Dynamic Structural Models in Marketing and Economics

Andrew Ching

Johns Hopkins Carey Business School

Introduction

- A common framework to study consumers' or firms' forward-looking behavior is discrete choice dynamic programming models (DDP).
  - Ching and Osborne (2020): Consumer Stockpiling Problem.
  - Ching and Ishihara (2019): Demand for used and new video games.
- Consumer Learning Models
  - Ching et al. (2016): The Effects of Publicity on Demand (Prescription Drugs).
  - Ching and Lim (2020): Correlated Learning and Late-Mover Advantage.

Ching & Osborne (2020), Ching & Ishihara (2019)

- How forward-looking are economic agents?
- Discount factor is often set to the rational expectations benchmark – consistent with the prevailing interest rate.
- Without any model restrictions, DDP are not identified (Magnac and Thesmar 2002, Rust 1994).
- Our approach is to use exclusion restrictions, which is closely related to Fang and Wang (2015) and Abbring and Daljord (2020).

Ching and Osborne (2020): Consumer Stockpiling Problems

- Ching, Andrew T. and Matthew Osborne (2020) "Identification and Estimation of Forward-looking Behavior: The Case of Consumer Stockpiling," *Marketing Science*, vol.39(4), pp.707-726.
- Related works:
- Erdem, Imai and Keane (2003), Hendel and Nevo (2006)

Our Contributions (1)

- Propose a set of assumptions that can provide identification of the parameters of stockpiling models, with a focus on the discount factor.
  1. **Exclusion restrictions**: storage cost is function of the number of packages held rather than inventory within a package.
  2. **Consumption rates**: exogenous, and it takes several periods to use up a package (often many periods in practice).
- In standard formulations of stockpiling models, one assumes that storage costs are a **continuous** function of inventory.
- Stockpiling is widely studied in marketing, due to its importance for dynamic pricing.

How Inventory Can Create Exclusion Restrictions

- In many product categories, storage costs are a function of the number of packages held (ie, laundry detergent).
- Consider a consumer on her last bottle of detergent.
- As the amount in the bottle decreases, the possibility of a stockout happening in the near future increases.
- Since the cost of storing one bottle is constant, the consumer's storage cost doesn't change.
- A myopic consumer's purchase probability will be flat until she runs out.
- A forward-looking consumer's purchase probability should increase as inventory drops.
  - With continuous and increasing storage costs, purchase probability increases as inventory drops, **even for a myopic individual**.

Empirical Application

- Estimate dynamic structural model (including discount factors) on household-level scanner data.
- Population-average estimate: 0.71
- Significant heterogeneity across individuals:
  - 25th percentile: 0.67.
  - 75th percentile: 0.80.
- Find evidence that the discount factor is heterogeneous and varies with demographic characteristics:
  - Higher income households, older households, and larger households are more forward-looking.
  - However, the effect of demographics are relatively small: most heterogeneity is unobserved.

Implications for Pricing

Hi-Lo Pricing vs EDLP:

- In storable goods markets, the value to firms of using periodic promotions depends on how much consumers consider the future.
  - Hendel and Nevo (2013), Hong, McAfee and Nayyar (2002), Pesendorfer (2002), Sobel (1984)
- Fixing the discount factor to be too high may overstate the value of Hi-Lo pricing.

Temporary price discounts:

- Drive category expansion rather than brand switching (Erdem, Imai and Keane (2003)).
- Substantially increase consumption for several weeks after the promotion (Sun (2005)).
- Increasing promotion depth has higher returns than frequency (Osborne (2018)).

Talk Outline

- Develop stylized stockpiling model and derive properties.
- Discuss identification with observed inventory.
- Discuss identification with unobserved inventory.
- Artificial data experiments (time permitting).
- Application to scanner data.

Identification in a Simple Model

- Single product available in one package size.
- Package size is $b = 8$ units.
- Consumption need is fixed at $c = 1$ (we can relax this).
- Purchase price is $p > 0$ (fixed over time for now).
- Price coefficient: α.
- Stockout cost is ν.
- Discount factor is β.
- State is consumer inventory level I.
- Consumer decision is to buy or not buy a single package.
- Maximum storage capacity is $M = 3$ packages.
- Consumption utility γ.

Consumer Choice Specific Utility

- Utility of not buying:

$$
\begin{aligned}
v_{it,0} &= \gamma \mathbf{1}(I \geq c) - \nu \mathbf{1}(I < c) - SC(B) + \\
&\quad \beta V(\min\{I - c, 0\}) + \epsilon_{it,0}
\end{aligned}
$$

  Notation: B is number of packages held at end of the period.

- Utility of buying.

$$v_{it,1} = \gamma - \alpha p - SC(B) + \beta V(I + b - c) + \epsilon_{it,1}$$

- Storage Cost (nonparametric)

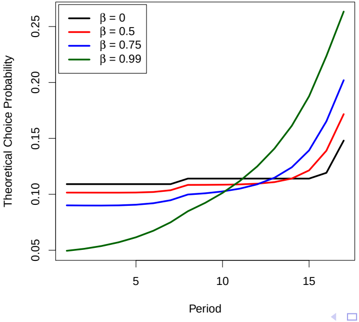
$$SC(B) = \omega_B$$

Properties of the Model

- If storage costs are weakly convex and weakly increasing, and $Var(\epsilon_{it,1} - \epsilon_{it,0})$ is sufficiently small, we can prove the following propositions:
  1. The expected future value of making a purchase, $\beta[V(I + b) - V(I)]$, is decreasing in I for $I \geq 0$. It is strictly decreasing if $\beta > 0$ and 0 if $\beta = 0$.
  2. The expected future value of purchase, $\beta\left[V(I + b) - V(I))\right]$, is increasing in β for $I \geq 0$ and sufficiently small storage costs.

Identification with Observed Inventory (Intuition)

- With logit errors, *all* model parameters can be expressed in terms of observed choice probabilities at different levels of inventory.
- Intuitive argument: set $I = 16$ in period 0, plot purchase probability (inventory decreases by 1 every period).



Identification with Observed Inventory

- Formal argument follows Ching and Ishihara (2018).
- We wish to identify $M + 3$ parameters.
  - $\boldsymbol{\theta} = (\omega_1, ..., \omega_M, \alpha, \nu, \beta)$
- Define the choice-specific value of buying j packages as:

$$\begin{aligned} v_j(I; \boldsymbol{\theta}) &= -\alpha p \mathbf{1}\{j = 1\} - \omega_{B(j,I,1)} - \nu \mathbf{1}\{I = 0\} \\ &\quad + \beta V(\max\{I + bj - 1, 0\}). \end{aligned}$$

- Under logit errors, we can express choice probabilities at I in terms of choice-specific values as

$$\Delta \log(\hat{P}(I; \boldsymbol{\theta})) = \log(\hat{P}(I)) - \log(1 - \hat{P}(I)) = v_1(I; \boldsymbol{\theta}) - v_0(I; \boldsymbol{\theta}).$$

Identification with Observed Inventory (cont'd)

- It is possible to derive formulas for all the model parameters in terms of $\Delta \log(\hat{P}(I;\boldsymbol{\theta}))$.
- Defining

$$\hat{\Phi}(I) \equiv \log\left(1 + \frac{\hat{P}_1(I+b)}{\hat{P}_0(I+b)}\right) - \log\left(1 + \frac{\hat{P}_1(I)}{\hat{P}_0(I)}\right),$$

- The formula for the discount factor is

$$\hat{\beta} = \frac{\Delta \log(\hat{P}(I+2)) - \Delta \log(\hat{P}(I+1))}{\Delta \log(\hat{P}(I+1)) - \Delta \log(\hat{P}(I)) + \hat{\Phi}(I+1) - \hat{\Phi}(I)}$$

Identification with Observed Inventory (cont'd)

- Price coefficient: $\hat{\alpha} = -\frac{\Delta \log(\hat{P}(Mb))}{p}$.
- Stockout cost and storage cost (first package):

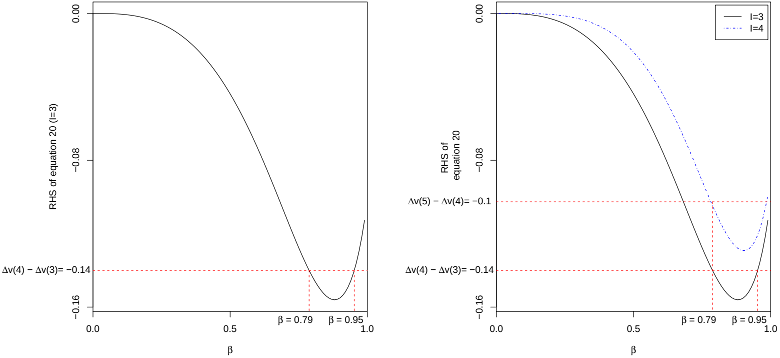
$$\begin{bmatrix} \frac{\hat{\beta}^{b-1}-1}{1-\hat{\beta}} & \frac{2\hat{\beta}-\hat{\beta}^{b}-1}{1-\hat{\beta}} \\ \frac{\hat{\beta}^{b}-1}{1-\hat{\beta}} & \frac{\hat{\beta}-\hat{\beta}^{b+1}}{1-\hat{\beta}} \end{bmatrix} \begin{bmatrix} \hat{\omega}_1 \\ \hat{\nu} \end{bmatrix} = \begin{bmatrix} \hat{\alpha}p - \Delta \log(\hat{P}(0)) + \hat{\beta} h_0(\hat{\beta}) \\ \hat{\alpha}p - \Delta \log(\hat{P}(1)) + \hat{\beta} h_1(\hat{\beta}) \end{bmatrix},$$

- Storage costs for 2+ packages:

$$\begin{aligned} \hat{\omega}_B \quad = \quad & -\hat{\alpha}p + \hat{\omega}_1 + \hat{\beta}(V((B-1)b+1) - V((B-2)b+1)) \\ & -\Delta \log(\hat{P}((B-1)b+2)) \end{aligned}$$

- The h functions and value function differences can be expressed solely in terms of choice probabilities.

Why we need multiple exclusion restrictions



The $\Delta v(I)$ is $v_1(I) - v_0(I)$

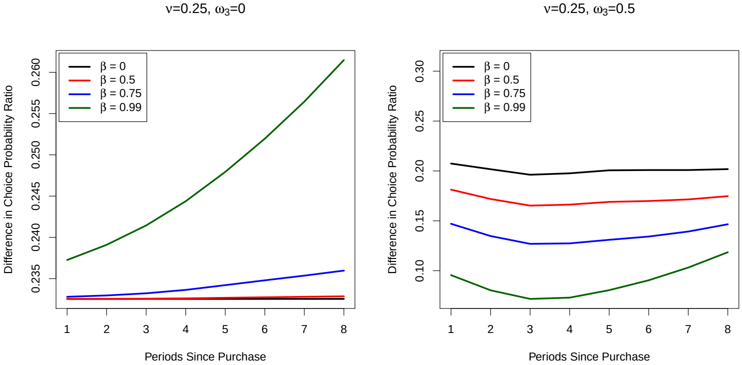
Price Variation and Multiple Packages

- Forward-looking consumers will be more likely to purchase multiple packages in response to price cuts as inventory drops.
- To show this we run a numerical experiment allowing for 2 prices, $p = 1$ and $p = 2$, and allow consumers to purchase up to 2 packages.
- We compute the following for τ periods following a purchase in period 0:

$$\frac{Prob_t(\text{buy 2 units}|p=1)}{Prob_t(\text{buy 1 units}|p=1)} - \frac{Prob_t(\text{buy 2 units}|p=2)}{Prob_t(\text{buy 1 units}|p=2)}$$

- Note: Storage costs and the discount factor can have **opposite** effects on the above moment.
  - Caveat: can break down if β is sufficiently low and ω sufficiently high.

Difference in Prob(2)/Prob(1) at High vs Low Price



Unobserved Heterogeneity

- So far we have assumed that all parameters are fixed across the population.
- If we allow unobserved heterogeneity, one approach is to rely on the ability to construct the purchase hazard and consumption rate moments on an individual basis.
- We want T to go to infinity at a rate that is much faster than N.
- With infinite amounts of data, we could consistently estimate the purchase hazard for each individual, and then construct individual level estimates of the parameters.
- With finite amounts of data, we have to rely on distribution assumptions wrt the unobserved heterogeneity.

Artificial Data Experiment

- To investigate identification in a more realistic setting we simulate data under different parameterizations and recover the parameters.
- The simulated data has 500 households, and 600 periods.
- We assume in period 1 everyone starts with 0 inventory.
- Since in real data initial inventory is unobserved (and probably positive), we conduct estimation on the final 400 periods.
- We use periods 201 to 400 to simulate initial inventories, and periods 401 to 601 to estimate parameters. (In estimation we start simulating initial inventory at 0)
- We use 100 simulated paths of consumption shocks for each household in constructing the likelihood.

Artificial Data Experiment: Price Process, Parameters

- We need price variation to estimate the price coefficient (α).
- We assume that there are 3 possible prices with the following transition process

	$p_t = 0.5$	$p_t = 1$	$p_t = 2$
$p_{t-1} = 0.5$	0.1	0	0.9
$p_{t-1} = 1$	0	0.1	0.9
$p_{t-1} = 2$	0.1	0.1	0.8

- Base parameters values are $\alpha = 1$, $\pi_c = 0.5$,$\nu = 0.1$, $\eta = 1$.

Artificial Data Experiments: Results

Parameter	No Storage Costs Est	S.E.	Truth	ω_3 Free Est	S.E.	Truth	ω_1, ω_2, ω_3 Free Est	S.E.	Truth
Price Coeff (α)	1.004	0.007	1	1.002	0.014	1	1.001	0.015	1
Stockout Cost (ν)	0.098	0.01	0.1	0.103	0.006	0.1	0.101	0.036	0.1
Discount Factor (β)	0.957	0.016	0.95	0.957	0.032	0.95	0.954	0.067	0.95
ω_1	-	-	0	-	-	0	0.105	0.059	0.1
ω_2	-	-	0	-	-	0	0.243	0.051	0.25
ω_3	-	-	0	0.499	0.055	0.5	0.508	0.142	0.5
π_c	0.489	0.007	0.5	0.49	4.72e-04	0.5	0.5	0.002	0.5

Parameter	$\beta = 0.001$ Est	S.E.	Truth	$\beta = 0.6$ Est	S.E.	Truth	$\beta = 0.99$ Est	S.E.	Truth
Price Coeff (α)	1.002	0.007	1	1.002	0.009	1	1	0.014	1
Stockout Cost (ν)	0.096	0.022	0.1	0.1	0.011	0.1	0.102	0.006	0.1
Discount Factor (β)	0.001	0.149	0.001	0.619	0.052	0.6	0.994	0.034	0.99
ω_1	-	-	0	-	-	0	-	-	0
ω_2	-	-	0	-	-	0	-	-	0
ω_3	0.479	0.061	0.5	0.47	0.044	0.5	0.488	0.059	0.5
π_c	0.492	0.001	0.5	0.496	2.80e-04	0.5	0.494	4.03e-04	0.5

Application to Scanner Data

- To test our method in real data we use IRI scanner data on laundry detergents to estimate actual discount factors.
- We use the years 2001 - 2007 for estimation.
- Complications of using field data:
  1. There are many brands, so some simplification of the state space is necessary.
  2. Even with 3 or 4 continuous state variables, the model becomes computationally burdensome if we use nested fixed point.
  3. We also need to model persistent unobserved consumer heterogeneity (in brand preferences, price sensitivities, etc)

Estimation Method

- Since we cannot estimate a model with brand specific inventory, where consumers track brand specific prices, we apply the simplifications of Hendel and Nevo (2006)
  - All utility from brand consumption occurs at the time of purchase.
  - Inclusive value sufficiency
- We use the method of Imai, Jain and Ching (2009) to estimate the model.
  - Alleviates computational burden of including persistent unobserved heterogeneity.

Highlights of IJC method

- In the conventional approach, the value functions need to be solved at every trial parameter vector (θ^{*r}).
- The value functions computed at past parameter vectors are simply thrown away!
- Imai, Jain and Ching (2009) (IJC) algorithm:
  - In each MCMC iteration, the value function is only partially solved (at the minimum, only apply the Bellman operator once). We call them **pseudo-value functions**.
  - Store those pseudo-value functions evaluated at past parameter vectors, and use them to approximate the value functions at the current parameter vector nonparametrically.
  - This nonparametric approximation can be computationally much cheaper than the method of successive approximation.

IJC Algorithm

- Outer loop (MCMC algorithm)
  - Similar to the conventional Bayesian approach.
  - Use the likelihood constructed based on **pseudo alternative specific value functions**, $\tilde{V}_j^r$. (thus, we also call the likelihood the pseudo-likelihood).
- Inner loop (Key innovation of the IJC algorithm)
  - Approximate the expected future value at θ^{*r} by the weighted average of the past pseudo-value functions.
  - Apply the Bellman operator once to get pseudo-value function evaluated at θ^{*r}, and store it.

IJC: Inner loop

- Let $H^r = \{\theta^{*l}; \tilde{V}^l(s, p^l; \theta^{*l}), \forall s\}_{l=r-N}^{r-1}$ be the outcome of the algorithm to iteration $r-1$.

- For each s, the expected future value at the current parameter value (θ^{*r}) is approximated as

$$\hat{E}^r_{p'}[V(s, p'; \theta^{*r})] = \sum_{l=r-N}^{r-1} \tilde{V}^l(s', p^l, \theta^{*l}) \frac{K_h(\theta^{*r} - \theta^{*l})}{\sum_{k=r-N}^{r-1} K_h(\theta^{*r} - \theta^{*l})},$$

where $K_h()$ is a kernel with bandwidth $h > 0$.

- Pseudo alternative specific value functions are then

$$\tilde{V}^r_j(s, p_j; \theta^{*r}) = \begin{cases} \alpha_j - \gamma p_j + \beta \hat{E}^r_{p'}[V(s, p'; \theta^{*r})] & \text{if } s_j < \bar{S}_j - 1, \\ \alpha_j - \gamma p_j + G_j + \beta \hat{E}^r_{p'}[V(s, p'; \theta^{*r})] & \text{if } s_j = \bar{S}_j - 1, \end{cases}$$

$$\tilde{V}^r_0(s, p_j; \theta^{*r}) = \beta \hat{E}^r_{p'}[V(s, p'; \theta^{*r})].$$

IJC: Inner loop (cont'd)

- Simulate one draw of price vector, p^r, from the known price distribution. Apply the Bellman operator once and obtain the pseudo-value function:

$$\tilde{V}^r(s, p^r; \theta^{*r}) = E_\epsilon \max_j \{\bar{U}_{ijt}(s, p^r; \theta^{*r}) + \epsilon_{ijt} + \hat{E}_{p'}[V^l(s, p'; \theta^{*r})]\}.$$

We store $\{\theta^{*r}; \tilde{V}^r(s, p^r; \theta^{*r}), \forall s\}$ and update the outcome to H^{r+1}.

