Internet Advertising with Information Congestion

An Empirical Investigation

Dong Ook CHOI (KDI)

e-mail: dochoi@kdi.re.kr

March 25, 2016

Korea Academic Society of Industrial Organization

Outlines

Research Objective I

How does *the limitation in consumer's attention span* affect the Internet advertising markets?

- Economics of a platform w/ an additional externality.
- Structural estimation.

Information congestion

Information congestion occurs due to *the limitation in consumer's attention span*:

- Typically, consumers cannot process all the ad messages.
  - Consumer's attention is not an unlimited resource.
- Advertisers keep sending messages if $WTP \geq mc$.
  - Firms compete for consumers' attention.
  - Advertisers do not pay consumers for accessing attention.
  - A common property problem.
- junk mails or spam mails.

Information congestion

Information congestion = *an externality*

- Mismatch between the advertisers and the consumers:
  - Possible screening out of the right message for a consumer.
  - Loss for the corresponding advertiser.

Monopoly Platform

What if there is a platform between advertisers and consumers?

Monopoly Platform

Competitive bottleneck:

- e.g.) single-homing readers and multi-homing advertisers.
- Competing platforms become *de facto* monopolists.

Monopoly Platform and Information Congestion

If there is a monopoly media firm, the congestion should be *perfectly* priced out (Anderson and de Palma, RAND 2009).

The monopoly media firm would set its ad level equal to *the average attention level* when maximizing its profit.

- (A1) Homogeneous viewers.
- (A2) Single-homing viewers.

Monopoly Platform and Information Congestion

- I relax (A1) and show that information congestion occurs even with the monopoly platform.
  - Heterogeneity of consumers in attention span.
  - Single-homing viewers: a monopoly platform.
  - No theoretical result yet in previous studies.

Theories on Information Congestion

Van Zandt (RAND 2004), Anderson and de Palma (RAND 2009)

- Seminal studies on information congestion.
- Message senders and receivers (no platforms).
- Imposing tax for msg delivery would resolve congestion.
  - The e-mail tax suggested by Bill Gates.

Contribution to IC theory

By introducing information congestion into media platform,

- Endogenous ad cost (ad price or quantity is set by platforms)
- Endogenous consumers' decision to participate.

Theories on Two-Sided Market and Media Econ

- Two-sided market theory:
  - Network externalities in one side of the market from the other side.
  - Rochet and Tirole (2003), Caillaud and Jullien (2003), Armstrong (2006).
- Media economics:
  - Anderson and Coate (2005), Anderson and Gabszewicz (2006), Ambrus et al. (2012).
  - Negative impact of ads on viewer demand and positive impact of viewership on ad demand.
- My study introduces an additional externality of limited attention.

Empirical Literature

Two-sided market empirical studies:

- Rysman (RES 2004) – Yellow pages.
- Wilbur (MktSci 2008) – Television.
- Choi et al. (RIO 2012) – Search engines.

My contribution:

- Heterog. in consumer information in ad demand function.
- Individual choice data in consumer market.

Online Display Advertising

① Big banner
- size : 400*100
- file size : 20K
- price : CPM $3

② Main text
- size : 18 words including spaces
- price: $2,500 / week

③ Margin banner
- size : 90*180
- file size : 10K
- price : $5,000/day

④ Right banner
- size : 235*110
- file size : 20K
- price : CPM $2
- note : stops after play once/ < 5sec

⑤ Floating ad
- size : 600*285/reminder 90*100
- file size : 30K/reminder 10K
- price : CPM $10
- note : transparent/no border line
/less than 8 sec

Data

Internet banners (display ads) on 6 websites:

- ResearchAd (http://www.researchad.com).
- Ad price, ad quantity, contract periods, number of sections, ...
- "Naver.com," "Daum.net," "Nate.com," "Yahoo.com/kr," "Empas.com," and "Paran.com."

