat{\beta}_1^{t-1} \to \beta_1$ and $p^t \to p^*$
- Anderson and Taylor (1979) showed that if $[(X^t)^T X^t]^{-1} \to 0$ and $\lambda_{\max}^t = O(\lambda_{\min}^t)$ then w.p.1 $\widehat{\beta}^{t-1} \to \beta$ and $p^t \to p^*$
- Christopeit and Helmes (1980) relaxed the second condition to $(\lambda_{\max}^t)^{(1+\delta)/2} = O(\lambda_{\min}^t)$ for some $\delta > 0$
- Lai and Wei (1982) further relaxed the second condition to $\log(\lambda_{\max}^t) = o(\lambda_{\min}^t)$
- Lai and Wei (1982) also provided an example in which $\log(\lambda_{\max}^t)/\lambda_{\min}^t$ converges to a positive limit and $\widehat{\beta}^t$ converges to a limit that is not equal to β

Simpler Case: Monopoly

- Sternby (1977) considered a "Bayesian embedding"
  - It is shown that, w.p.1, $[(X^t)^T X^t]^{-1}$ converges, and $\widehat{\beta}^t$ converges
  - It is also shown that, w.p.1, $[(X^t)^T X^t]^{-1} \to 0$ if and only if $\widehat{\beta}^t \to \beta$, and a sufficient condition is given for this to occur
  - However, sufficient condition is in terms of process behavior and not process input
- Kumar (1990) and Chen and Hu (1998) used this approach to established consistency of various controls, and under additional conditions, consistency of parameter estimates
- Nassiri-Toussi and Ren (1994) showed examples for which estimates converge w.p.1 to incorrect values or w.p.1 do not converge
- We do not know in general when monopoly (duopoly) estimates and controls converge

Learning with Modeling Error

- Discrete time $t = 0, 1, 2, \ldots$
- At each time t
  - each player i has some estimate $\widehat{H}_i^{t-1}$ of the parameters of the (incorrect) model of player i
  - each player i chooses a decision $x_i^t \in \widehat{B}_i(\widehat{H}_i^{t-1})$
  - each player i observes data, including decisions of other players $Y_i^t = F(x^t)$
  - each player i updates estimate $\widehat{H}_i^t = \phi(Y_i^0, \ldots, Y_i^t)$

Bertrand Competition

- Consider linear demand case:
  $d_i(p_i, p_{-i}) = \beta_{i,0} + \beta_{i,i} p_i + \beta_{i,-i} p_{-i} + \varepsilon_i$ for $i = \pm 1$
- Each seller i uses a model of own demand as a function of own price only
  $\delta_i(p_i) = \alpha_{i,0} + \alpha_i p_i$
- $\widehat{H}_i^{t-1}$ represents estimates $\widehat{\alpha}_{i,0}^{t-1}$, $\widehat{\alpha}_i^{t-1}$ of $\alpha_{i,0}$, α_i
- At each time t, seller i chooses price

$$p_i^t = -\frac{\widehat{\alpha}_{i,0}^{t-1}}{2\widehat{\alpha}_i^{t-1}}$$

- Questions:
  - Does $\widehat{\alpha}^t$ converge?
  - Does p_i^t converge, and if so, to p_i^*?

Bertrand Competition

Slope known

- Each seller i uses $\alpha_i = \beta_{i,i}$
- $\widehat{H}_i^{t-1}$ represents estimate $\widehat{\alpha}_{i,0}^{t-1}$ of $\alpha_{i,0}$
- At each time t, seller i chooses price

$$p_i^t = -\frac{\widehat{\alpha}_{i,0}^{t-1}}{2\alpha_i}$$

- Results:
  - Estimates

$$\widehat{\alpha}^t \rightarrow \alpha_i^* := \frac{2\beta_{i,i}\left(2\beta_{i,0}\beta_{-i,-i} - \beta_{-i,0}\beta_{i,-i}\right)}{4\beta_{i,i}\beta_{-i,-i} - \beta_{-i,i}\beta_{i,-i}}$$

  - Controls $p_i^t \rightarrow p_i^*$
  - $d_i(p_i, p_{-i}) = \beta_{i,0} + \beta_{i,i}p_i + \beta_{i,-i}p_{-i} + \varepsilon_i$ and $\delta_i(p_i) = \beta_{i,0} + \beta_{i,i}p_i + \beta_{i,-i}p_{-i}^*$

Bertrand Competition

Intercept known

- Each seller i uses $\alpha_{i,0} = \beta_{i,0}$
- $\widehat{H}_i^{t-1}$ represents estimate $\widehat{\alpha}_i^{t-1}$ of α_i
- At each time t, seller i chooses price

$$p_i^t = -\frac{\alpha_{i,0}}{2\widehat{\alpha}_i^{t-1}}$$

- Results for symmetric case:
  - Estimates $\widehat{\alpha}^t \to \beta_{i,i} + \beta_{i,-i}$
  - Controls $p_i^t \to -\frac{\beta_{i,0}}{2(\beta_{i,i}+\beta_{i,-i})}$
  - Resulting revenue is the same as though sellers collude to maximize the sum of their revenues