Modeling Simplifications

- Consumption rates were hard to identify (even with a fixed discount factor) so we calibrate from the data.
  - We perform a robustness exercise where we increase consumption rate by 25% and re-estimate the model.
  - Average discount factor estimated to be 0.77 (vs 0.73 in the baseline).
- With 5 package sizes, allowing a flexible storage cost function significantly increases computational burden.
- We would have to track the composition of inventory (in terms of bottle sizes), as well as make an assumption on the order in which different sizes of bottles are used.
- We assume storage costs are zero until a point ω_i (which we estimate), when they become infinite. (We are relaxing this)

Model Specifications

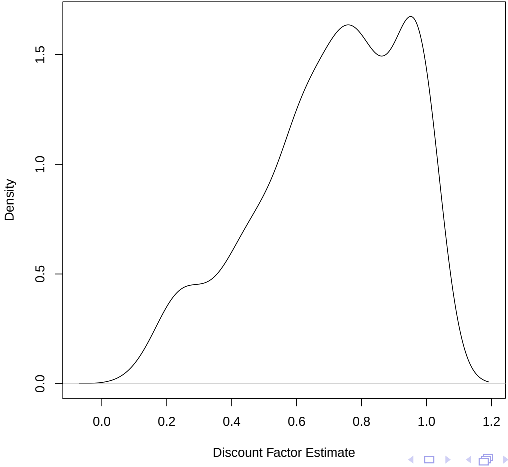
- All specifications include the following demographic variables at the household level: income, age, education and size.
- Two primary specifications: forward-looking and myopic.
  - DIC of forward-looking model: 40511 (marginal log-likelihood: -19585.4)
  - DIC of fixing $\beta = 0.9995$: 40773 (marginal log-likelihood: -19768.0)

Estimates (Forward-Looking Model)

Parameter	1st Quartile	Median	Mean	3rd Quartile
Price Coefficient	-0.29	-0.24	-0.27	-0.21
	[-0.31, -0.28]	[-0.26, -0.23]	[-0.29, -0.26]	[-0.22, -0.2]
Stockout Cost	0.29	0.39	0.48	0.5
	[0.24, 0.36]	[0.31, 0.49]	[0.37, 0.66]	[0.4, 0.67]
Discount Factor	0.62	0.94	0.71	0.99
	[0.14, 0.89]	[0.81, 0.98]	[0.58, 0.82]	[0.98, 1]
Fixed Cost of Purchase	-	-	-1.83	-
			[-1.91, -1.77]	
Log-likelihood	-19585.37			

Notes: This table shows average moments of the posterior distribution of the population distribution of the dynamic parameters. For example, the median columns shows the average of the population median of a given parameter, where the average is taken across MCMC draws. Square brackets show 95% confidence intervals.

Kernel Density of Individual-Specific Discount Factor Estimates



Marginal Effects of Demographic Variables

Parameter	Baseline	HH Income	HH Head Age	HH Head College	HH Size
Price Coefficient	-0.25	0.042	-0.045	0	-0.006
		[0.006, 0.077]	[-0.077, -0.01]	[-0.036, 0.036]	[-0.048, 0.032]
Stockout Cost	0.469	-0.036	-0.176	0.258	-0.181
		[-0.246, 0.158]	[-0.445, 0.139]	[-0.024, 0.557]	[-0.312, -0.001]
Discount Factor	0.563	-0.435	0.497	-0.037	0.243
		[-0.993, 0]	[0.001, 0.966]	[-0.838, 0.754]	[-0.204, 0.889]

Notes: This table shows the estimated impact of changing one of the demographic dummy variables from zero to one on a particular parameter. The respective demographic dummy variables are defined to be 1 under the following conditions: Income above $35,000; age of household head above 55; household head has a college degree; size of household is more than 2 individuals. The baseline column shows the predicted value of a parameter at the mode of the demographic distribution.

Counterfactual

- Simulate choices under the following scenarios: i) estimated discount factor, ii) discount factor set to 0.9995, items i) and ii) with 50% higher promotional frequency, and 50% deeper promotions for 100 oz Tide.
- We do this under the scenario that consumer expectations update, and assuming they do not (not yet in the paper).

Increased Promotional Depth				
	Estimated Discount Factor		$\beta = 0.9995$	
Counterfactual	Quantity	Revenue	Quantity	Revenue
Updated Expectations	502.72	1018.48	576.95	1183.98
	[448, 560]	[777.49, 1256.67]	[514, 641]	[916.38, 1452.15]
Expectations from Data	507.56	1034.54	599.05	1255.78
	[452, 566]	[794.98, 1269.93]	[532, 669]	[972.66, 1542.56]
Increased Promotional Frequency				
	Estimated Discount Factor		$\beta = 0.9995$	
Counterfactual	Quantity	Revenue	Quantity	Revenue
Updated Expectations	49.34	232.99	55.19	276.34
	[-11, 110]	[-214.35, 679.05]	[-7, 118]	[-171.78, 734.91]
Expectations from Data	49.62	234.77	55.83	279.28
	[-11, 111]	[-213.24, 681.19]	[-7, 119]	[-172.65, 735.99]

Conclusion

- We develop an identification argument for the discount factor in stockpiling models, and estimate it in scanner data based on two assumptions:
  1. Storage cost only varies with number of packages held.
  2. It takes several periods for consumers to use up a package
- Empirical results suggest individuals are more myopic than the rational expectations benchmark.
  - We are working on new counterfactuals to quantify the impact on pricing.
- Q: Do low discount factors reflect some unmodeled behavioral phenomena?
  - inattention, cost of learning the optimal policy, etc
- We are conducting follow up work to test these ideas in field data.

Ishihara and Ching (2019): Demand for New and Used Digital Goods

- Ishihara, Masakazu and Andrew T. Ching (2019) “Dynamic Demand for New and Used Durable Goods without Physical Depreciation: The Case of Japanese Video Games,” *Marketing Science*, vol.38(3), pp.392-416.
- Related works:
- Gillingham, Iskhakov, Munk-Nielsen, Rust & Schjerning (2022), Berkovec (1985)

Motivation

Video game publishers' attempt to fight against used video games

In Japan

- Used video game lawsuit in 1998-2002
- Publishers tried to kill off the used game market (but lost in 2002).

In the U.S.

- Recent expansion of the used game market, e.g., 7-Eleven
- Publishers put restrictions on online features of games (require activation code)

Motivation (contd)

Why are producers so concerned about the used game markets?

Producers claim that specific features of the goods make the competition fiercer (as opposed to cars, etc.):

1. Products physically depreciate very negligibly
   - New and used copies could be almost identical
2. Owners might quickly become satiated and sell
   - Used goods might become available soon after a new good release

Used Goods Markets

Used goods markets might influence new good sales in two ways:

1. **Substitution effect**
   - Used goods markets provide consumers with substitutes for new goods
   - Affects new good sales **negatively**
2. **Resale effect**
   - Used goods markets provide consumers with future selling opportunity
   - Affects new good sales **positively**

Two Main Challenges

1. Are consumers forward-looking?

   - Whether consumers are forward-looking is an important factor that determines the resale effect

     Example: if consumers are myopic, no resale effect

   - We use an exclusion restriction to estimate the discount factor (Chevalier and Goolsbee 2009)

2. How to conceptualize "durability"

   - Previous literature just focuses on physical depreciation (e.g., Esteban and Shum 2007; Chen et al. 2010; Engers, Hartmann and Stern 2009; Schiraldi 2009; Tanaka 2009)

Concept of "Durability"

For products such as video games, CDs/DVDs, there could be different sources of depreciation of consumption values for owners and potential buyers:

1. **Satiation-based** depreciation (for owners)

   - Owners eventually become bored
   - Speed of satiation may depend on product characteristics:

     Example: story-based games (e.g., mystery games) versus non-story based games (e.g., sports games)

2. **Freshness-based** depreciation (for potential buyers)

   - Hype and buzz

Research Questions

1. Does the existence of used goods markets help or hurt new-good producers?
2. Are consumers forward-looking?
3. How does the speed of satiation depend on product characteristics? How quickly does consumption value decrease due to freshness-based depreciation?

To answer these questions,

- We develop a model of forward-looking consumers' buying <u>and</u> selling decisions and apply it to the Japanese video game market (software).
- We extend the Bayesian Markov chain Monte Carlo algorithm by Imai, Jain, and Ching (2009) to a non-stationary model.

Background: Japanese Video Games

Since 80s, the video game industry has rapidly become one of the most important sectors in the entertainment industry

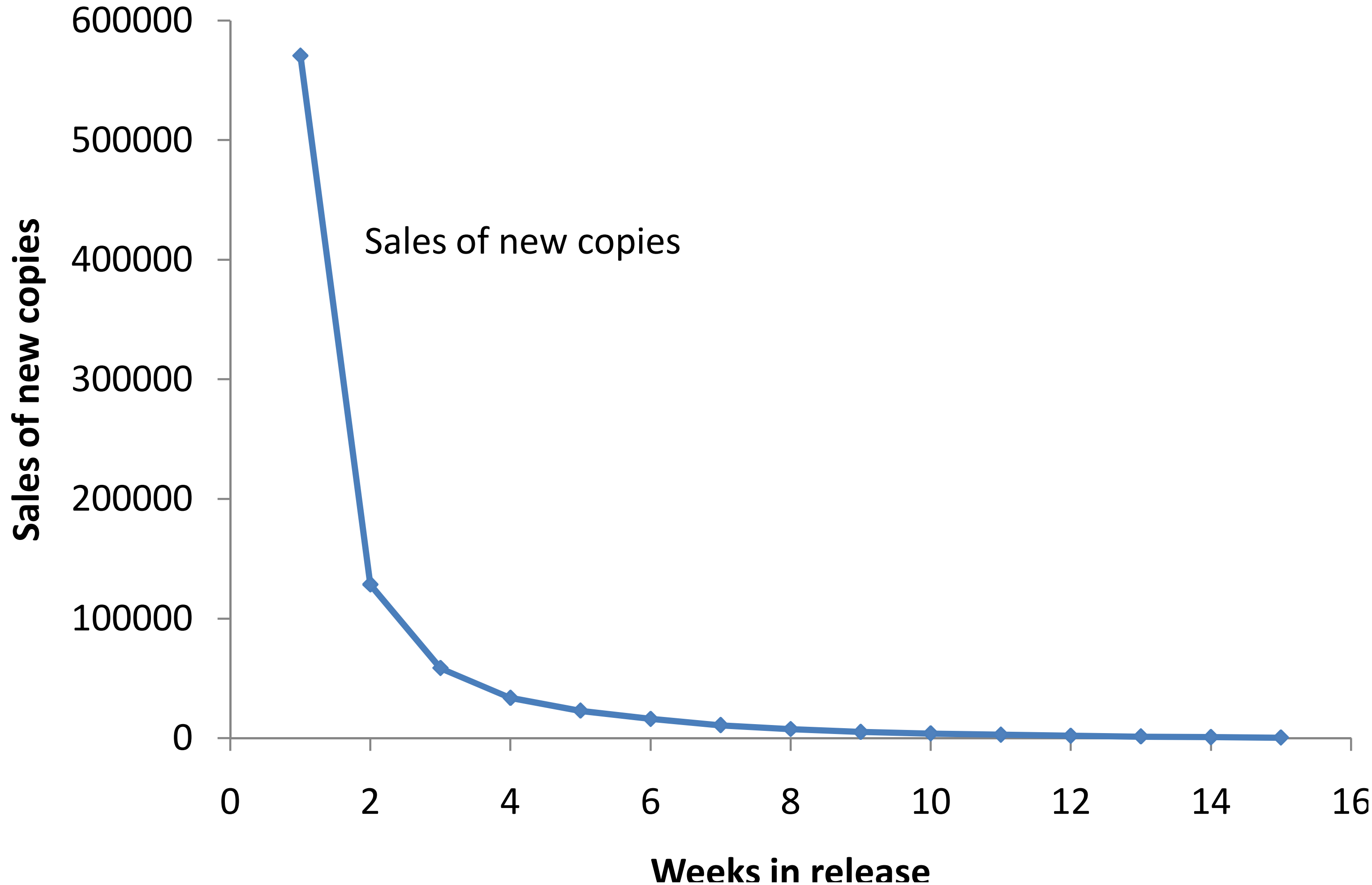
- The size of new-game markets (hardware,software,equipments): \$5.5 billion in 2009
- The size of used-game markets (only software): \$1 billion in 2009
- Most of used game trading is controlled by video game retailers (about 85%)
- No rental market

Data

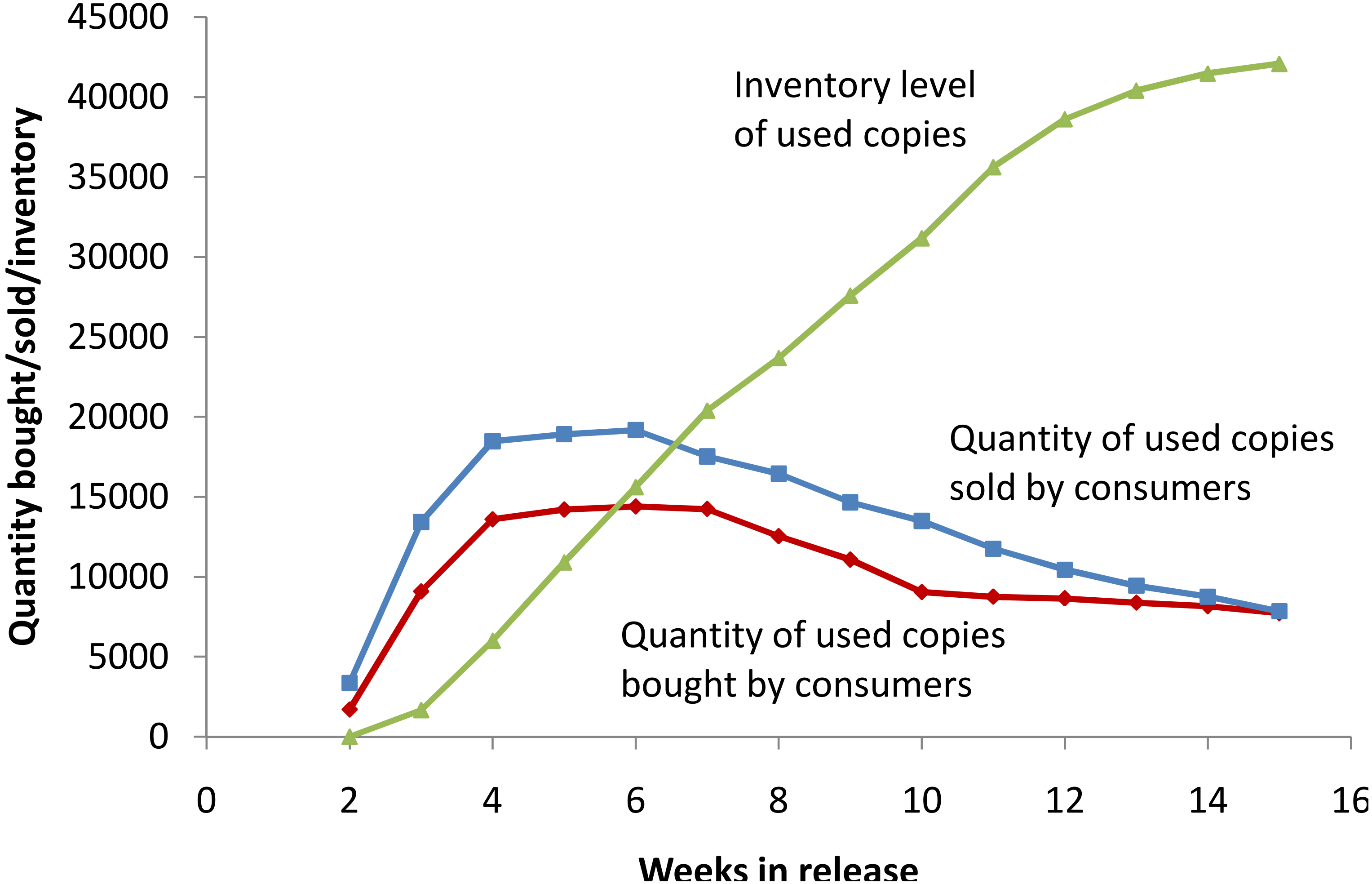
20 home video games (software) released in Japan between 2004 and 2008 (data sources: Weekly Famitsu Magazines, Annual Video Game Industry Reports)

- Sales: weekly aggregate sales of new and used copies, and quantity of used copies sold by consumers
- Prices: price of a new copy, weekly price and resale value of a used copy
- Product characteristics: average critic and user rating, dummies for story-based games and multiplayer games