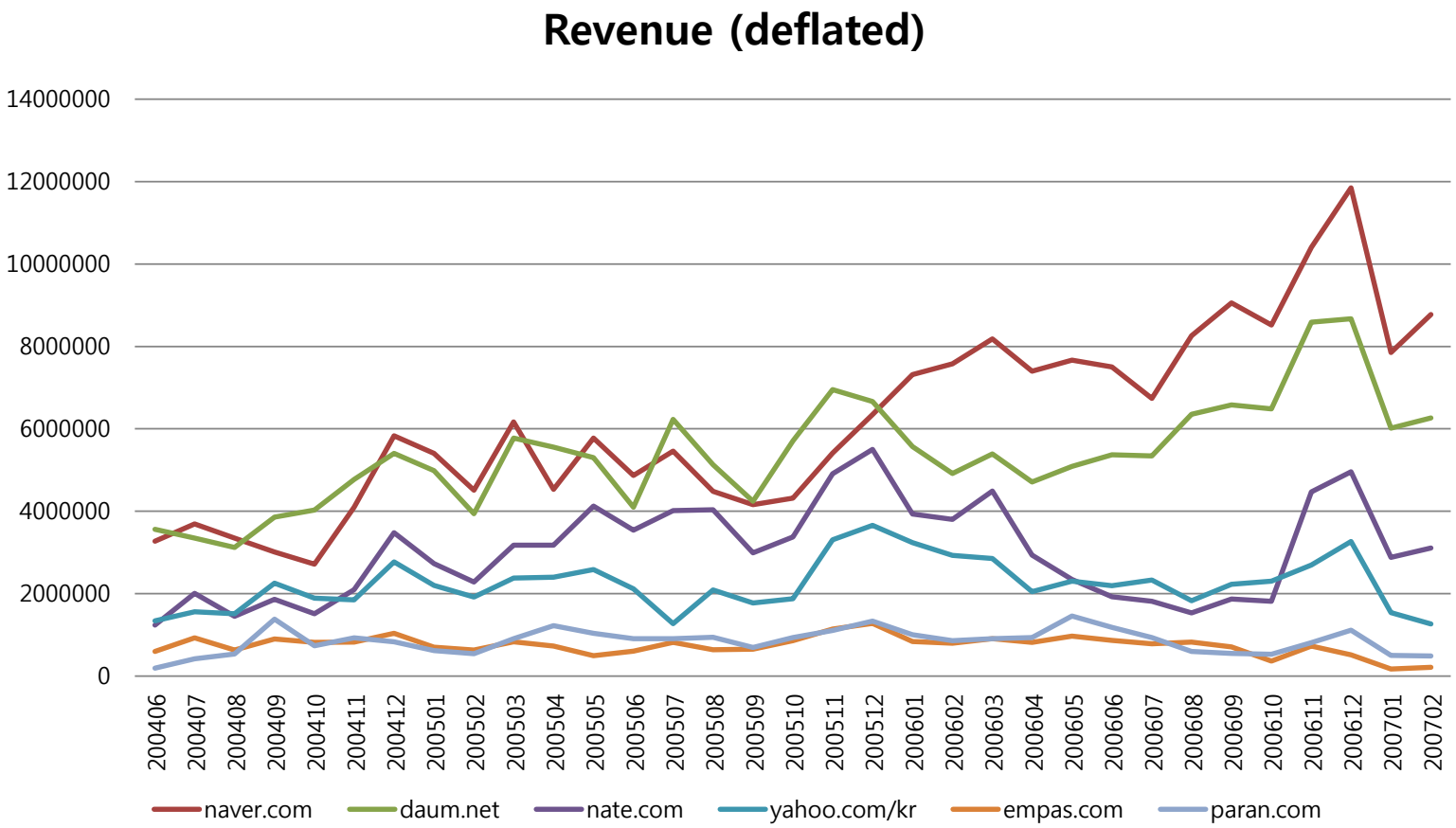
Internet traffic data on user side:

- Nielsen KoreanClick (www.koreanclick.com).
- Demographics (age, edu, income) and performance metrics (PV, duration, days ...).