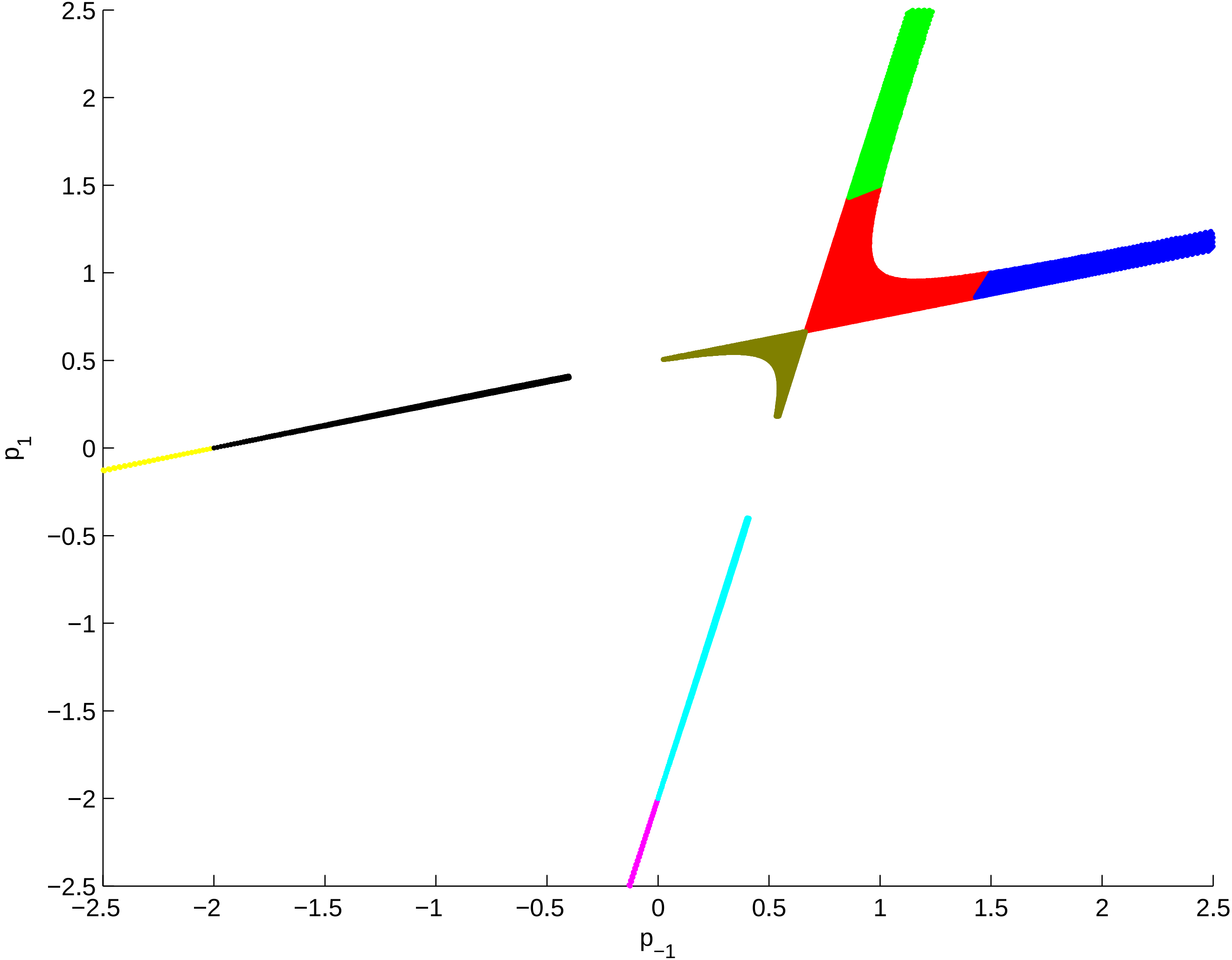
Bertrand Competition

Neither intercept nor slope known

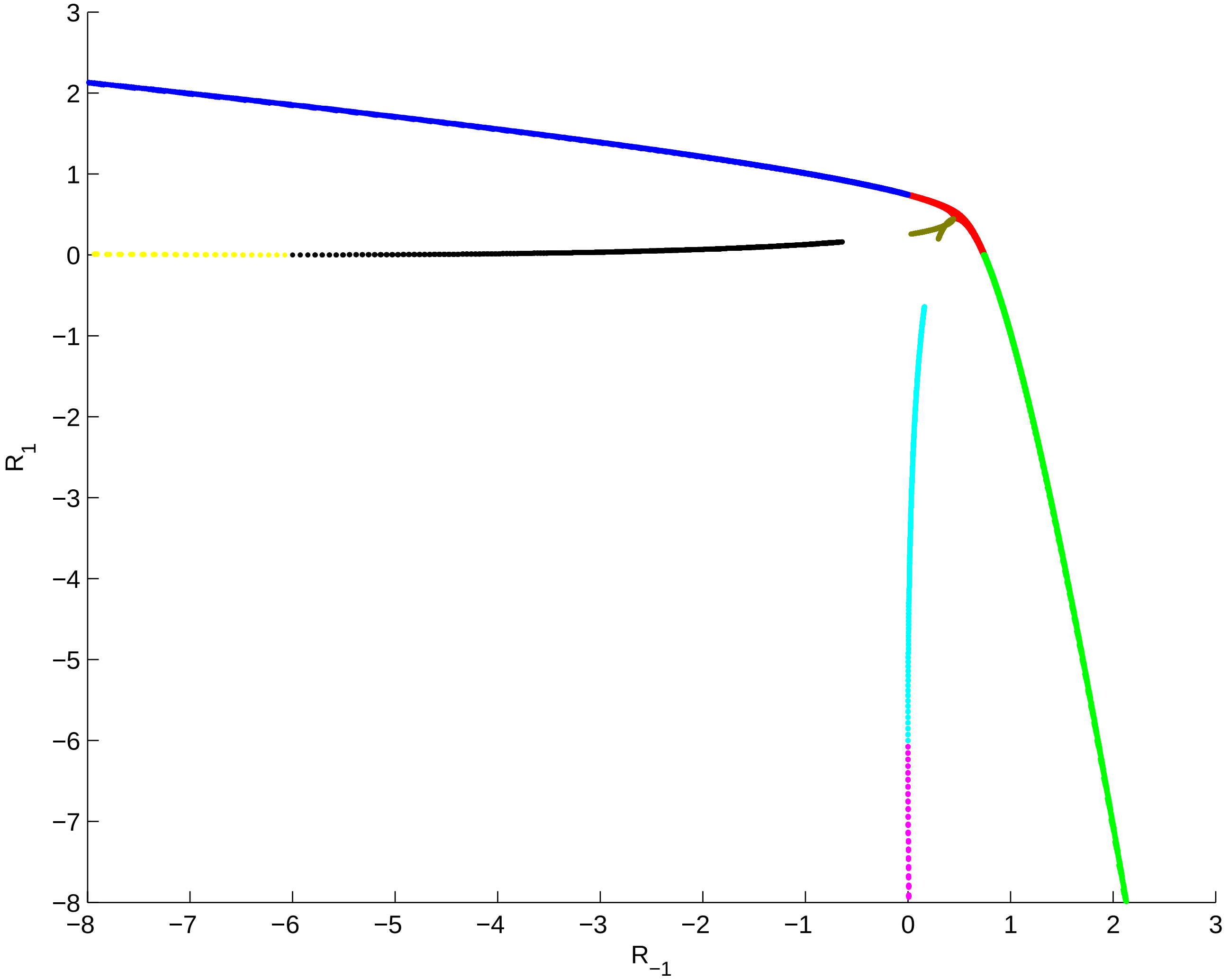
- $\widehat{H}_i^{t-1}$ represents estimate $\widehat{\alpha}_{i,0}^{t-1}$, $\widehat{\alpha}_i^{t-1}$ of $\alpha_{i,0}$, α_i
- At each time t, seller i chooses price