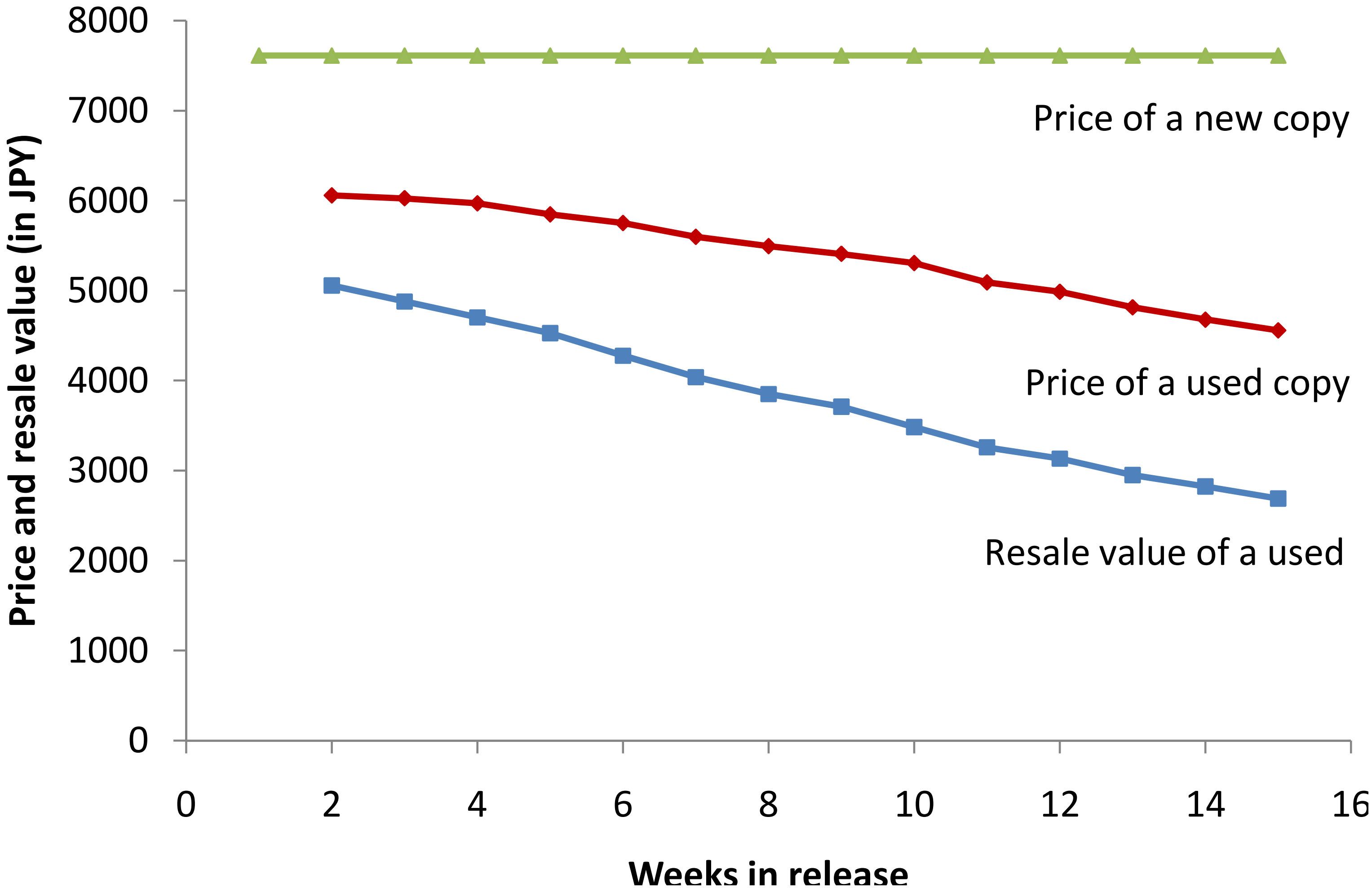
New Copies: Average Sales



Used Copies: Average Quantities



Average Prices and Resale Value



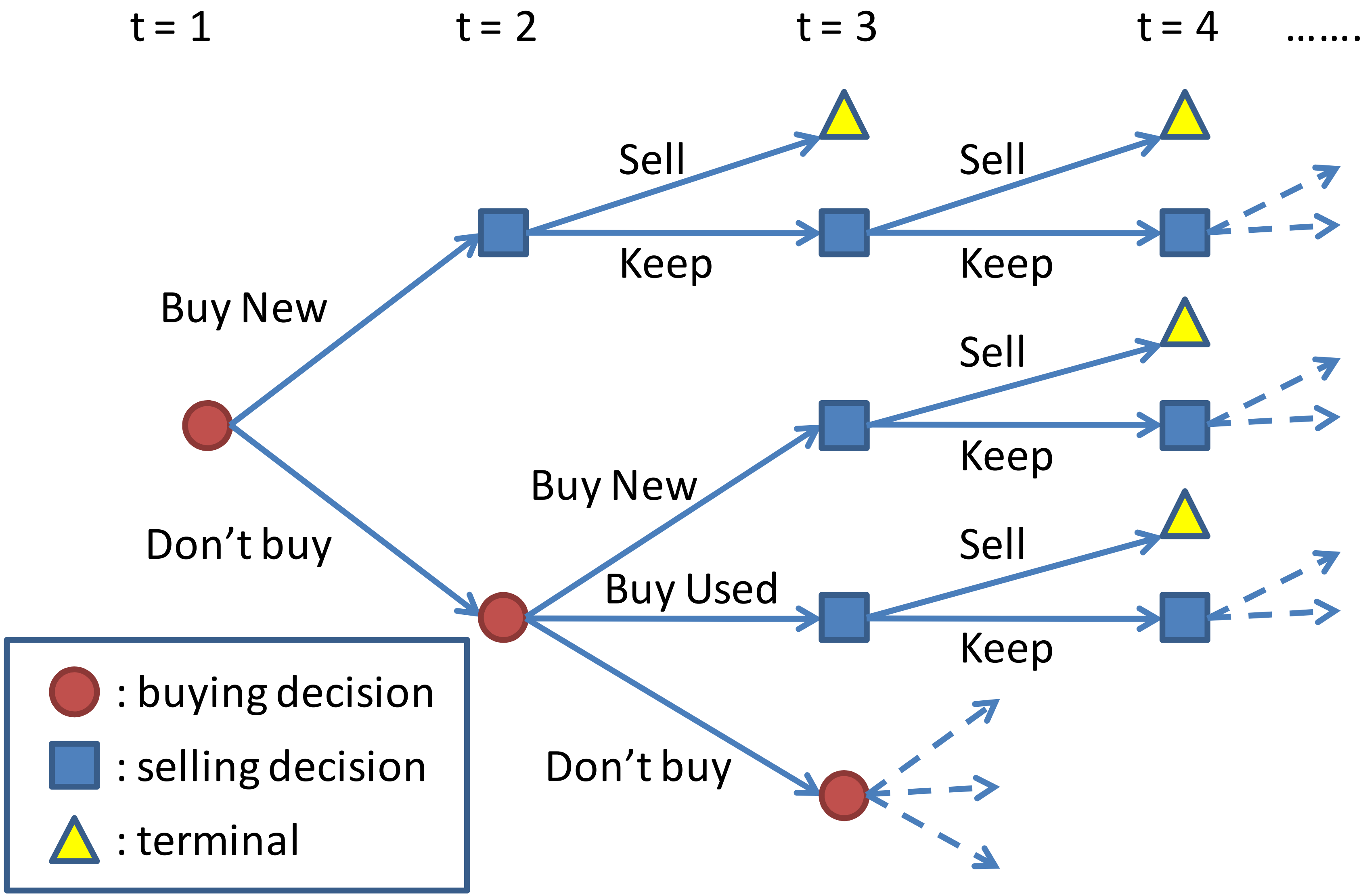
Summary Statistics

	Average	S.D.	Min	Max
Price of new copies (in JPY)	7,613.1	629.1	7,140	9,240
Price of used copies (in JPY)	4,515.3	1,087.8	2,219	7,433
Resale value of used copies (in JPY)	2,828.1	1,182.7	1,036	6,547
Sales of new copies	100,650.4	259,022.3	2,772	2,236,881
Sales of used copies	7,184.6	6,478.8	458	62,734
Quantity sold by consumers	8,121.4	8,436.8	1,012	55,830
Inventory of used copies	31,022.5	28,347.7	0	129,462
Market size (installed base)	14,866,067.6	6,097,167.2	746,971	20,822,775
Weekly # new game introduction	7.01	4.02	0	17
Dummy for story-based games	0.700	0.470	0	1
Dummy for multi-player games	0.450	0.510	0	1
Critic rating (in 10-point scale)	8.99	0.656	7.75	10
User rating	56.4	9.20	41.6	67.4

Note: USD 1 ≈ JPY 100

* user rating is a standardized score against a set of video games released in the same year (by Enterbrain, Inc.)

Decision Tree



Model

Four important features:

1. Consumers first make buying decisions, then make selling decisions after purchase
2. Consumers form expectation about future prices of new and used goods and resale value of used goods
3. Satiation-based depreciation and freshness-based depreciation
4. Impact of used-good availability on consumer buying decisions

Utility Function for Buying Decisions

The per-period utility function for consumer i who haven't bought the product up to time t:

$$u_{ijt}^{g} = \begin{cases} \pi_0 + \pi_1 \ln(C_t^g) + \epsilon_{i0t}^g & \text{no purchase } (j = 0) \\ v^g(t) - \alpha p_{1t}^g + \xi_{1t}^g + \epsilon_{i1t}^g & \text{new copy } (j = 1) \\ v^g(t) - \alpha p_{2t}^g - l(Y_t^g) + \xi_{2t}^g + \epsilon_{i2t}^g & \text{used copy } (j = 2) \end{cases}$$

- $v^g(t)$: consumption values to potential buyers at time t
- $v^g(1) = \gamma^g$ and $v^g(t+1) = (1 - \phi_t)v^g(t)$
- $l(Y_t^g) = \lambda_0 + \lambda_1 \exp(-\lambda_2 Y_t^g)$: any additional cost for buying a used good (e.g., psychological costs, search costs); Y_t^g is the inventory level of used copies
- C_t^g: cumulative number of competing games since game g's release

Utility Function for Selling Decisions

The per-period utility function for type-i owners who have owned for τ periods up to time t:

$$w^g_{ikt}(\tau) = \begin{cases} \tilde{v}^g(t,\tau) + e^g_{i0t} & \text{keeping } (k=0) \\ \alpha r^g_t - \mu + \xi^g_{st} + e^g_{i1t} & \text{selling } (k=1) \end{cases}$$

- $\tilde{v}^g(t,\tau)$: consumption values to owners who have owned for τ periods at time t
- $\tilde{v}^g(t,0) = v^g(t)$ and $\tilde{v}^g(t+1,\tau+1) = (1-\kappa^g_\tau)\tilde{v}^g(t,\tau)$
- r^g_t is the resale value at time t
- μ is any additional cost for selling the product (e.g., transaction costs, endowment effects)

Dynamics of Consumer Selling Decisions

Drop g superscript. State vector: $s_w = (r, Y, \xi_s, t, \tau)$.

The Bellman equation for selling decisions by consumer i who has kept the product for τ periods is given by

$$W_i(s_w) = \max\{W_{i0}(s_w) + e_{i0t}, W_{i1}(s_w) + e_{i1t}\}$$

Alternative-specific value functions

$$W_{ik}(s_w) = \begin{cases} \tilde{v}(t,\tau) + \beta E[W_i(s'_w)|s_w] & \text{keep } (k=0) \\ \alpha r_t - \mu + \xi_{st} & \text{sell } (k=1) \end{cases}$$

Dynamics of Consumer Buying Decisions

Let $p = (p_1, p_2)$. State vector: $s_v = (p, r, Y, C, \xi_1, \xi_2, t)$

The Bellman equation for buying decisions by type-i consumer is given by

$$V_i(s_v) = \max\{V_0(s_v) + \epsilon_{i0t}, V_1(s_v) + \epsilon_{i1t}, V_2(s_v) + \epsilon_{i2t}\}$$

Alternative-specific value functions

$$V_j(s_v) = \begin{cases} \pi_0 + \pi_1 \ln(C_t) + \beta E[V_i(s'_v)|s_v] & \text{don't buy } (j=0) \\ v(t) - \alpha p_{1t} + \xi_{1t} + \beta E[W_i(s'_w)|s_v] & \text{new copy } (j=1) \\ v(t) - \alpha p_{2t} - l(Y_t) + \xi_{2t} + \beta E[W_i(s'_w)|s_v] & \text{used copy } (j=2) \end{cases}$$

Estimation Strategy

Two-stage approach:

Stage 1: Estimate the processes for consumer expectations about future prices, resale value, inventory level, and cumulative number of newly introduced games, by linear regressions with lagged values and product characteristics

Stage 2: Assuming consumers' expectation follows the process estimated in stage 1, estimate the rest of the structural parameters

- Control for the potential price endogeneity by the pseudo-policy function approach (Ching 2010):
  - Approximate the pricing policy functions by a polynomial of observed and unobserved state variables
  - Jointly estimate it with the demand-side model

IJC algorithm

IJC algorithm (Imai et al. 2009): A Bayesian MCMC algorithm for single-agent dynamic discrete choice models with infinite time horizon.

- In each MCMC iteration, solve for value functions partially (pseudo-value functions) instead of fully solving for them. Store pseudo-value functions
- Use stored pseudo-value functions to nonparametrically approximate the expected future values at current parameter values
- As the number of MCMC iterations increases, pseudo-value functions will converge to the true ones

However, this algorithm cannot directly be applied to a finite-horizon problem!

Modified IJC algorithm

We extend the IJC algorithm to a non-stationary model with stochastic continuous state variables:

- In each MCMC iteration, pseudo-value functions are solved for each time period, but only at **one randomly chosen state vector** per time period
- Pseudo-value functions are stored period by period
- In approximating the expected future values at time t, use the pseudo-value functions evaluated at time $t+1$

This new algorithm has the potential to lower the computational cost, especially when the model has multiple stochastic continuous state variables.

Identification

Satiation-based depreciation (for owners):

- Time-series variations in the quantities sold by consumers across games, controlling for the variation in resale value and the size of owners (both observed)

Freshness-based depreciation (for potential buyers):

- The average declining rate of the sales of new (and used) games over time, after controlling for new (and used) prices

Discount factor:

- Exclusion restriction: resale values affect the continuation value for buying decisions, but do not affect the per-period utility for buying decisions.

Estimates: Preference

	full model		model w/o ppf	
	mean	s.d.	mean	s.d.
preference parameters				
discount factor (β)	0.878	0.001	0.897	0.002
price sensitivity (α)	5.99E-04	2.52E-05	5.94E-04	1.21144E-05
costs for buying used goods				
intercept (λ_0)	0.060	0.012	0.319	0.015
inventory (λ_1)	1.36	0.000	1.36	2.40E-04
inventory (λ_2)	2.94E-04	7.72E-05	3.10E-04	8.21E-05
costs for selling used goods (μ)	5.97	0.075	5.97	0.043
seasonal dummies (γ)				
golden week (early May)	0.101	0.014	0.030	0.032
christmas (late Dec.)	0.297	0.016	0.208	0.016
no-purchase option				
intercept (π_0)	-0.461	0.083	-0.574	0.080
competitive effects from other games (π_1)	0.333	0.019	0.309	0.022

Estimates: Depreciation Rates

	full model		model w/o ppf	
	mean	s.d.	mean	s.d.
depreciation rates				
potential buyers (φ)				
intercept	-0.382	0.112	-0.582	0.122
time since release	-6.20	0.14	-7.80	0.21
game owners (δ)				
intercept	0.547	0.028	0.706	0.039
story-based game	-0.389	0.018	-0.232	0.016
multi-player game	0.611	0.010	0.678	0.034
critic rating	-0.026	0.011	-0.036	0.027
user rating	0.016	0.002	0.012	0.005
ownership duration (logged)	-0.484	0.015	-0.551	0.019

Estimated Consumer Expectation Process $(t = 2)$

	price of used copies		resale value	
variable	estimate	s.e.	estimate	s.e.
price of new copies	0.784*	0.083	0.839	0.129
dummy for story-based games	208.9	114.9	100.5	177.8
dummy for multi-player games	150.1	120.3	140.4	186.2
critic rating	127.0	85.5	132.4	132.3
user rating	-12.8	6.31	-13.4	9.76
constant	-548.2	732.2	-1899.6	1133.2
Adjusted R-squared	0.921		0.840	
# observations	20		20	

Note: * $p < 0.05$

Estimated Consumer Expectation Process $(t > 2)$

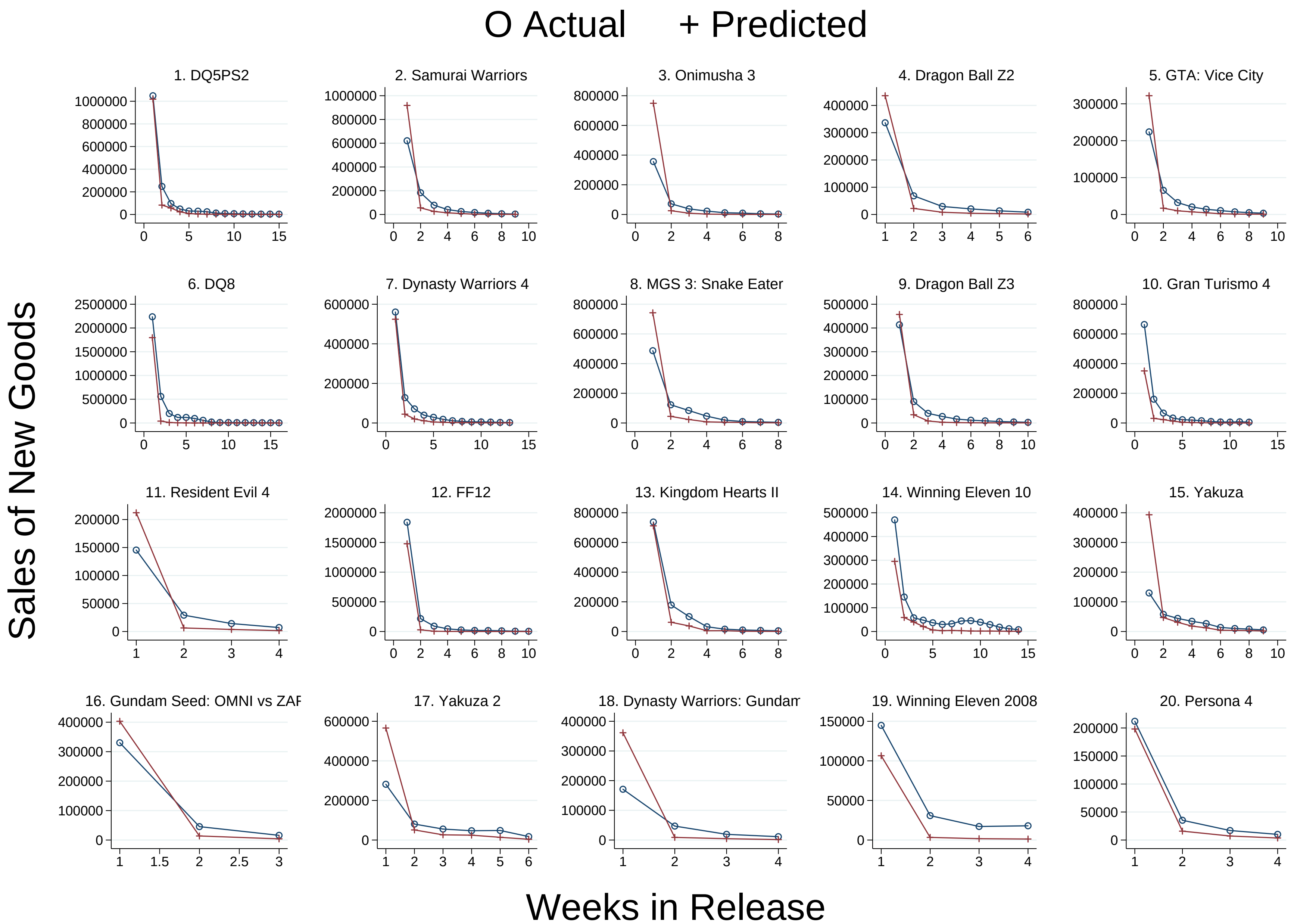
variable	price of used copies estimate	price of used copies s.e.	resale value estimate	resale value s.e.	inventory estimate	inventory s.e.
lagged value	0.958*	0.005	0.928*	0.005	0.958*	0.006
lagged inventory	-2.22E-03*	2.39E-04	-1.71E-03*	2.45E-04	-	-
dummy for story-based games	5.44	16.5	-26.6	16.8	1581.9*	472.5
dummy for multi-player games	-18.1	17.4	-14.3	17.7	-470.7	497.4
critic rating	29.0*	10.8	28.0*	11.0	1214.4*	306.8
user rating	-2.26*	0.731	-0.876	0.744	-93.8*	20.85
constant	54.6	97.8	-14.6	97.3	-4376.2	2733.4
Adjusted R-squared	0.987		0.988		0.984	
# observations	647		647		667	