Complete match of "both sides of the markets"

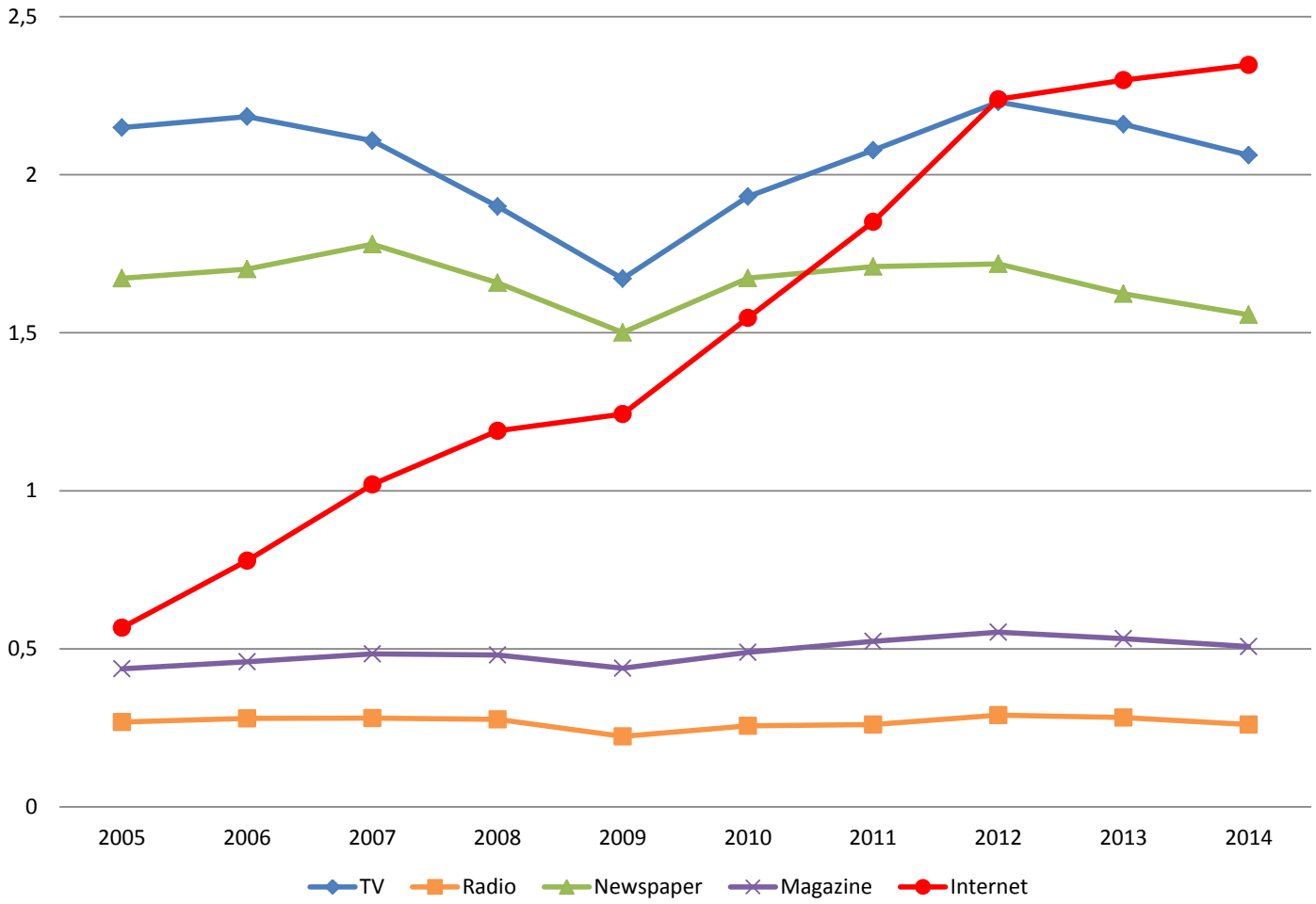
- June 2004 – February 2007.
- 198 obs. for aggregate level data (33 months x 6 websites)
- 1,157,246 obs. for individual choice and demographics data

Data



Revenue Growths of Banner Ads in Six Search Engines.