$$p_i^t = -\frac{\widehat{\alpha}_{i,0}^{t-1}}{2\widehat{\alpha}_i^{t-1}}$$

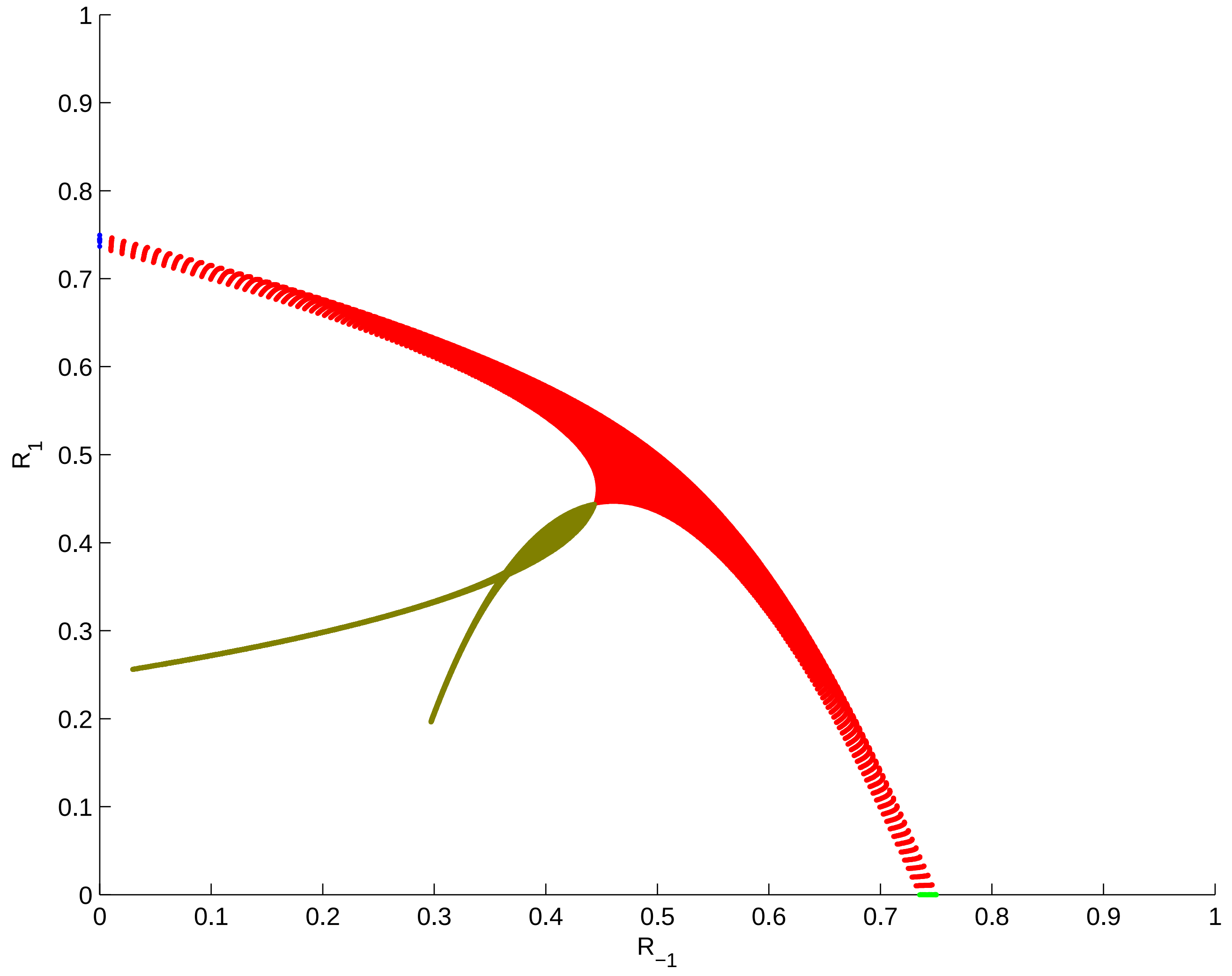
- Questions of convergence studied since 1970's (Kirman 1975, 1983, 1992, 1995)