Note: * $p < 0.05$

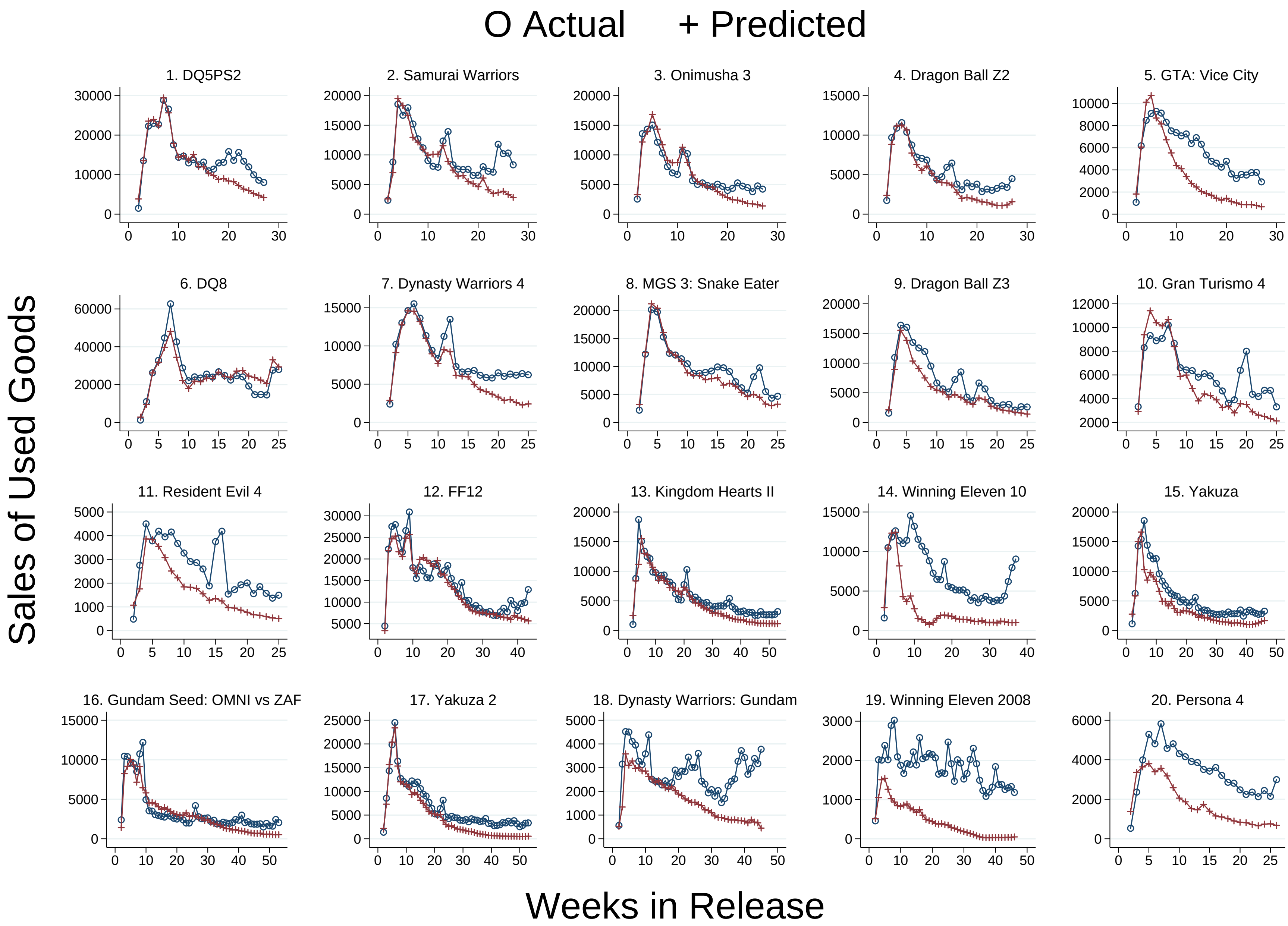
Estimation Results

- Freshness-based weekly depreciation rate: from the release week to 2nd week
  - about 41%
- Satiation-based depreciation
  - Ranges from 64% to 88%.
- Discount factor (β)
  - Estimated weekly discount factor is 0.88

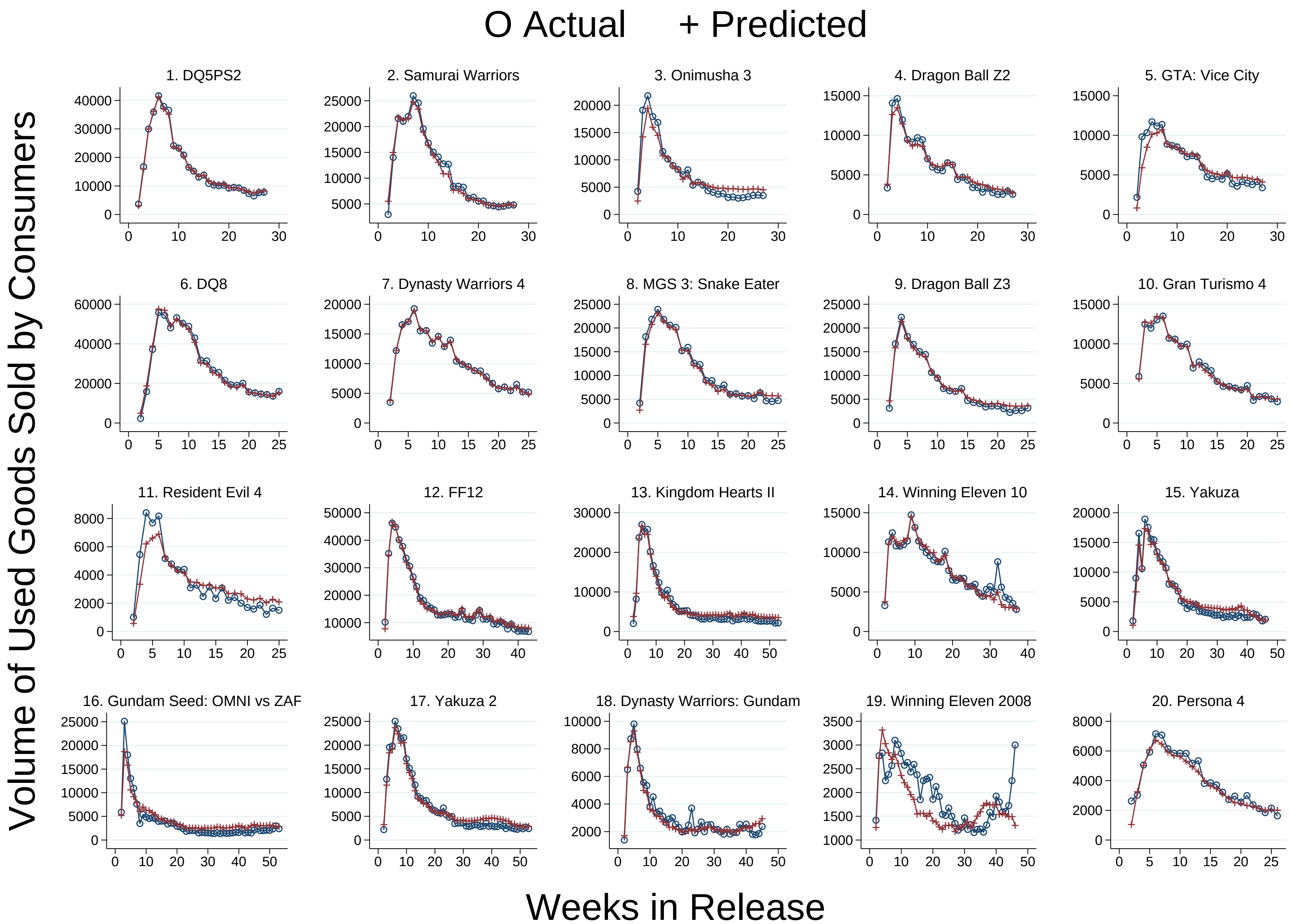
Model Fit: New Goods Sales



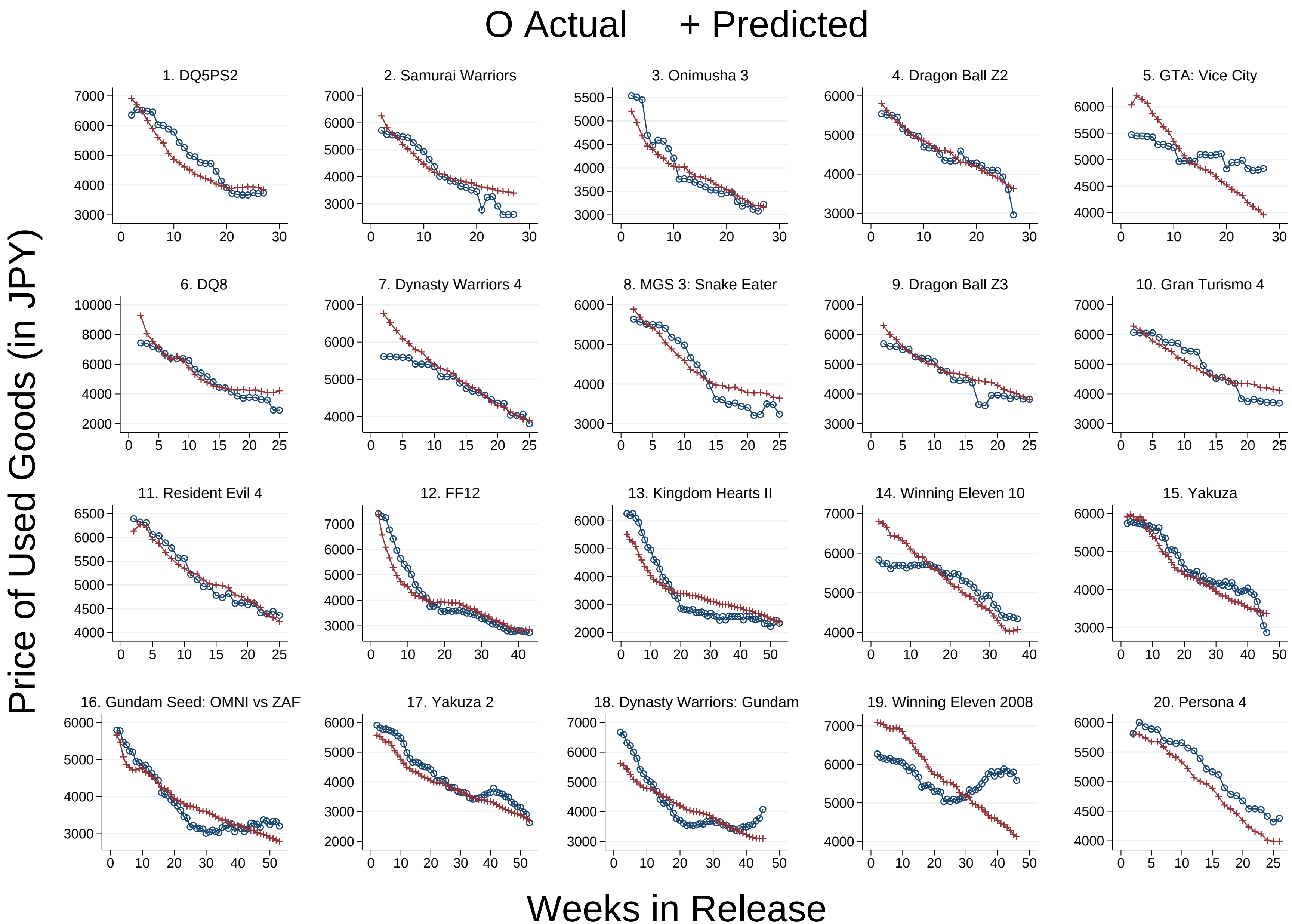
Model Fit: Used Goods Sales



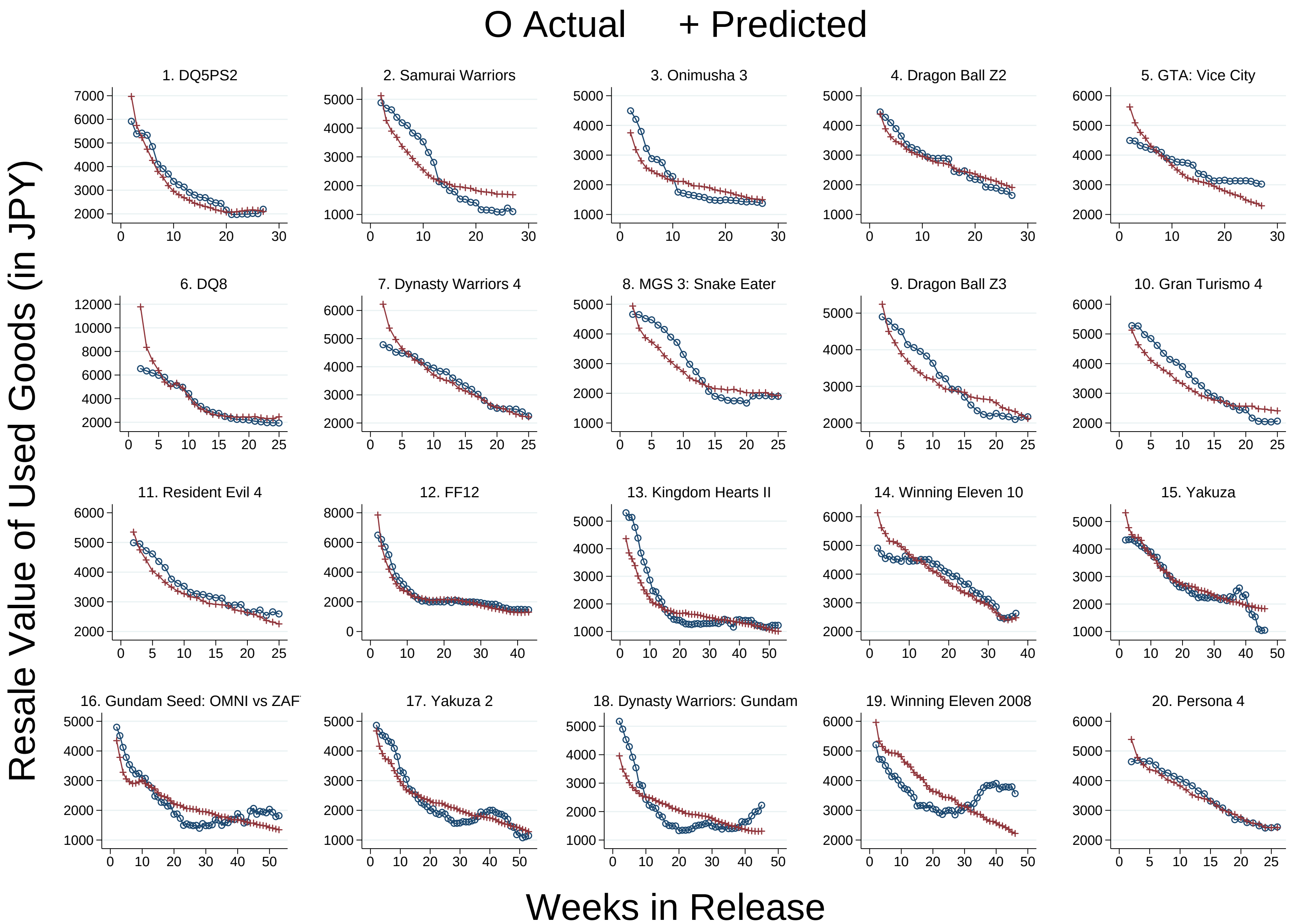
Model Fit: Used Goods Sold by Consumers



Model Fit: Price of Used Goods



Model Fit: Resale Value of Used Goods



Elimination of used game markets

Examine the overall impact of used-game markets on new-game revenues by eliminating used-game markets.

- Set both costs for buying and selling used games to be high so that the used-game market is effectively shut down.
- Simulate the model and examine the percentage change in new-game revenues.

Under the observed pricing strategy (flat pricing)

- On average, the elimination of the used-game market reduced revenues for a new game by 4% ($\approx$\$2.6 million).
- Sales of new copies dropped most in weeks right after release.

Flat Pricing vs Marginal Price-Skimming

In the presence of the used game market, is the observed flat pricing optimal?

Examine the profitability of an alternative pricing strategy, which reduces the price of new games marginally over time (by 0.1% per period).

Findings:

- On average, the marginal price-skimming strategy would increase the revenue for a new game by 0.5% ($\approx$\$0.3 million).
- But the revenues would decrease for some games.

Conclusion

- Develop a model of forward-looking consumers' buying and selling decisions, and apply it to the Japanese video game market.
- Extended the Bayesian MCMC algorithm by Imai, Jain, and Ching (2009) to a non-stationary model.
- Main findings:
    - Elimination of used game markets could reduce revenues for a new copy if publishers do not adjust their pricing strategy.
    - The observed flat pricing might not be optimal for some games.