Data



Korean advertising market (in billion USD).

Some Descriptive Statistics

Average Korean search engine:

- Serves 20mil users/month.
- Provides 18.4 sections of different contents.
- 70.5 advertisers post 326.1 ads/month.
- Avg. ad price is 18,581 USD for 6.8 days of contract.

Some Descriptive Statistics

Naver.com – largest search engine in South Korea,

- Shows about 100 ads/day.
- Page views/user/day $= 47$.
- Monthly CTR $= 0.4\%$.

Internet Display Ads Example

Internet Display Ads

Advantage of Website Banners for Studying Information Congestion.

- Unsolicited way of advertising:
  - Congestion happens when consumers ignore the message.
  - Consumers normally have ad-nuisance.
  - Advertisers are from various industries.
  - c.f. Yellow Pages or search ads – business stealing effect.
- No targeting or tracking technologies.
  - During study period (2004–2007).
  - Not using individual information.
  - Weak targeting: subsections for certain topics.

User's Attention-Giving Behavior

Table 1: The process of drawing consumers' attention according to advertising and media types.

Media type	Targeted	Nuisance[b]	Stage 1 (Recognition)	Stage 2 (Decision)	Stage 3 (Attention)	Stage 4 (Action)	Related Study
Direct mails	Y	Y	Receive	Open/Throw-away	Examine	Contact Advertisers (make a phone call, click on a banner, etc.)	Anderson and de Palma (2009) and Van Zandt (2004)
Telemarketing	Y	Y	Hear (Phone ring)	Answer/Ignore	Listen		
Website banners	N[a]	Y	Notice	Look/Ignore	Examine		Choi et al. (2012)
Search ads	N[a]	N	Search	(none)	Examine		*
Newspaper	N[a]	Y	Read	Read/Ignore	Examine		*
Television	N[a]	Y	Watch	Watch/Switch	Exam/Listen		Wilbur (2008)
Yellow Pages	N[a]	N	Search	(none)	Examine		Rysman (2004)
Magazines	Y	N	Read	Read/Ignore	Examine		Kaiser and Wright (2006)

[a] I consider only general type of media here. However, it is always possible for these media to target a certain group of people.
[b] This is the net effect of recipients' attitude towards the ad messages. "Y" means that the recipients feel nuisance when they are exposed to signals of message delivery.

Reduced-Form Analysis

Some results on market and user characteristics:

- Contents affect ad quantity.
  - e-Commerce (+), community (+).
  - Game (–), mini-hompage (–).
- Contract periods and user demographics affect user visits.
  - Avg. contract period (–), male% (–), age (–)
  - Student% (+), education level (+).
- Ad quantity negatively affects duration and revisit.
- N. of sections positively affects duration and revisit.

Reduced-Form Analysis

Some results on congestion effects:

- Ad quantity negatively affects CTR.
- 1% increase in total ad quantity on a website decreases the advertiser's average ad price by 0.29%.

A Model of Information Congestion

– If ad quantity is excessive, msgs can crowd out each other

$$Pr\left(msg\_process\right) = \min\left[1, \frac{m}{A}\right]$$

m: viewer's attention span – num. of msgs processed

A: total messages sent.

- No spill-over effect among messages (no complementarity or substitutability)

The Model - Advertising Demand

Advertiser i, Platform j, Period t.

- Assumption 1: Single-Homing Users.
- Assumption 2: Proportionality.
- Assumption 3: Small Advertisers.