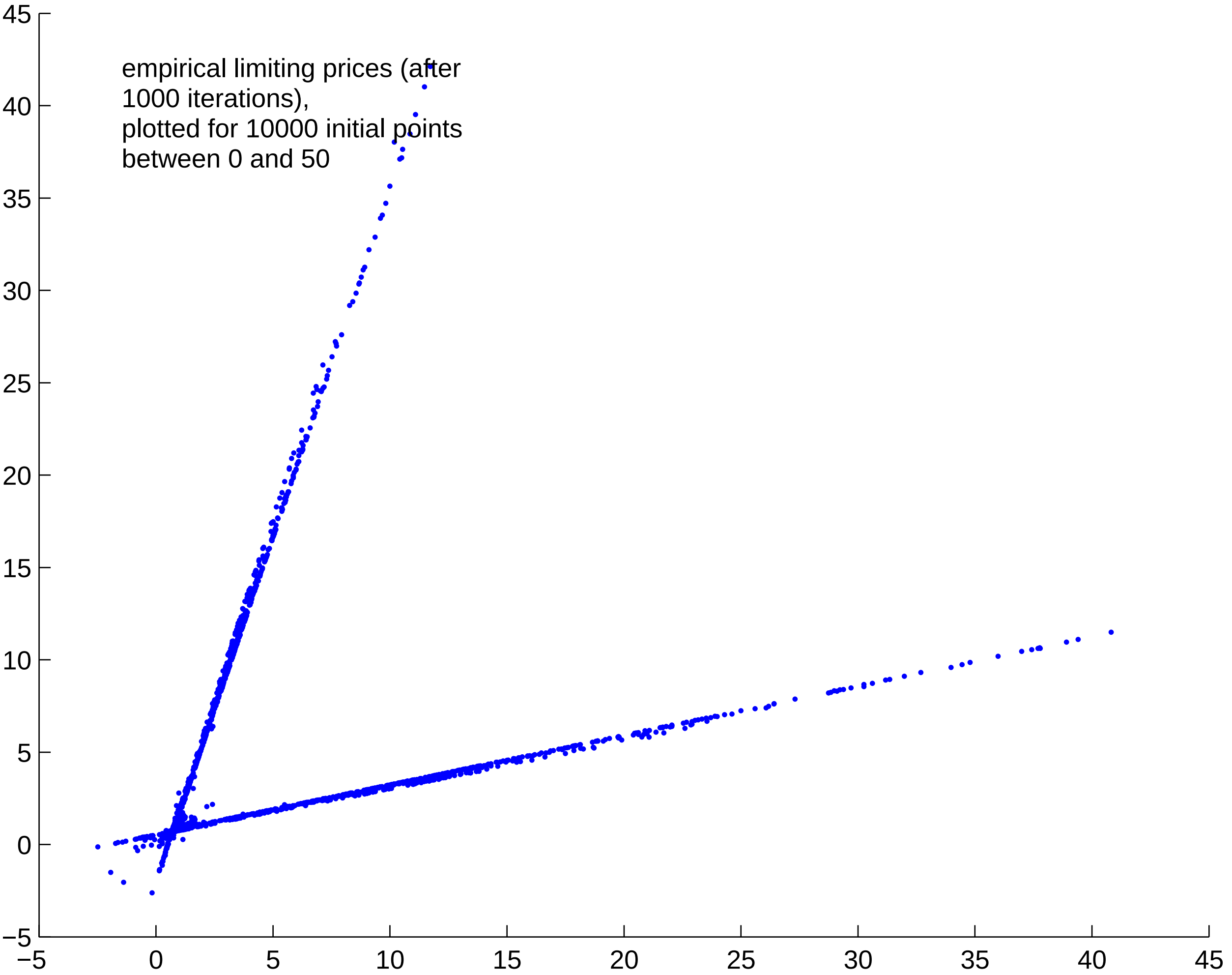
2.5
2
1.5
1
0.5
0
−0.5
−1
−1.5
−2
−2.5
−2.5
−2
−1.5
−1
−0.5
0
0.5
1
1.5
2
2.5
p_1
p_{-1}