Pseudo-Pricing Policy Function

The state space of the equilibrium model consists of

$$S = (\xi_1, \xi_2, \xi_s, Y, C, v(t,0), M_t^d, \{v(t,\tau), M_t^s(\tau)\}_{\tau=1}^{t-1}).$$

I used the following functional form:

$$\begin{aligned}
\ln p_{2t} \quad &= \quad \omega_{10} + \omega_{11} v(t,0) + \omega_{12} \frac{1}{(t-1)} \sum_{\tau=1}^{t-1} v(t,\tau) + \omega_{13} M_t^d \\
& \quad + \omega_{14} \frac{1}{(t-1)} \sum_{\tau=1}^{t-1} M_t^s(\tau) + \omega_{15}(\xi_{2t}^d - \xi_{1t}^d) + \omega_{16}\xi_t^s + \omega_{17} Y_t + \omega_{18} C_t + \nu_t^p. \\
\ln r_t \quad &= \quad \omega_{20} + \omega_{21} v(t,0) + \omega_{22} \frac{1}{(t-1)} \sum_{\tau=1}^{t-1} v(t,\tau) + \omega_{23} M_t^d \\
& \quad + \omega_{24} \frac{1}{(t-1)} \sum_{\tau=1}^{t-1} M_t^s(\tau) + \omega_{25}(\xi_{2t}^d - \xi_{1t}^d) + \omega_{26}\xi_t^s + \omega_{27} Y_t + \omega_{28} C_t + \nu_t^r.
\end{aligned}$$

Estimates: Pseudo-Policy Functions

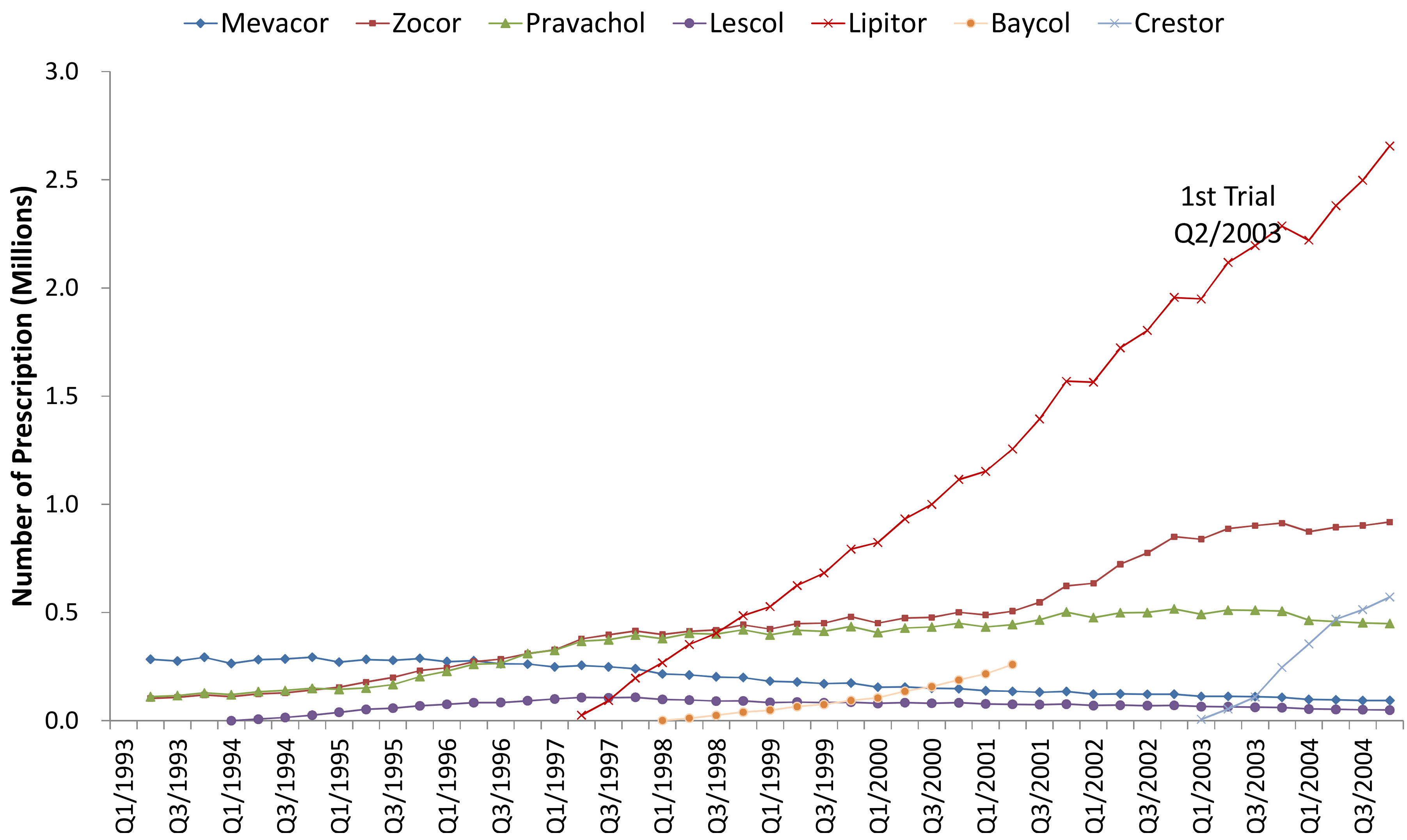
	mean	s.d.
pseudo-pricing policy function parameters		
price of used goods		
intercept (ω_{10})	7.78	0.048
consumption value for potential buyers (ω_{11})	0.221	0.007
average consumption value for owners (ω_{12})	-0.055	0.025
unobserved shock to buying (ω_{13})	0.002	0.005
unobserved shock to selling (ω_{14})	0.005	2.02E-05
size of potential buyers (ω_{15})	1.23E-08	4.26E-10
average size of owners (ω_{16})	1.67E-07	4.66E-08
inventory of used goods (ω_{17})	-2.76E-06	1.18E-07
cumulative # competing games (ω_{18})	-2.05E-03	5.07E-05
s.d.(v^p)	0.100	0.003
resale value of used goods		
intercept (ω_{20})	6.56	0.084
consumption value for potential buyers (ω_{21})	0.393	0.013
average consumption value for owners (ω_{22})	0.027	0.041
unobserved shock to buying (ω_{23})	0.007	0.008
unobserved shock to selling (ω_{24})	1.62E-04	3.63E-05
size of potential buyers (ω_{25})	3.21E-08	9.92E-10
average size of owners (ω_{26})	2.88E-07	8.10E-08
inventory of used goods (ω_{27})	-5.25E-06	8.73E-20
cumulative # competing games (ω_{28})	-0.003	9.18E-05
s.d.(v^r)	0.174	0.004

Ching and Lim (2020)
Correlated Learning & Late-mover Advantage

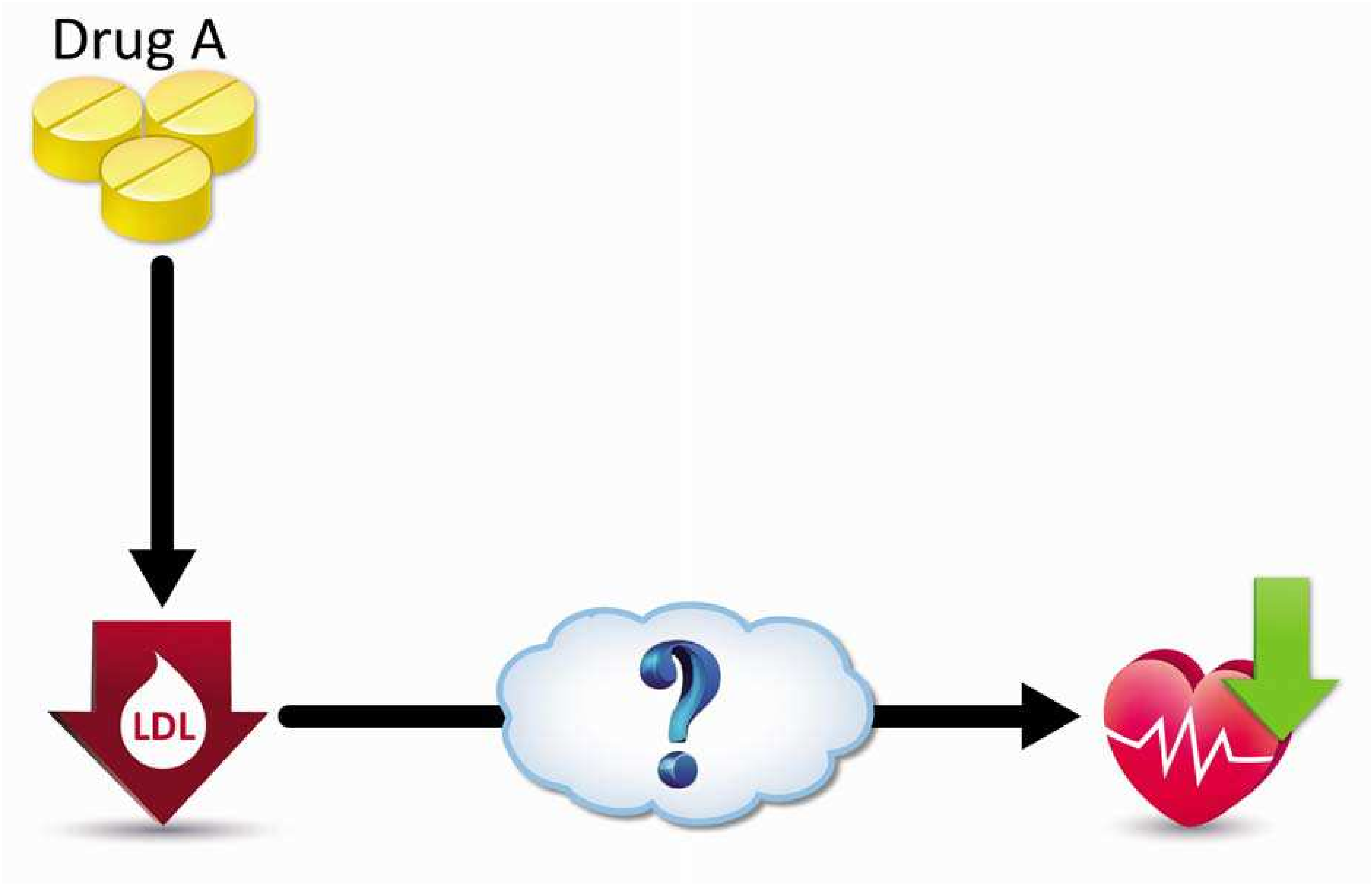
- Ching, A.T., R. Clark, I. Horstmann and H. Lim (2016), "The Effects of Publicity on Demand: The Case of Anti-cholesterol Drugs," *Marketing Science*, vol.35(1), pp.158-181.
- Ching, Andrew T. and Hyunwoo Lim (2020), "A Structural Model of Correlated Learning and Late-mover Advantages: The Case of Statins," *Management Science*, vol.66(3), pp.1095-1123.
- Related works:
- Ching (2010a, 2010b), Ching and Ishihara (2010, 2012)
- Ching, Liu and Hermosilla (2019) "Structural Models in the Prescription Drug Market," (with Manuel Hermosilla and Qiang Liu), *Foundations and Trends® in Marketing*, vol.13(1), pp.1-76. https://ssrn.com/abstract=3348196

  provides a survey of the structural models in the prescription drugs market.
- Ching, Erdem and Keane (2013, 2017) provide surveys of structural learning models in economics and marketing.
  - Ching, Erdem and Keane (2017) "Empirical Models of Learning Dynamics: A Survey of Recent Developments," Chapter 8 in *Handbook of Marketing Decision Models*
  - Ching, Erdem and Keane (2013) "Learning Models: An Assessment of Progress, Challenges and New Developments," (with Tülin Erdem, Michael Keane), *Marketing Science*, vol.32(6), pp.913-938.