Look Function, $L_{jt}(a_{ijt}, \phi_{jt}) = a_{ijt}^{\alpha}\phi_{jt}^{\beta}$: Expected exposure/contact

Total Profit of Advertiser i:
$\Pi it = \pi_{1t}L_{1t} - P_{1t}a_{i1t} + ... + \pi_{Jt}L_{Jt} - P_{Jt}a_{iJt}$.

By FOC and aggregation: $A_{jt} = \left(\frac{P_{jt}}{\alpha\pi_{jt}\phi_{jt}^{\beta}}\right)^{\frac{1}{\alpha-1}}$.

- π_{jt}: Advertiser i's profit per Look.
- P_j: Ad price of platform j.
- A_j: Ad quantity of platform j.
- ϕ_j: Expected number of ad noticing in j.

Inverse Demand for Advertising

$$\ln P_j = \alpha^p \ln A_j + \beta^p \ln \phi_j + \mathbf{X_j}\omega^p + \upsilon_j.$$

- X_j: Exogenous variables of j.
- υ_j: i.i.d. Normal distribution.
- I expect $-1 < \alpha^p < 0$ and $\beta^p > 0$.

Attention span of users

- Expected user demand: $\phi_j = g(A_j) \cdot U_j$
- U_j: actual number of visitors.
- msg processing rate: $g(A_j) = \int \min\left(1, \frac{m}{A_j}\right) dF(m)$.
- User k has a limited attention span, m.
  - m is i.i.d. random variable following $F(\cdot)$ and $f(\cdot)$.

Simulated moments

Simulating "hypothetical" attention span of users (m_{jk})

$$\ln m_{jkt} = \mathbf{D}_k' \alpha^m + c_j^m + \zeta_{jkt}.$$

- $\mathbf{D}_k$: gender, age, education level.
- c_j^m: website-specific constant.
- $\zeta_{jk} = \sigma^m \epsilon_{jk}$
- Draw ϵ_{jk} from standard normal dist.

Estimation of Ad Demand

Generalized method of moments (GMM)

- Orthogonality condition: $E[Z'\Lambda] = 0$, where $\Lambda = (\upsilon)$ is a vector of error terms.
- GMM objective function: $\Lambda' Z W^{-1} Z' \Lambda$,
  - where Λ is a vector of error terms, Z is a vector of instruments.
  - $W = E[Z'\Lambda\Lambda' Z]$ is a weighting matrix.
  - A set of parameters $\{\alpha^p, \beta^p, \omega^p, \alpha^m, c^m, \sigma^m\}$ is chosen to minimize this.

Identification

- Endogenous variables: A_{jt} and ϕ_{jt} (or U_{jt}).
- Excluded instruments
  - Average behavior of advertisers: average contract periods.
  - Average characteristics of users: gender, job, age, edu, region.
  - Overall market condition: total Internet users.
- In our case, we need to consider panel property of our data set (cf. Nevo, Emet 2001).
  - Some of our exog. characteristics have variation over time.
    - Number of sections, age of websites.
    - Some dummies such as *lmail* and *mhpg*.

Comparison w/ Rysman (2004)

- Rysman (2004) assumes the constant rate of decreasing returns in the aggregate ad level
  - Both models have the structural parameter of the congestion effect but Rysman (2004) cannot identify the parameter separately.
  - the variation of consumers' responses to the ad level change is the source of identification in my model.
- Interpretation of the congestion effect is different.
  - Rysman's interpretation is mostly about business stealing effects.
  - In my model, the only bottleneck is the limited attention span of each individual.

User Demand

- Mixed-logit model (Train, 2003) using individual choice data.
- Can allow for alternative-specific fixed effects (cf. BLP, 1994).
- Maximum likelihood estimation (MLE).