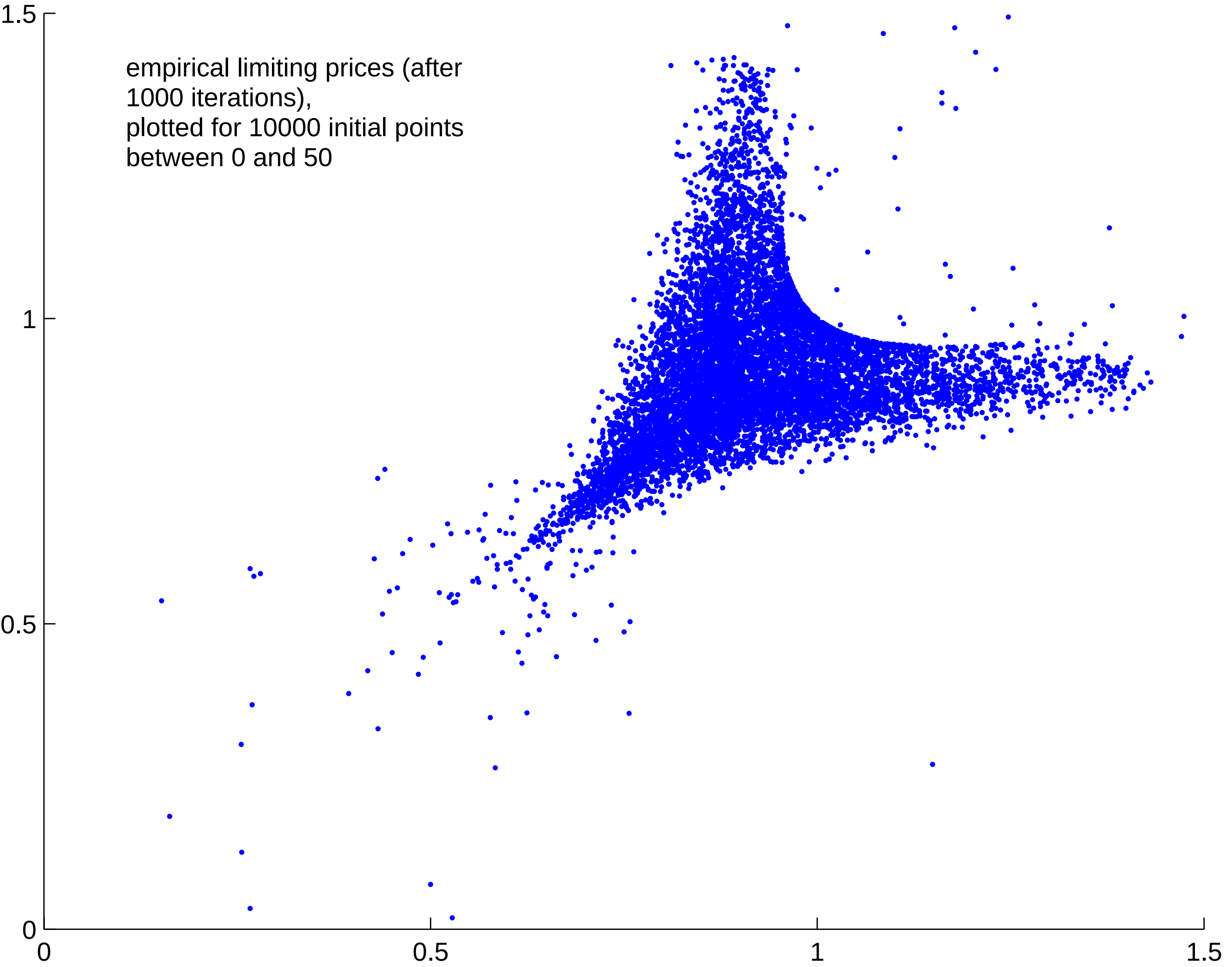


3
2
1
0
−1
−2
−3
−4
−5
−6
−7
−8
R_1
−8
−7
−6
−5
−4
−3
−2
−1
0
1
2
3
R_−1



1
0.9
0.8
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0
R_1
0
0.1
0.2
0.3
0.4
0.5
0.6
0.7
0.8
0.9
1
R_{-1}


empirical limiting prices (after 1000 iterations), plotted for 10000 initial points between 0 and 50
45
40
35
30
25
20
15
10
5
0
−5
−5
0
5
10
15
20
25
30
35
40
45


empirical limiting prices (after
1000 iterations),
plotted for 10000 initial points
between 0 and 50
0
0.5
1
1.5
0
0.5
1
1.5

Strategic Buyers and Learning

- Most dynamic pricing models consider demand that depends on price at the same time only
  - Buyer behavior is not affected by past prices
  - Buyers do not try to take into account future prices
- In most practical settings, buyers are more sophisticated than in the models

Strategic Buyers and Learning

- Some recent research considers models of buyers who take into account future prices, either through exogenous probability distribution of future prices, or through pricing policy of seller
  - probability distribution of future prices is exogenous (does not depend on seller or buyer decisions), and
  - buyers know the correct probability distribution of future prices, or
  - buyers know the pricing policy of seller over entire time horizon
- In most practical settings, buyers are less sophisticated than in the models

Some Related Literature

- Peleg, Lee and Hausman (2002): Buyer who can buy product at contract price and also can buy it from online market with exogenous price.
- Wu, Kleindorfer and Zhang (2002): Buyer who can buy option for product and also can buy it from spot market with exogenous price. Buyer knows seller's option pricing policy and also seller knows the buyer's decision policy.
- Levin, McGill and Nediak (2006): Buyer who knows seller's pricing policy over entire time horizon, and who dynamically decides when to purchase. Establishes sufficient condition for equilibrium.
- Ovchinnikov and Milner (2006): Buyer whose purchasing decision is made based on seller's decision and buyer's own decision one period ago.

Strategic Buyers and Learning

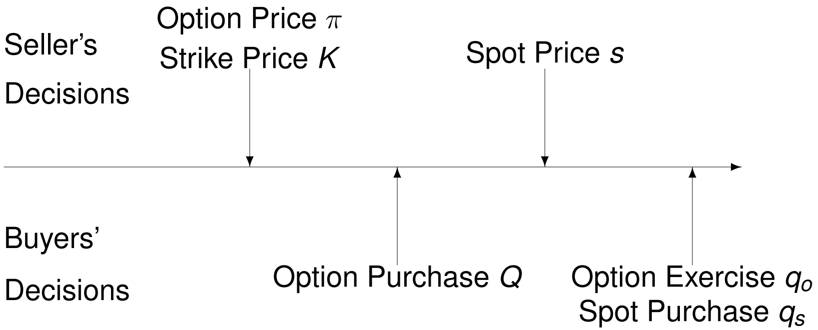
We would like to consider models in which buyers are more sophisticated than in traditional revenue management models, but

- spot price distribution may depend on seller and buyer decisions (not exogenous)
- buyers do not know spot price distribution
- buyers do not know pricing policy of seller over entire time horizon
- buyers observe spot price data and attempt to learn spot price distribution