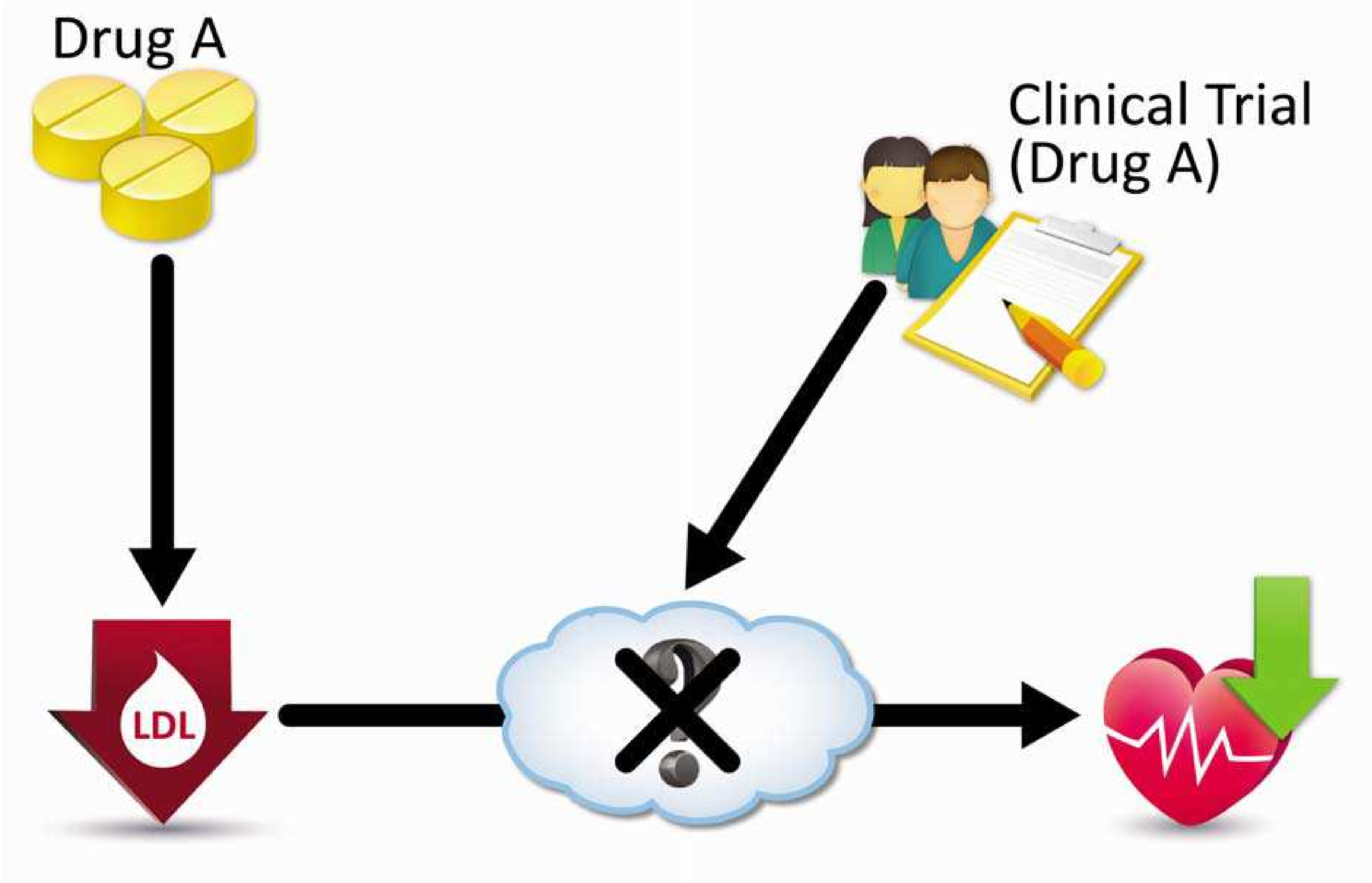
Sales of Statins from 1993 to 2004

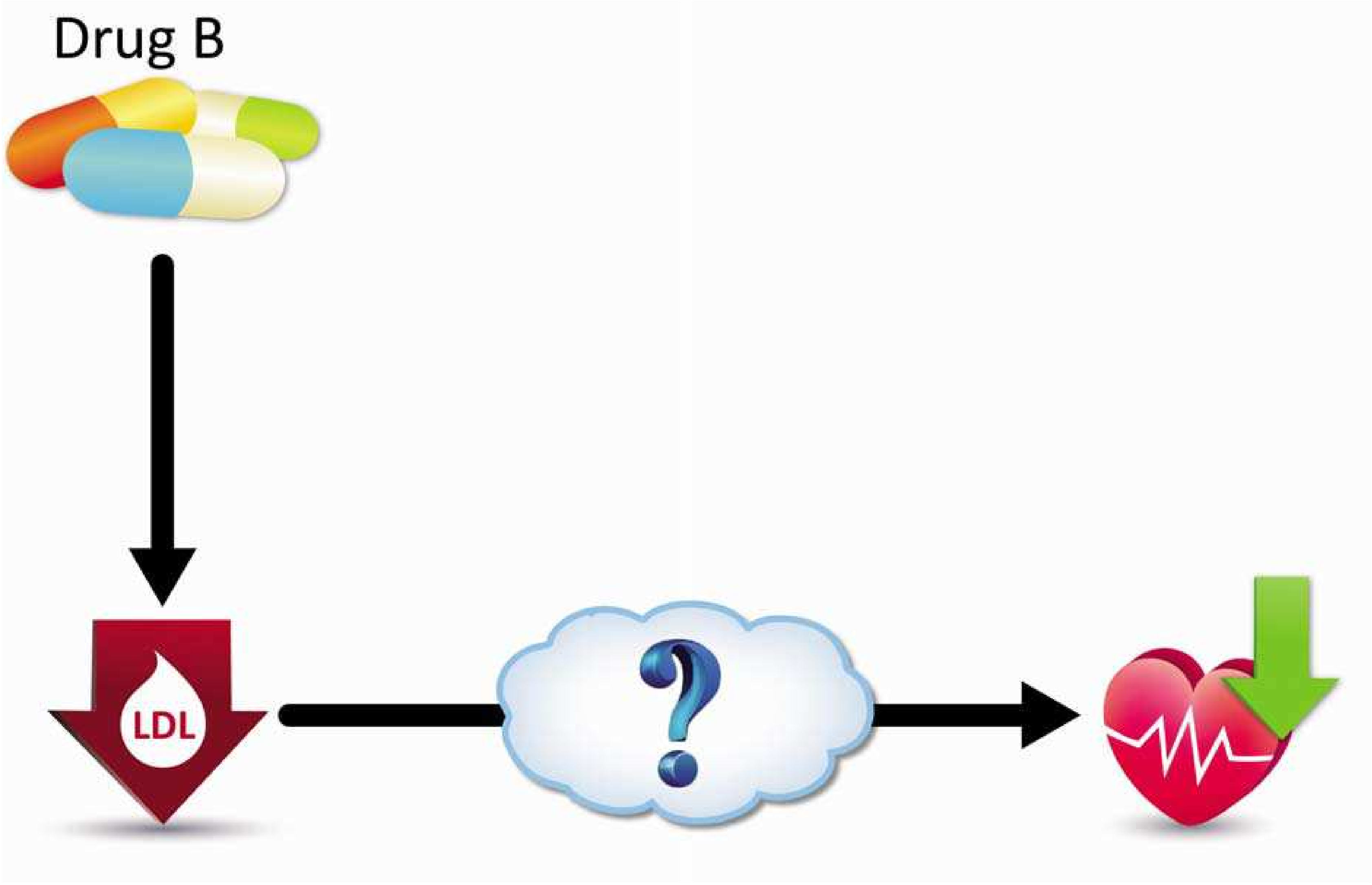


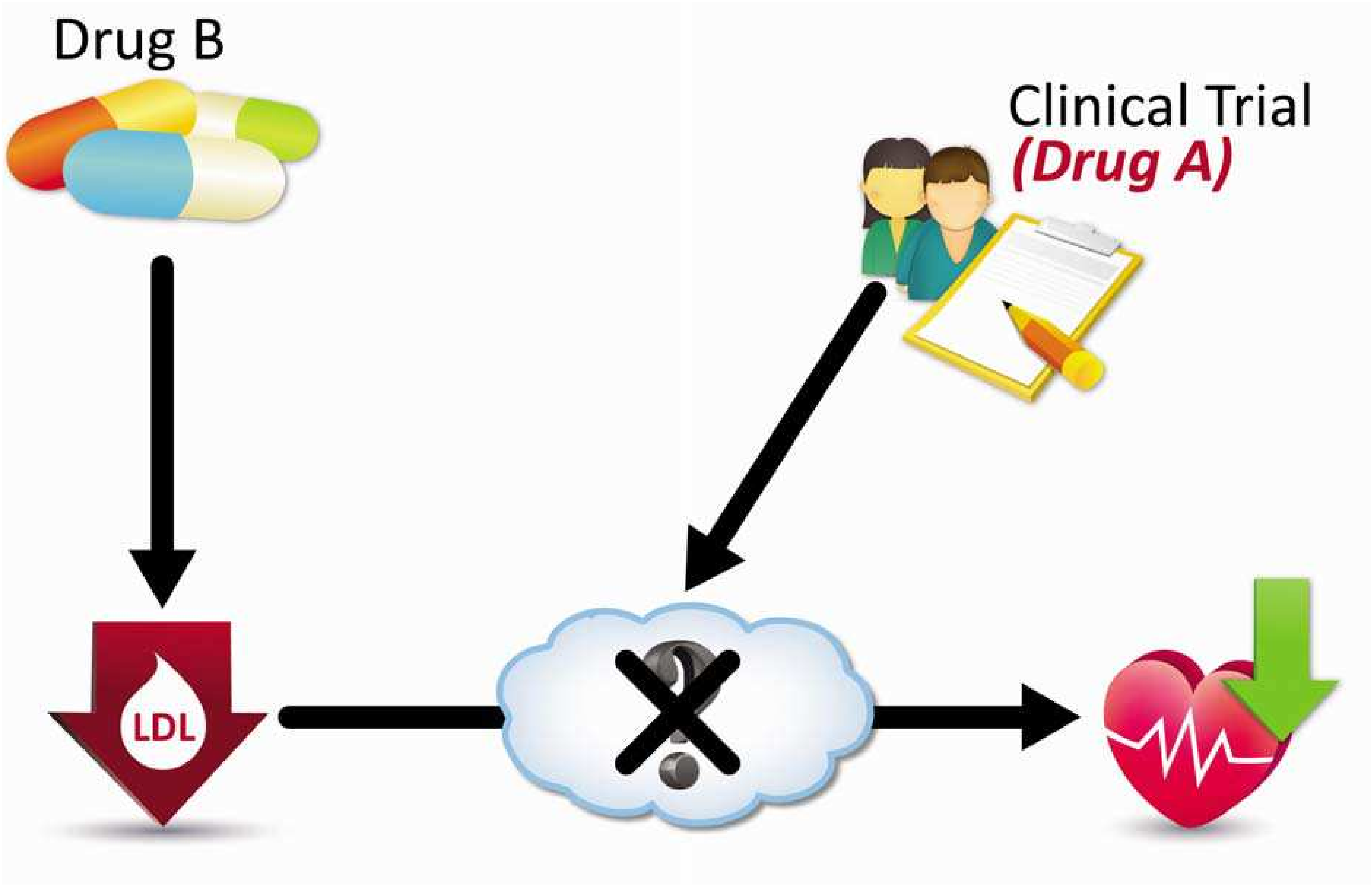
The Role of Clinical Evidence



The Role of Clinical Evidence




Drug B
LDL
?


Drug B
Clinical Trial
(Drug A)
LDL

Research Objectives

- Propose a structural model of correlated learning with indirect inference to explain late-mover advantage
- Our model focuses on a class of products with the following two features
  - products which build on a common fundamental technology
  - consumers can observe some (not all) product attributes of a product

Literature Review

- Janakiraman et al. (2009) study correlated learning across brand within a category.
  - This study assumes firms know the true quality of their products. (similar to Erdem & Sun (2002))
    - The effectiveness of advertising or detailing does not depend on consumption experience or clinical trial results.
    - Firms do not need to use consumption experience or clinical trials to learn about the true quality if the assumption is valid.
  - These implications are rejected by Azoulay (2002) and Venkataraman & Stremersch (2007).
- Ching & Ishihara (2010) incorporate clinical trials when modeling informative advertising, but do not consider correlated learning and late mover advantages.

Demand, Detailing, Switching and Discontinuing rates

- Quarterly Canadian data for each statin between Q2 1993 and Q4 2004 from IMS Canada
  - Prescription volume, Detailing spending
- Quarterly data on *switching* and *discontinuing* between Q2 1993 and Q4 2004 from Ontario Health Insurance Program (OHIP)
  - On average, 2.1% of statin users who switch from a given statin to another statin per quarter → Switching costs is high.
  - On average, 15% of statin users quit taking statin per quarter → Refilling costs (or the hassel costs) is also high.

Landmark Clinical Trials

- It is very difficult for physicians to learn about drugs' efficacy in heart disease risks from patient's feedback.
- Collect 14 landmark clinical trials reporting the efficacy of statins in reducing heart disease risks between 1993 and 2004.
- The number of patients consists of 1,600 to 20,000 and the follow-up period ranges from 2 to 6 years.
- They provide observable signals (to researchers) on how efficient a statin is in reducing heart disease risks.
  - More advanced than Ching & Ishihara (2010) who only use qualitative outcome of comparison trials.

Landmark Clinical Trials

Title	Publication Date	Drugs Studied	# of Subjects	Follow-up Period	Efficiency Raito
4S	Dec, 1994	Zocor	4,444	5.2 years	0.21
WOSCOPS	Nov, 1995	Pravachol	6,595	4.8 years	0.27
CARE	Oct, 1996	Pravachol	4,159	4.8 years	0.22
Post-CABG	Jan, 1997	Mevacor	1,351	4.2 years	0.22
AFCAPS/TexCAPS	May, 1998	Mevacor	6,605	5.3 years	0.30
LIPID	Nov, 1998	Pravachol	9,014	5.6 years	0.20
GISSI Prevention	Dec, 2000	Pravachol	4,271	1.9 years	0.23
LIPS	Jun, 2002	Lescol	1,677	3.1 years	0.24
HPS	Jul, 2002	Zocor	20,536	5 years	0.21
PROSPER	Nov, 2002	Pravachol	5,804	3.2 years	0.13
ALLHAT-LLT	Dec, 2002	Pravachol	10,355	4.8 years	0.12
ASCOT-LLA	May, 2003	Lipitor	10,305	3.2 years	0.28
ALERT	Jun, 2003	Lescol	2,102	5.1 years	0.11
CARDS	Aug, 2004	Lipitor	2,838	3.9 years	0.32

Publicity Data

- 2,754 articles mentioning “statin” from “Canadian Accessible Sources” in Factiva between year 1986 and 2004
- Classify articles along three dimensions:

  1. Lowering cholesterol levels
  2. Reducing heart disease risks
  3. Side effects

- Overcome the ambiguity of single dimensional coding scheme.
- Details are provided in Ching, Clark, Horstmann and Lim (2016).

Canadian Accessible Sources

- Online sources
- Canadian TV news, newspapers and magazines
- US TV news from big 4 TV stations (ABC, CBS, FOX and NBC) and CNN
- 8 US newspapers with more than 500,000 daily circulations
- 25 top selling US magazines

Classification

- For articles specifically mention drug names, classify each article into comparison vs. non-comparison for each dimension.
  - For non-comparison articles, we assign each drug "+1" or "-1" if the article shows positive or negative attitude towards the focal drug.
  - For comparison articles, we assign "+1" to the drug which the article favors most for each dimension. All other compared drugs are assigned "-1."

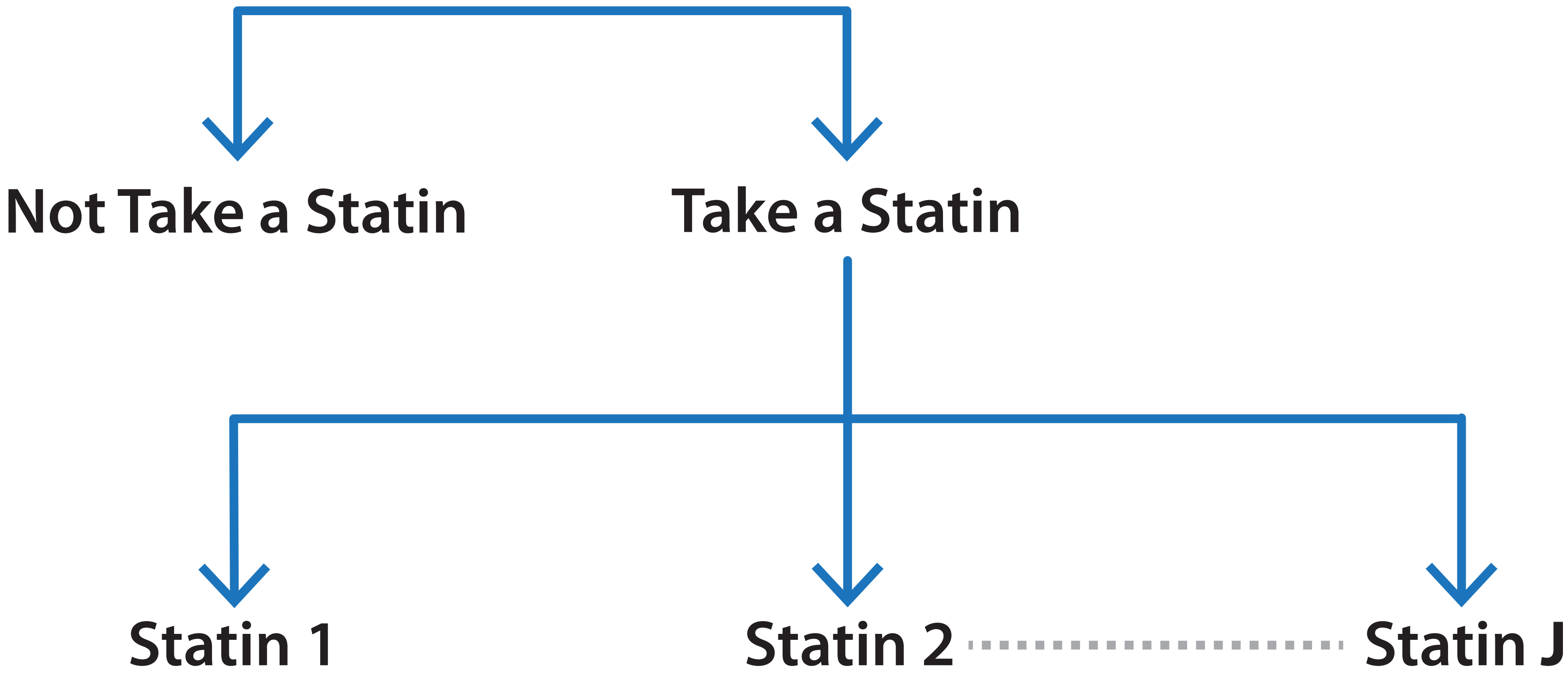
Summary of Publicity Variables

Lowering Cholesterol Levels	# of Months	Non-comparison					Comparison				
		# of Articles	Values				# of Articles	Values			
			Mean	Std. Dev.	Min.	Max.		Mean	Std. Dev.	Min.	Max.
Mevacor	198	255	1.80	1.93	0	9	6	-0.04	0.20	-1	0
Zocor	173	470	3.31	3.55	0	17	14	-0.04	0.29	-1	1
Pravachol	171	262	1.85	2.00	0	11	35	-0.22	1.45	-13	1
Lescol	130	33	0.24	0.62	-1	3	2	0.00	0.12	-1	1
Lipitor	94	707	7.53	7.00	0	28	58	0.15	1.57	-4	11
Baycol	41	9	0.22	0.65	0	3	0	0.00	0.00	0	0
Crestor	23	120	5.22	5.95	0	20	16	0.70	1.11	0	4

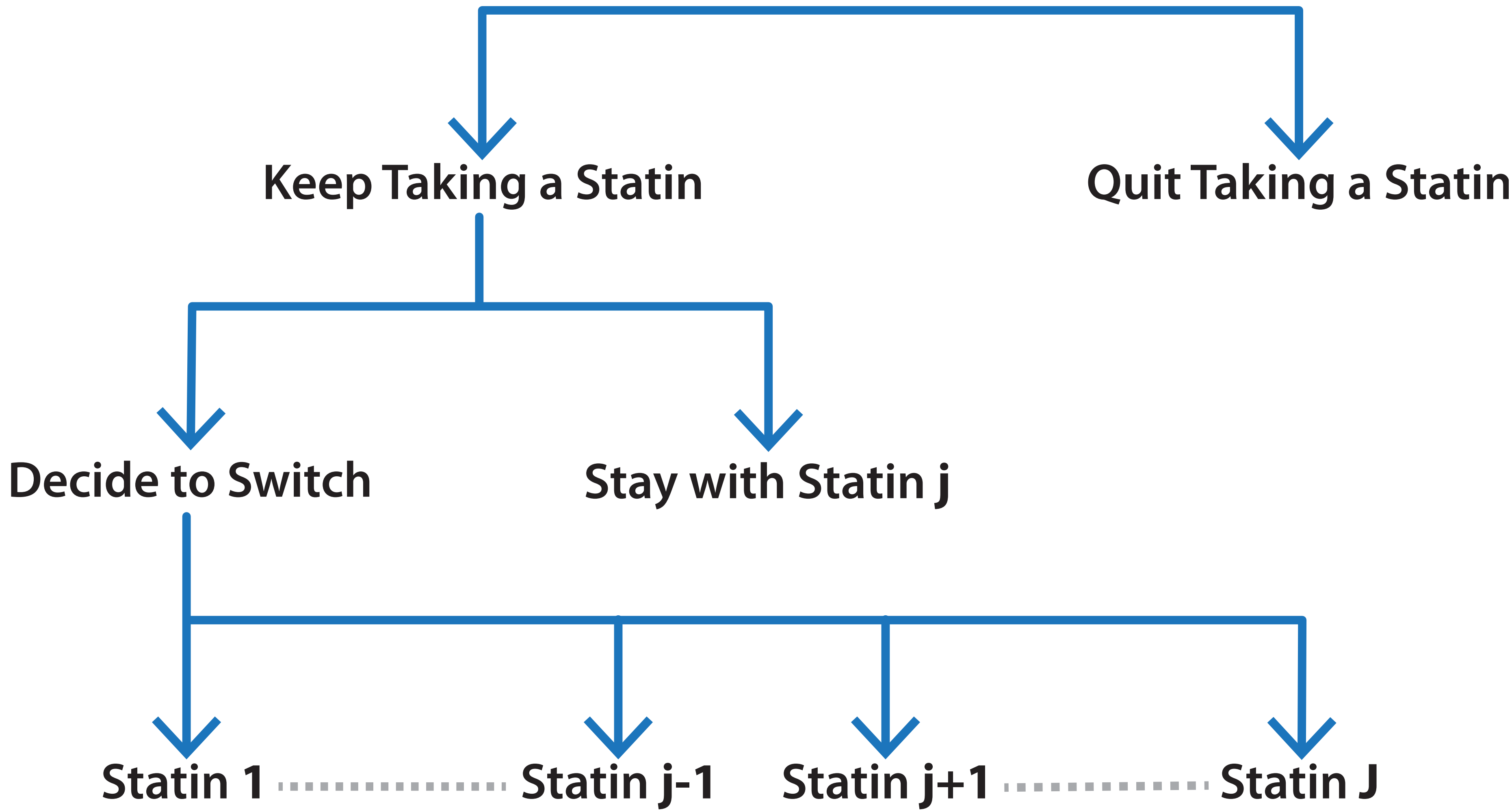
Reducing Risks of Heart Disease	# of Months	Non-comparison					Comparison				
		# of Articles	Values				# of Articles	Values			
			Mean	Std. Dev.	Min.	Max.		Mean	Std. Dev.	Min.	Max.
Mevacor	198	41	0.25	0.73	-1	4	0	0	0	0	0
Zocor	173	94	0.56	1.39	-3	9	2	0.01	0.12	0	1
Pravachol	171	81	0.54	1.11	-1	7	32	-0.23	1.49	-15	0
Lescol	130	3	0.02	0.15	0	1	0	0	0	0	0
Lipitor	94	92	0.98	2.18	0	15	32	0.32	1.83	-1	15
Baycol	41	1	0.02	0.16	0	1	0	0	0	0	0
Crestor	23	7	0.30	0.63	0	2	0	0	0	0	0

Side-Effects	# of Months	Non-comparison					Comparison				
		# of Articles	Values				# of Articles	Values			
			Mean	Std. Dev.	Min.	Max.		Mean	Std. Dev.	Min.	Max.
Mevacor	198	5	0.04	0.22	0	2	5	0.01	0.22	-1	2
Zocor	173	16	0.08	0.48	-1	4	4	0.01	0.21	-1	2
Pravachol	171	15	0.08	0.38	-1	3	10	0.08	0.39	0	2
Lescol	130	0	0.00	0.00	0	0	5	0.04	0.23	0	2
Lipitor	94	20	0.07	0.42	-1	1	9	0.03	0.40	-2	2
Baycol	41	2	0.05	0.31	0	2	0	0	0	0	0
Crestor	23	72	-0.07	0.49	-5	0	1	-0.04	0.21	-1	0

Joint Decision Process for Potential Patients



Joint Decision Process for Existing Patients



Efficacies of Statins

- q_j^c denotes the true efficacy in lowering cholesterol levels of drug j.
  - We assume that q_j^c is known to physicians.
  - A meta-analysis provides us with such information.
- q_j^h denotes the true efficacy in reducing heart disease risks of drug j.
  - We assume that q_j^h is uncertain to physicians.
  - Physicians learn about this efficacy from landmark clinical trials.