Strategic Buyers and Learning

- Single product type
- Single seller
- Multiple homogeneous buyers
- Buyers can buy product through two channels: option contracts and spot market
    - If a buyer buys options, then in next stage buyer has right to purchase product at specified strike price
    - Also, buyer can purchase product in next stage at spot price

Timeline of Decisions



Seller and Buyer Behavior

- Seller's objective to maximize expected profit
- Buyers' utility function, $U(q)$, is quadratic in the amount q of product obtained:

$$U(q) := -\frac{1}{2b}q^2 + \frac{a}{b}q$$

(This is the revenue function of a retailer who faces a linear demand function $D = 2a - 2bp$ and who sells q units of product)

- Demand level a is random and b is deterministic, and both seller and buyers observe a after the buyers' first decision and before seller's second decision

Buyers' Second Decision

- Buyers know a, b
- Option price π, strike price K, number Q of options bought, spot price s, already decided
- Buyers decide

$$q_o(a,s,Q) := \text{Number of options out of } Q \text{ exercised}$$
$$q_s(a,s,Q) := \text{Quantity of products bought at spot price } s$$

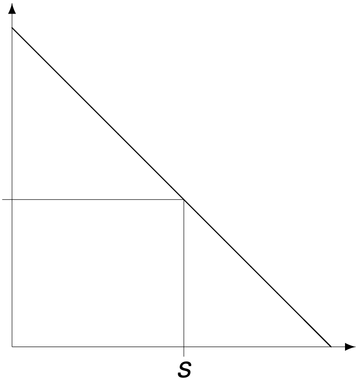
- Buyers choose q_o, q_s by solving

$$\max_{0 \le q_o \le Q, 0 \le q_s} \{U(q_o + q_s) - Kq_o - sq_s\}$$

Buyers' Second Decision

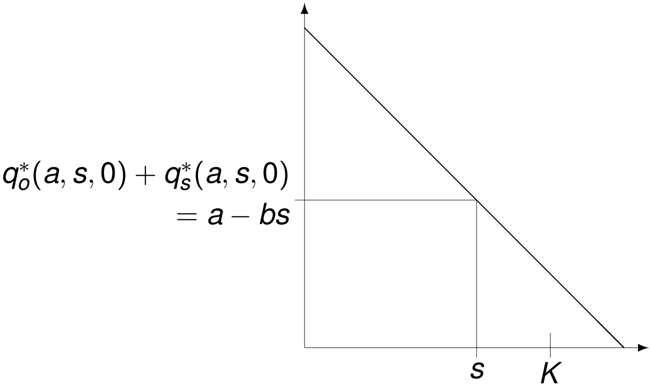
Without option contracts ($Q = 0$),
$q_o^*(a, s, 0) = 0$, $q_s^*(a, s, 0) = a - bs$

$$q_o^*(a, s, 0) + q_s^*(a, s, 0) = a - bs$$



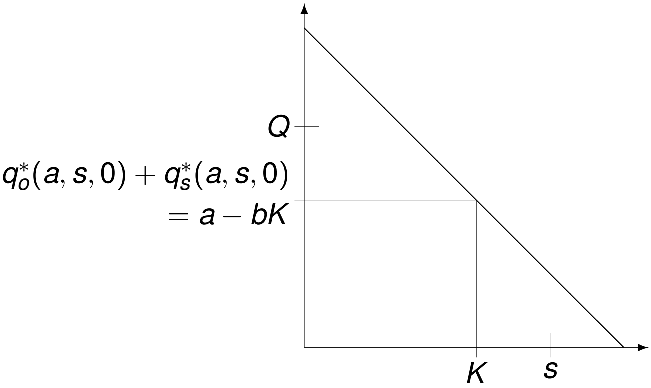
Buyers' Second Decision

With $Q > 0$, and $K \geq s$,
$q_o^*(a, s, Q) = 0$, $q_s^*(a, s, Q) = a - bs$



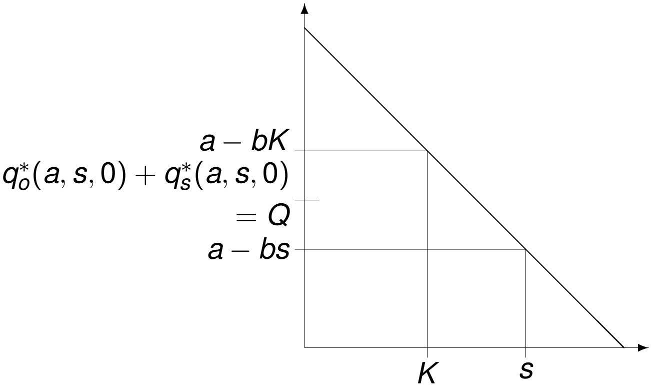
Buyers' Second Decision

With $Q > 0$, $K < s$, and $Q \geq a - bK > a - bs$,
$q_o^*(a, s, Q) = a - bK$, $q_s^*(a, s, Q) = 0$



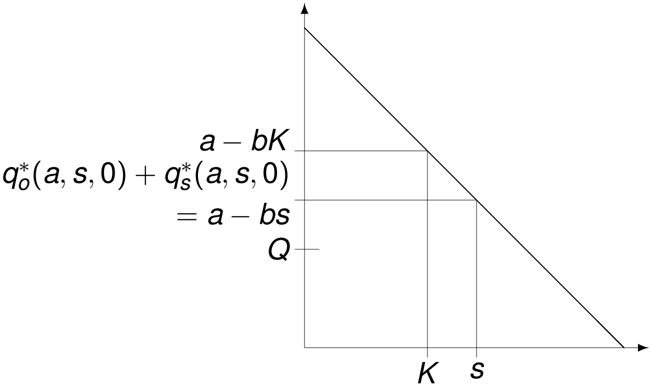
Buyers' Second Decision

With $Q > 0$, $K < s$, and $a - bK \geq Q \geq a - bs$,
$q_o^*(a, s, Q) = Q$, $q_s^*(a, s, Q) = 0$



Buyers' Second Decision

With $Q > 0$, $K < s$, and $a - bK > a - bs > Q$,
$q_o^*(a, s, Q) = Q$, $q_s^*(a, s, Q) = a - bs - Q$



Seller's Second Decision

- Seller knows a, b
- Option price π, strike price K, number Q of options bought, already decided
- Seller knows buyers' aggregate behavior
- Seller chooses spot price s by solving

$$\max_{s \geq 0} \left\{ K q_o^*(a, s, Q) + s q_s^*(a, s, Q) - c[q_o^*(a, s, Q) + q_s^*(a, s, Q)] \right\}$$

- Chosen spot price is function $s^*(a, Q)$ of a and Q
- For $Q_1 \leq Q_2$, and any a,

$$s^*(a, Q_1) \quad \geq \quad s^*(a, Q_2)$$

Buyers' First Decision

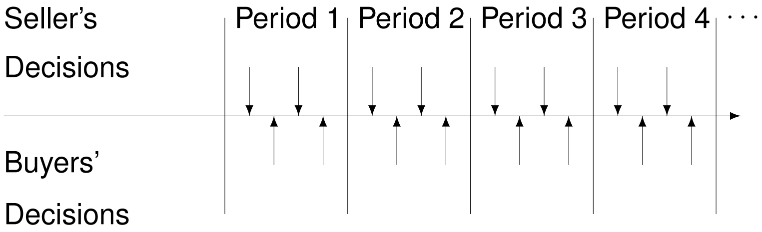
- Buyers uncertain regarding future value of random variable a
- Seller's second decision $s^*(a, Q)$ depends on a
- Buyers also uncertain regarding future decision of seller
- In this study, buyers not just uncertain regarding future decision of seller because of uncertainty in future value of a, but buyers do not know function $s^*(a, Q)$
- In particular, buyers do not know dependence of $s^*(a, Q)$ on Q, because each buyer may contribute very little to Q, so that dependence of $s^*(a, Q)$ on each buyer's Q-decision is insignificant

Buyers' First Decision

- Buyers attempt to learn a probability distribution for future decision of seller as though seller's future decisions are exogenously distributed (independent of buyers' first decision)

Sequence of Replications

Consider sequence of 2-stage problems



Buyers' First Decision

- Buyers observe spot price s_t in each period t
- Buyers forecast spot price distribution with empirical distribution constructed by observed previous spot prices:

$$H_n(p) \quad := \quad \frac{1}{n}\sum_{t=1}^{n} \mathbb{I}_{\{s_t \leq p\}}$$

Buyers' First Decision

- Buyers know probability distribution F (finite mean) of a, b
- Option price π, strike price K, already decided
- In period n, buyers forecast spot price s with empirical distribution H_n
- Buyers decide number Q of options to buy
- Buyers choose Q by solving

$$\max_{Q\geq 0} \quad \Big\{ B(\pi, K, H_n, Q) \;:=\; \int_0^\infty \int_0^\infty \big[U(q_o^*(a,s,Q) + q_s^*(a,s,Q)) - \pi Q - K q_o^*(a,s,Q) - s q_s^*(a,s,Q) \big] \, dH_n(s) dF(a) \Big\}$$

Buyers' First Decision

- It follows that

$$\frac{\partial B(\pi, K, H_n, Q)}{\partial Q} = -\pi + \int_{Q+bK}^{\infty} \int_{0}^{\infty} \left[(s-K)^+ - \left(s - \frac{a-Q}{b} \right)^+ \right] dH_n(s) dF(a)$$

- $\partial B(\pi, K, H_n, Q)/\partial Q$ is Lipschitz continuous and nonincreasing in Q
- The optimizer Q_n^* to buyer's optimization problem follows:
  - If $\partial B(\pi, K, H_n, 0)/\partial Q < 0$, then $Q_n^* = 0$
  - If $\partial B(\pi, K, H_n, 0)/\partial Q \geq 0$, then $Q_n^* \geq 0$

Long-Run Behavior

Note that

- Number Q_n^* of options is chosen based on empirical distribution of previous spot prices $s_1, \ldots, s_n$,

$$Q_n^* \quad = \quad g(s_1, s_2, \ldots, s_n))$$

- Given $s_1, \ldots, s_n$ (and thus H_n and Q_n^*), actual distribution $A(\cdot; Q_n^*)$ of next spot price s_{n+1} is the distribution of $s^*(a_{n+1}, Q_n^*)$
- Even if sequence $\{a_n\}$ is independent and identically distributed, spot price sequence $\{s_n\}$ is neither independent nor identically distributed,

$$s_{n+1} \quad = \quad s^*(a_{n+1}, g(s_1, s_2, \ldots, s_n))$$

Long-Run Behavior

Questions:

- Do the sequences $\{H_n\}$, $\{Q_n^*\}$, $\{A(s; Q_n^*)\}$ converge?
- If so, can we characterize the limits, and how are the limits related to each other?
- How do the limits compare with an equilibrium in the case in which buyer knows seller's behavior in the second stage (given by function $s^*(a, Q)$)?