Mean Cholesterol Reduction (mmol/L)

	Daily Dose (mg)					Mean
	5	10	20	40	80	
Mevacor	N/A	1.02	1.40	1.77	2.15	1.59
Zocor	1.08	1.31	1.54	1.78	2.01	1.66
Pravachol	0.73	0.95	1.17	1.38	1.60	1.28
Lescol	0.46	0.74	1.02	1.30	1.58	1.16
Lipitor	1.51	1.79	2.07	2.36	2.64	2.22
Crestor	1.84	2.08	2.32	2.56	2.80	2.44

- Law et al. (2003) summarize (non-landmark) clinical trials investigating the efficacy in lowering cholesterol levels of statins.

Let q_j^h be the true efficacy in reducing heart disease risks of drug j

$$q_j^h = q_j^c \cdot \beta_j,$$

- q_j^c is the efficacy in lowering the cholesterol level;

- β_j is the "efficiency ratio".

Physicians know q_j^c, but are uncertain about β_j (and hence uncertain about q_j^h).

$$E[q_j^h|I(t)] = q_j^c \cdot E[\beta_j|I(t)].$$

Initial Prior Beliefs

Initial prior beliefs on "efficiency ratio" are constructed as follows (before any landmark trials are available)

$$\begin{pmatrix} \beta_1 \\ \beta_2 \end{pmatrix}_{t=0} \sim N\left(\begin{pmatrix} \underline{\beta} \\ \underline{\beta} \end{pmatrix}, \sigma_\beta^2 \begin{pmatrix} 1 & \rho_0 \\ \rho_0 & 1 \end{pmatrix}\right),$$

where $\underline{\beta}$ is the mean initial prior belief about the efficiency ratio of each statin.

Quality Signal

Let β_j be the true mean level of the efficiency ratio for drug j. A noisy but unbiased observable signal from clinical trial l for drug j is

$$\tilde{\beta}_{jl} = \beta_j + \zeta_l,$$

where $\zeta_l \sim N(0, \sigma_\zeta^2 / N_l)$ and N_l denotes the number of patients who participate in landmark clinical trial l.

Updating Process for Drug 1

Assume that a physician learns about clinical trial l for drug 1 at time t.

Her posterior belief on the efficiency ratio of drug 1 is

$$\beta_{1t+1} = \beta_{1t} + \frac{\sigma^2_{\beta 1t}}{\sigma^2_{\beta 1t} + \sigma^2_{\zeta 1l}} \cdot (\tilde{\beta}_{1l} - \beta_{1t}).$$

Her posterior variance on the efficiency ratio of drug 1 is

$$\sigma^2_{\beta 1t+1} = \frac{\sigma^2_{\beta 1t}\sigma^2_{\zeta 1l}}{\sigma^2_{\beta 1t} + \sigma^2_{\zeta 1l}}.$$

Updating Process for Drug 2

Assume that a physician learns about clinical trial l for drug 1 at time t.

Her posterior belief on the efficiency ratio of drug 2 is

$$\beta_{2t+1} = \beta_{2t} + \frac{\pi_t}{\sigma^2_{\beta 2t} + \sigma^2_{\zeta 1l}} (\tilde{\beta}_{1l} - \beta_{1t}).$$

where π_t is the covariance in prior beliefs about "efficiency ratio" of drug 1 and 2 at time t.

Her posterior variance on the efficiency ratio of drug 2 is

$$\sigma^2_{\beta 2t+1} = \sigma^2_{\beta 2t} - \frac{\pi_t^2}{\sigma^2_{\beta 2t} + \sigma^2_{\zeta 1l}}.$$

Types of Physicians

- We extend the model proposed by Ching and Ishihara (2010).
- Informative detailing is a means to build and maintain the measure of physicians.
- A physician is either well-informed or uninformed about drug j.
  - A well-informed physician knows the most current landmark trials of drug j ($I_j(t)$).
  - An uninformed physician only knows the initial prior ($\bar{I}_j(t)$).

Informative Detailing and Publicity

- The probability that a physician will learn the most updated clinical information about drug j at time t is

$$M_{jt} = \frac{exp(\alpha_0 + \alpha_d \cdot I\_STK\_detail_{jt} + \alpha_p \cdot STK\_rh_{jt})}{1 + exp(\alpha_0 + \alpha_d \cdot I\_STK\_detail_{jt} + \alpha_p \cdot STK\_rh_{jt})},$$

where $I\_STK\_detail_{jt}$ and STK_rh_{jt} denote the informative stocks of detailing and drug specific non-comparison publicity in reducing heart disease risks for drug j at time t, respectively.

Utility Function

Let patient i's utility of consuming statin j at time t be

$$U_{ijt} = \omega \cdot q_j^h + \lambda_j + \epsilon_{ijt},$$

where q_j^h denotes drug j's efficacy in reducing heart disease risks; b_j captures time-invariant brand specific preference.

Physician k's expected utility of prescribing drug j to patient i at time t becomes

$$E[U_{ijt}^k | I^k(t)] = \omega \cdot E[q_j^h | I^k(t)] + \kappa_d \cdot P\_STK\_detail_{jt} + \lambda_j + \epsilon_{ijt},$$

where $P\_STK\_detail_{jt}$ is a persuasive detailing goodwill stock for drug j at time t.

- The total demand for drug j at time t is expressed as follows:

$$d_{jt} = \hat{d}^1_{jt} + \hat{d}^2_{jt} + \hat{d}^3_{jt} + e_{jt}$$

where $\hat{d}^1_{jt}$, $\hat{d}^2_{jt}$, $\hat{d}^3_{jt}$ are estimated demand for drug j at time t from "new patients", "stayers" and "switchers", respectively; e_{jt} is a measurement error.

- Estimate the model using Maximum Likelihood.

Estimation (cont'd)

- Demand due to new patients is expressed as follows:

$$\hat{d}^1_{jt} = (m_t - \sum_{r=1}^{J} d_{rt-1}) \cdot P_t(statin) \cdot \sum_{k_{type}=1}^{2^H} P_t(k_{type}) \cdot P_t(j|statin, k_{type}).$$

- Demand due to switchers is expressed as follows:

$$\hat{d}^3_{jt} = \sum_{m=1,\neq j}^{J} \{d_{mt-1} \cdot S_{mt} \cdot \sum_{k_{type}=1}^{2^H} [P_t(k_{type}) \cdot \frac{exp(U_{kjt}(k_{type}))}{\sum_{r=1,\neq m}^{J} exp(U_{krt}(k_{type})))}]\}.$$

Identification

- Correlated Learning
  - "Initial prior" captures physician's belief prior to the release of any landmark trials.
  - Drugs entered the market at different point of time.
  - Hence, they faced different existing stock of trials at their entry dates.
  - The value of ρ identified by the differences of the initial sales for different drugs.
  - Changes in demand after a clinical trial is released also help.

Identification

- Informative Detailing
  - Variations in demand and detailing before and after each clinical trial release identify the informative effects.
- Persuasive Detailing
  - It picks up the data variation not accounted by learning and informative detailing.

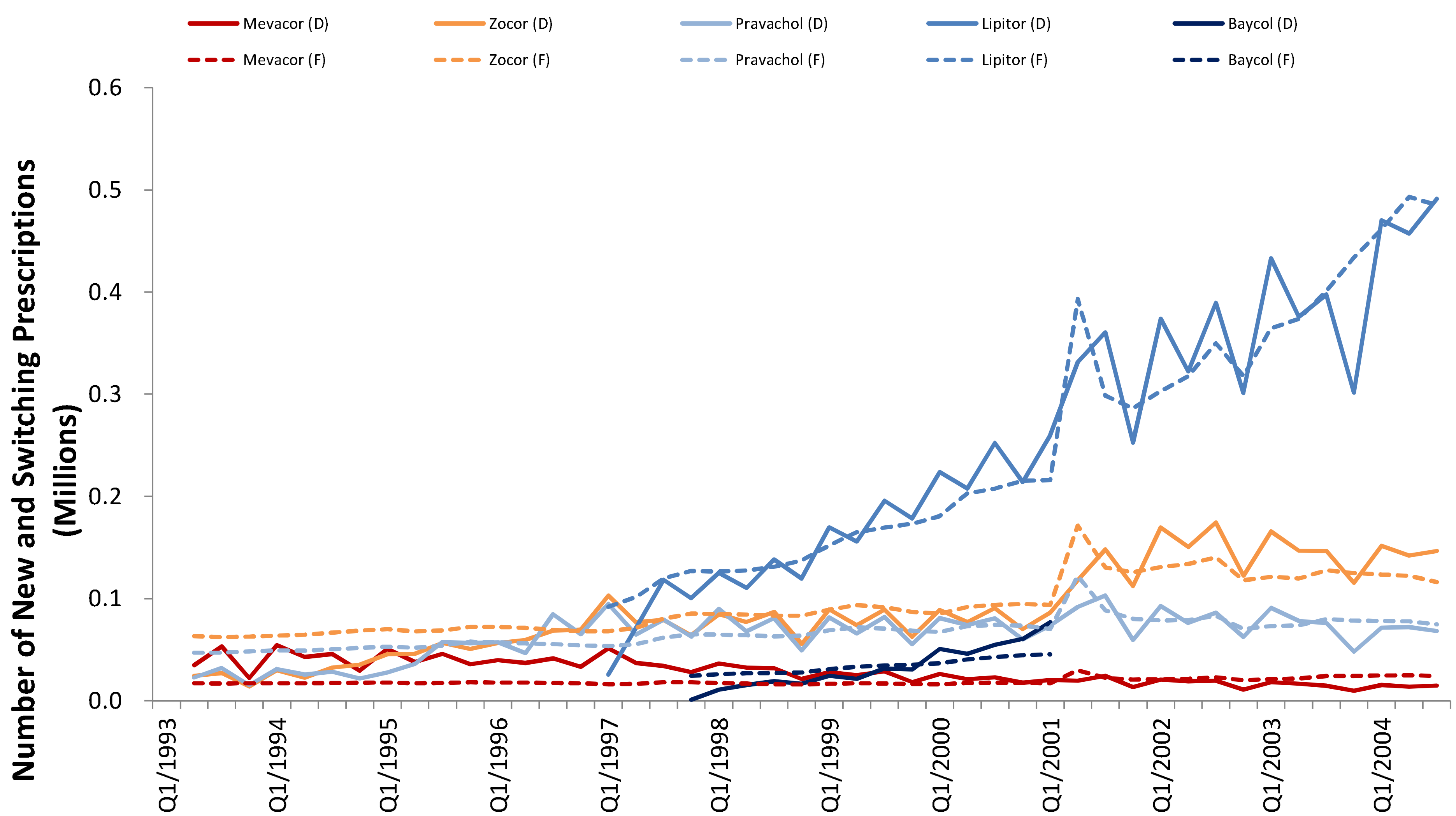
Results Table 1

Variable Descriptions	Estimates	S.E.
Statin Choice Stage		
Learning Parameters		
β (Initial Prior Belief on Efficiency Raito)	0.0851	0.0210
σ_β^2 (Initial Prior Variance on Efficiency Raito)	1.2435	0.0962
σ_ε^2 (Signal Variance from Different Design)	0.5227	0.2069
σ_ζ^2 (Signal Variance from 1,000 Patients)	5.4968	2.2340
ρ_0 (Correlated Learning Parameter in Initial Prior)	0.8247	0.0607
Parameters Determining Measure of informed Physicians		
α_0 (Constant)	-5.7250	0.6151
α_d (Informative Detailing)	2.5943	0.2180
δ_{d_inf} (Carryover Rate of Informative Detailing in Statin Choice)	0.8999	0.0143
δ_{rh_inf} (Carryover Rate of Informative Publicity in Statin Choice)	0.2142	0.0290
Utilty Parmaeters		
ω_h (Coefficient of Perceived Quality in Reducing Heart Disease)	2.0144	0.3112
κ_d (Persuasive Detailing)	1.0735	0.0828
δ_{d_per} (Carryover Rate of Persuasive Detailing in Statin Choice)	0.9272	0.0077
Log Likelihood	-2695.46	

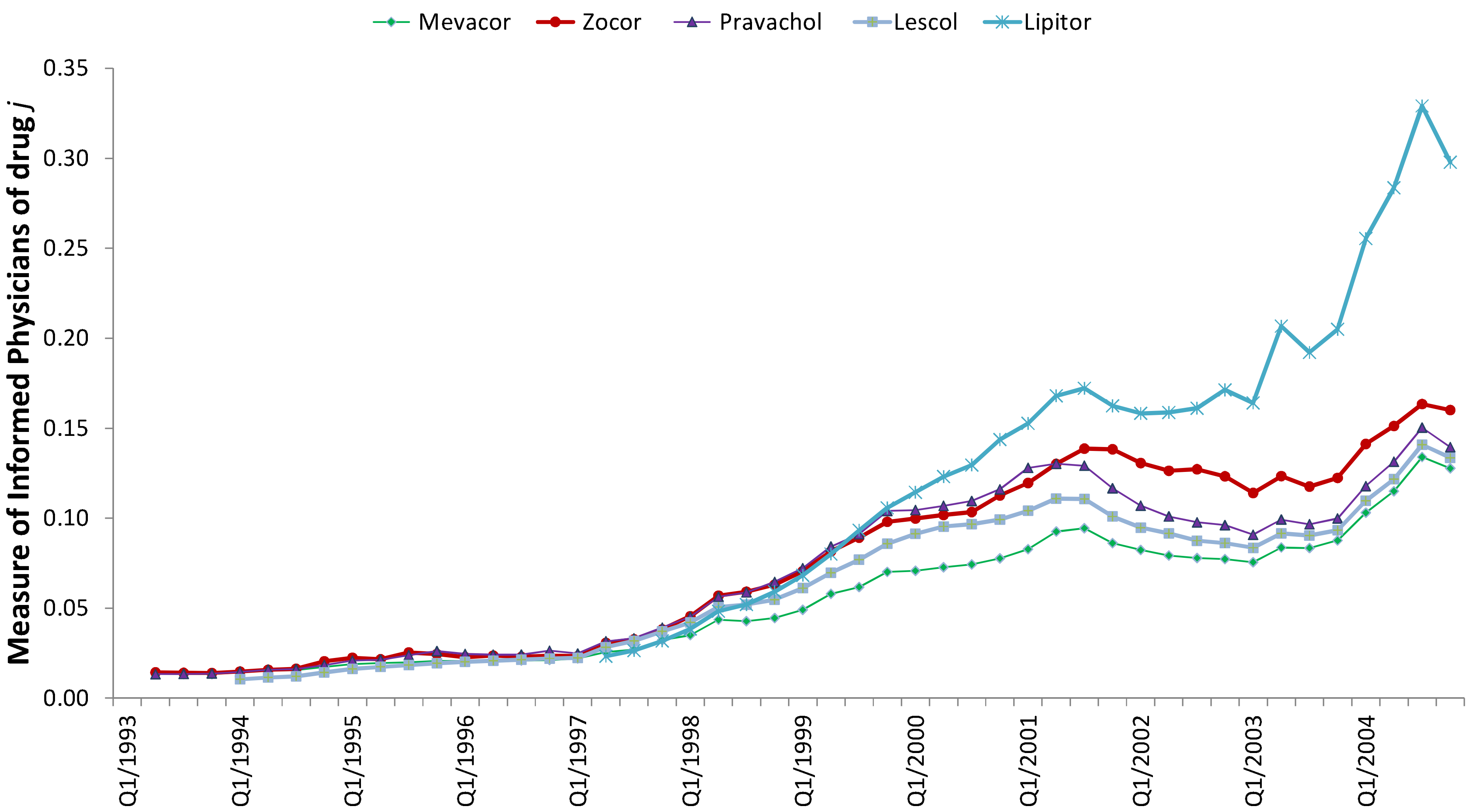
Results

- The estimate of the correlated learning parameter (ρ_0) is 0.82, which suggests a fairly high information spill-over.
- The estimates of both persuasive (κ_d) and informative (α_d) detailing parameters are positive and significant.
- Publicity in reducing heart disease risks (α_{rh}) has a significant impact on updating physicians about clinical trial information.