Convergence Results

Suppose that $\{Q_n^*\}$ converges to Q^*. Then, $\{A(\cdot; Q_n^*)\}$ converges weakly to $A(\cdot; Q^*)$, where

$$A(s; Q) \quad := \quad \int_0^\infty \mathbb{I}_{\{s^*(a,Q) \leq s\}} dF(a)$$

is actual spot price distribution given Q

Convergence Results

If $\{A(\cdot; Q_n^*)\}$ converges weakly to $A(\cdot; Q^*)$, then, $\{H_n\}$ converges weakly to $A(\cdot; Q^*)$ w.p.1

Existence and Uniqueness of Limit

- The function

$$\nabla_Q B(\pi, K, A(\cdot; Q), Q) := -\pi + \int_{Q+bK}^{\infty} \int_0^{\infty} \left[(s-K)^+ - \left(s - \frac{a-Q}{b} \right)^+ \right] A(ds; Q) dF(a)$$

  is Lipschitz continuous and strictly decreasing in Q
- Thus, either $\nabla_Q B(\pi, K, A(\cdot; 0), 0) < 0$, or there exists a unique $Q^* \geq 0$ such that

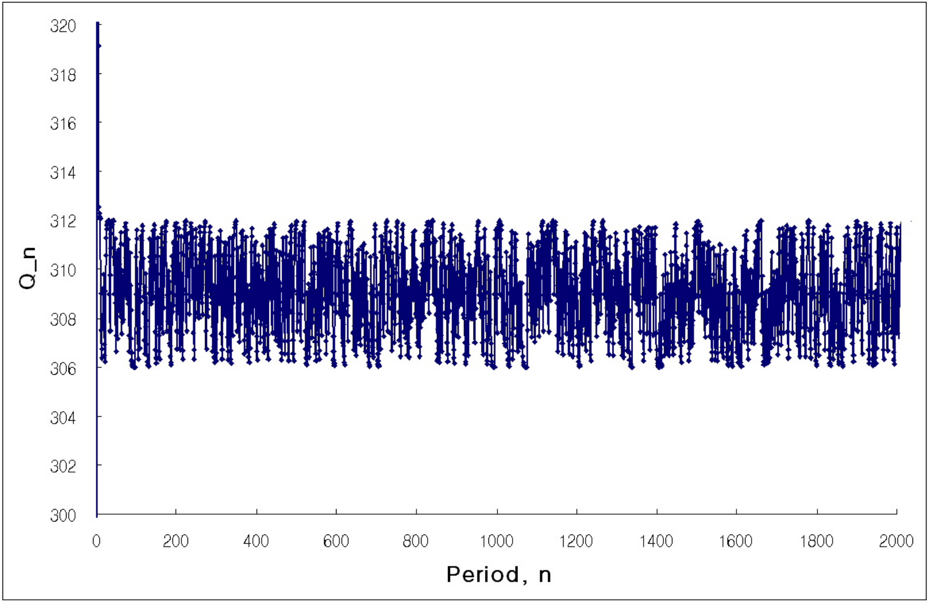
$$\nabla_Q B(\pi, K, A(\cdot; Q^*), Q^*) = -\pi + \int_{Q^*+bK}^{\infty} \int_0^{\infty} \left[(s-K)^+ - \left(s - \frac{a-Q^*}{b} \right)^+ \right] A(ds; Q^*) dF(a) = 0$$

Convergence Results

- If $\nabla_Q B(\pi, K, A(\cdot; 0), 0) < 0$, then w.p.1, $Q_n^* \to 0$, $A(\cdot; Q_n^*) \to A(\cdot; 0)$, $H_n \to A(\cdot; 0)$
- Otherwise, there exists $Q^* \geq 0$ such that $\nabla_Q B(\pi, K, A(\cdot; Q^*), Q^*) = 0$, and then w.p.1, there exists an interval $[Q_L, Q_U]$ such that $Q^* \in [Q_L, Q_U]$, and for all $Q \in [Q_L, Q_U]$ it holds that

$$\frac{\partial B(\pi, K, A(\cdot; Q^*), Q)}{\partial Q} = 0$$

  Also, for almost all $\omega \in \Omega$ and any $\varepsilon > 0$, there exists $N(\omega, \varepsilon) < \infty$ such that $Q_n \in [Q_L - \varepsilon, Q_U + \varepsilon]$ for all $n \geq N(\omega, \varepsilon)$, that is, w.p.1, $Q_n \to [Q_L, Q_U]$



Q_n
320
318
316
314
312
310
308
306
304
302
300
0
200
400
600
800
1000
1200
1400
1600
1800
2000
Period, n

Convergence Results

If $Q^* = Q_L = Q_U$, then w.p.1, $Q_n^* \to Q^*$

Existence of Full Information Equilibrium

Suppose that the buyer knows the seller's spot pricing policy $s^*(a, Q)$. Then the buyer decides to buy a quantity Q^E of options, such that

$$\begin{aligned} Q^E &= 0 \quad \text{if} \quad -\pi + \int_{2bK}^{\infty} \left[\frac{3}{4}\frac{a}{b} - K \right] dF(a) < 0 \\ Q^E &\geq 0 \quad \text{if} \quad -\pi + \int_{2bK}^{\infty} \left[\frac{3}{4}\frac{a}{b} - K \right] dF(a) \geq 0 \end{aligned}$$

Comparison of Limit and Full Information Equilibrium

$$Q^E \geq Q^*$$

Summary

- Setting with buyer learning of spot price distribution using empirical distribution is considered
- Spot prices are endogenous
- Spot price sequence is not i.i.d.
- Convergence of sequence $\{(Q_n^*, A(\cdot, Q_n^*), H_n)\}$ characterized for fixed π, K
- Limit compared with full information equilibrium

Summary

- Considered learning in various games without and with modeling error
  - Competitive dynamic pricing problem
  - Buyer buying options and on spot market
- Modeler attempts to learn about model parameters and other players' decisions with observed data
- Outcomes can be qualitatively different from results obtained under assumption of full information or assumption of no modeling error
- Many unsolved problems remain