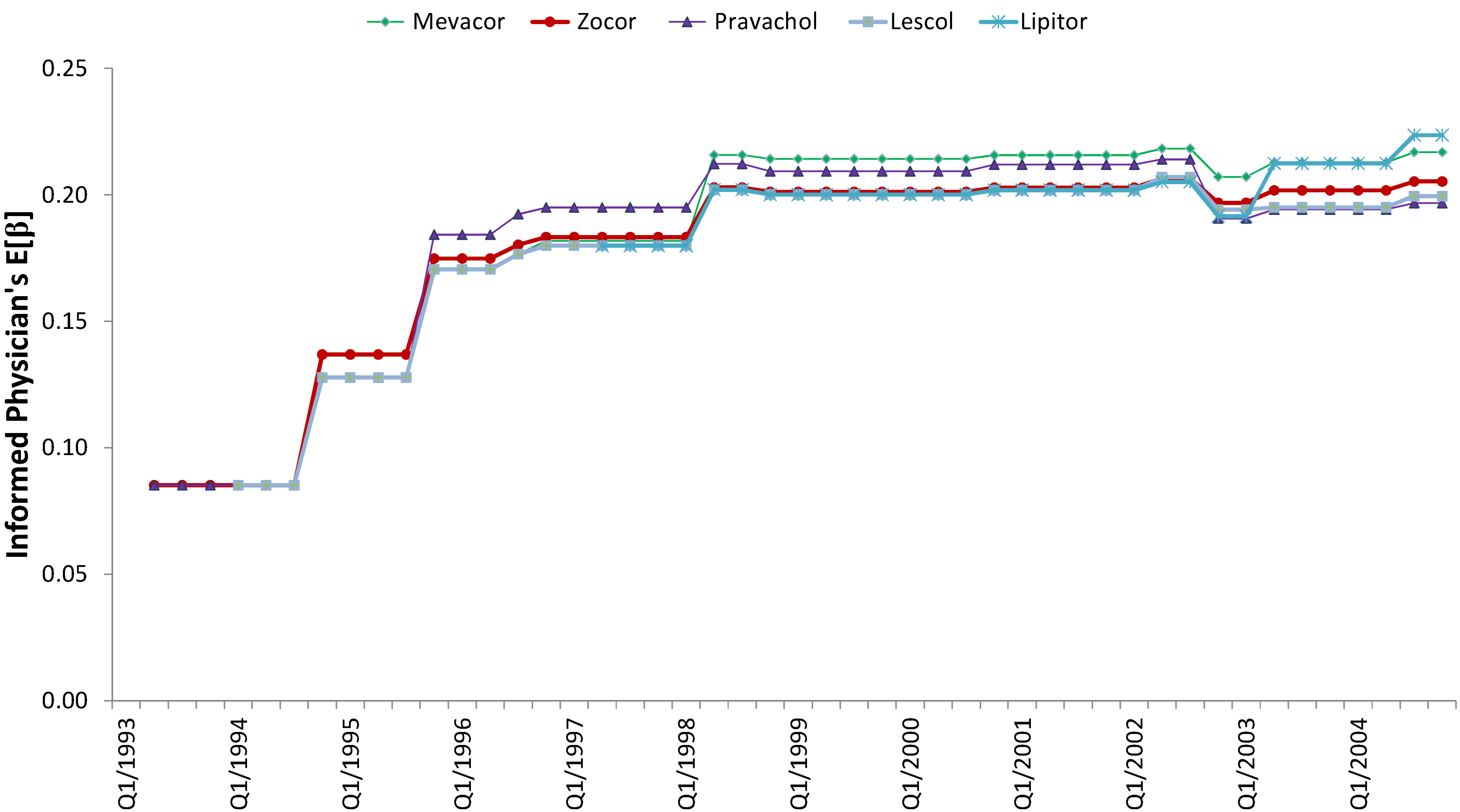
Goodness-of-fit



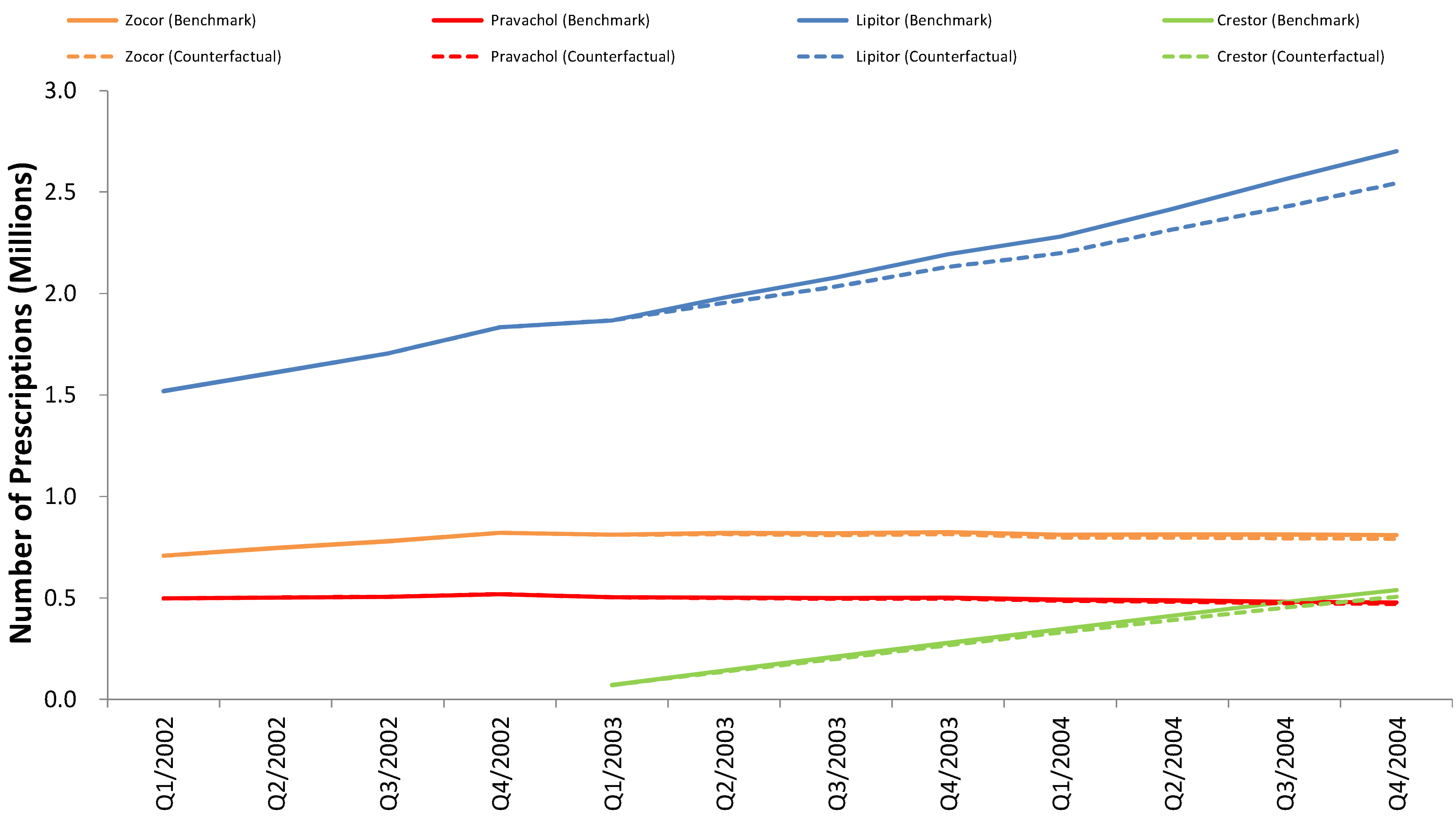
Measure of Well-informed Physicians



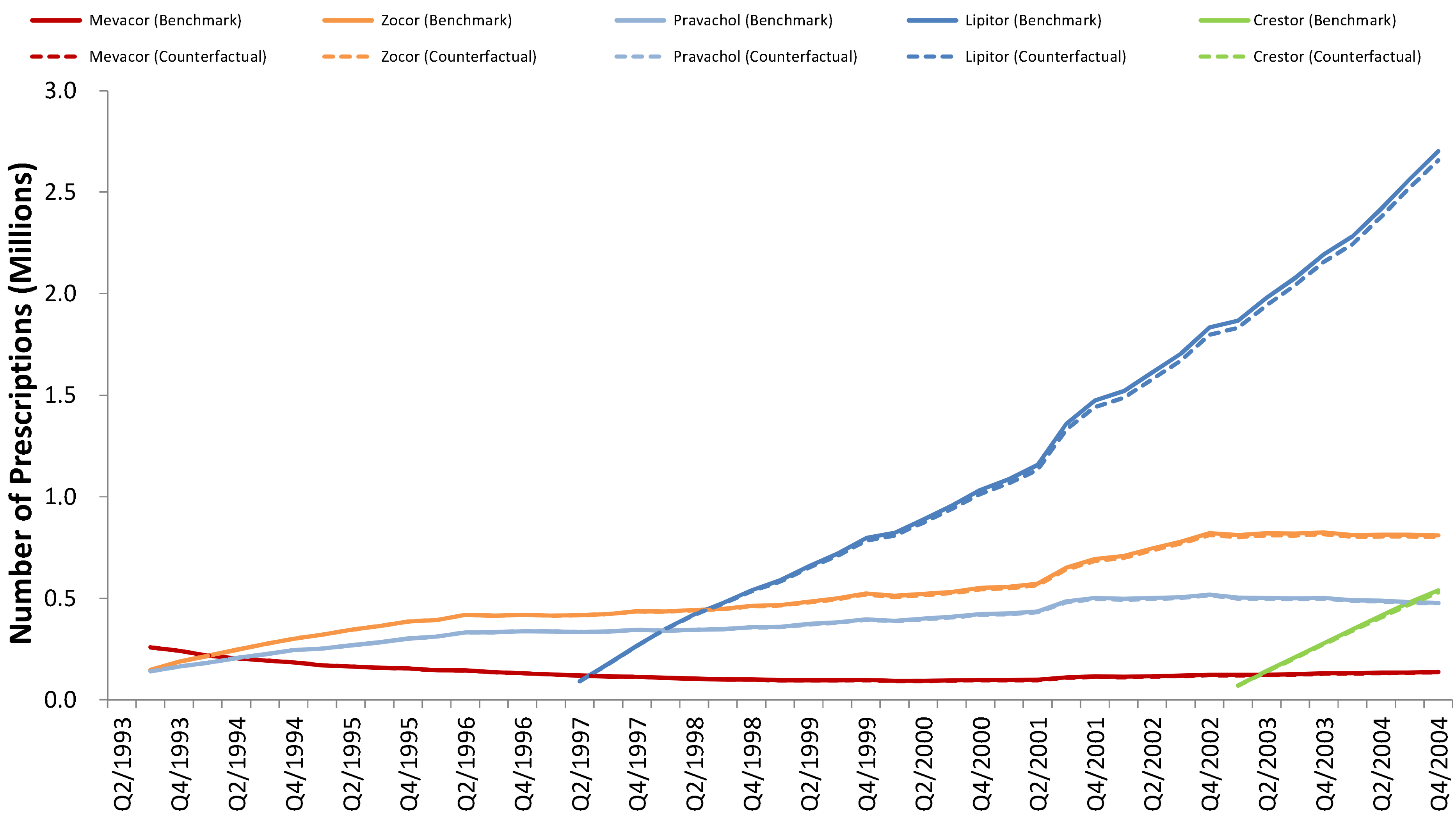
Learning of Well-informed Physicians



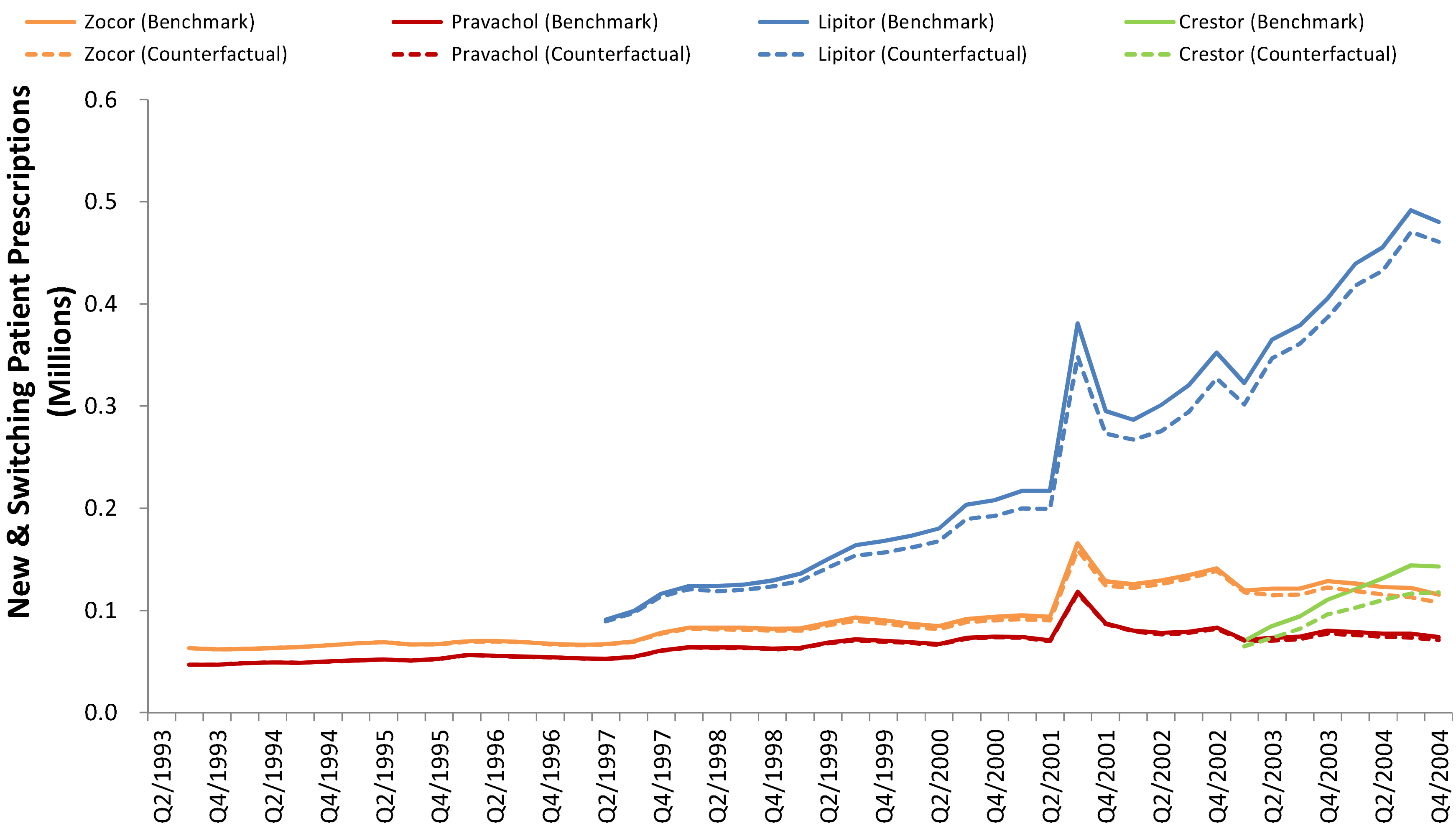
Expt 1: No Landmark Trials for Lipitor



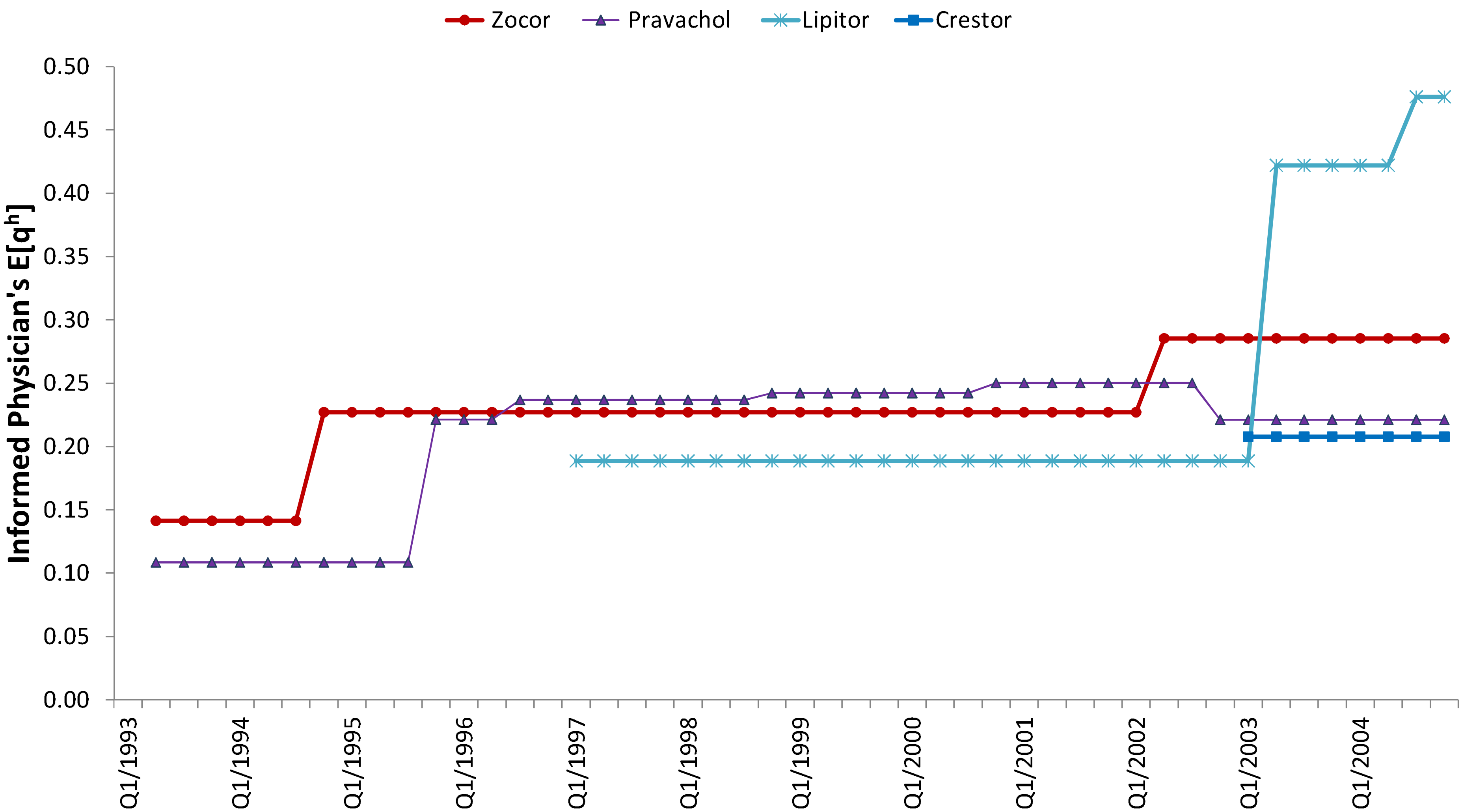
Expt 2: No Landmark Trials for Mevacor



Expt 3: No Correlated Learning



Expt 3: No Correlated Learning



Results

- First experiment: Without landmark clinical trials, the counterfactual new patients' demand for Lipitor is 2-5.8% lower for most quarters.
- Second experiment: Without landmark clinical trials of Mevacor, the counterfactual new patients' demand for Lipitor is 1.8-2.0% lower for most quarters.
- Third experiment: The result indicates that correlated learning plays a role for the early success of both Lipitor and Crestor
- Correlated learning is not the only driving force for the rapid success of Lipitor. Lipitor's superior efficacy in lowering cholesterol level, and detailing are also important factors.

Conclusion

- Our results suggest that late mover advantages can be generated by correlated learning.
- Although Lipitor can free-ride on incumbents' clinical trials, its own clinical trial still plays an important role in generating demand.
- This model can be extended to other market where some products qualities are uncertain, e.g., Computer, Tablet.