User Demand

- The utility function of user k when accessing website j is:
  - $u_{jkt} = \rho_k A_{jt} + \mathbf{X}_{jt}\lambda_k + c_j^u + \epsilon_{jkt}^u$,
  - $\rho_k \backsim i.i.d. \mathcal{N}(\rho + \mathbf{D}_{kt}\omega, \sigma^\rho)$.
  - $\lambda_k \backsim i.i.d. \mathcal{N}(\lambda, \sigma^\lambda)$.
  - $\mathbf{D}_{kt}$: demographics of k – gender, age, education.

User Demand Estimation

- The choice probability of platform j by user k is:
  - $P_{jkt} = \int \frac{e^{V_{jkt}}}{\sum_{l=1}^{K} e^{V_{lkt}}} \, dG(\rho, \lambda)$,
  - where $V_{jkt} = u_{jkt} - \epsilon^{u}_{jkt}$.
- Maximizing Simulated Log-Likelihood (MSL estimation):
  - $SLL = \sum_{t=1}^{T} \sum_{k=1}^{K} w_{jt} \ln \hat{P}_{jkt}$
  - where w_{jt} is a weight (aggregate market share) to resolve the bias from choice-based sampling (Manski and Lerman,1977).

Equilibrium Ad-Level and Market Power

Profit of the platform:
$\Pi^P_{jt} = P(A_{jt}, \phi(g(A_{jt}), U(A_{jt}|\mathbf{A}_{-jt})))A_{jt} - A_{jt}c_{jt}.$

$$A^*_{jt} = -\frac{P^*_{jt} - c_{jt}}{\frac{\partial P_{jt}}{\partial A_{jt}} + \frac{\partial P_{jt}}{\partial \phi_{jt}}\frac{\partial \phi_{jt}}{\partial g_{jt}}\frac{\partial g_{jt}}{\partial A_{jt}} + \frac{\partial P_{jt}}{\partial \phi_{jt}}\frac{\partial \phi_{jt}}{\partial U_{jt}}\frac{\partial U_{jt}}{\partial A_{jt}}},$$

$$\frac{P_{jt} - c_{jt}}{P_{jt}} = -\left[\underbrace{\varepsilon^P_A|_{g=\bar{g} \text{ and } U=\bar{U}}}_{\text{Traditional Market Power}} + \underbrace{\varepsilon^P_\phi \cdot \varepsilon^g_A}_{\text{Congestion Effect}} + \underbrace{\varepsilon^P_\phi \cdot \varepsilon^U_A}_{\text{Network Effect}}\right].$$

Efficient Ad-Level

Optimal level A^o_{jt}: social benefit equals mc (note that the user level is fixed at equilibrium level):

$$\left.\frac{\partial W_t}{\partial A_{jt}}\right|_{U(A^e_{jt})} = P_{jt}(A_{jt}, \phi_{jt}) + \frac{\partial g_{jt}}{\partial A_{jt}}(A_{jt}) \int_0^{A_{jt}} \frac{\partial P_{jt}}{\partial g_{jt}}(x, \phi_{jt})\, dx - c_{jt}.$$

Social benefit is lower than $P_{jt}(A_{jt}, \phi_{jt})$ due to the negative effect of information congestion.

Parameter Estimates in Ad Demand

The negative slope of the demand curve:

- $\hat{\alpha}^P = \frac{\partial P_j}{\partial A_j}|_{g=\bar{g} \text{ and } U=\bar{U}} = -.7294(.0564)$.
- Elasticity of price wrt ad quantity (the inverse of demand elasticity).
- A slope of advertiser's WTP curve when no externality.
- Consistent to our expectation and significant, $-1 < \hat{\alpha}^P < 0$.

Network externality from user side (effect of ϕ_{jt}):

- $\hat{\beta}^P = \frac{\partial P_j}{\partial \phi_j} = .8864(.0464)$.
- Consistent to our expectation and significant, $\hat{\beta}^P > 0$.
- Expected user demand is shown to be positively related with ad price.

Proposition

Proposition

If website users have a homogenous attention span and $\alpha^p + 1 < \beta^p$*, Information congestion is perfectly priced out. i.e.* $A^*_{jt} \leq m, \forall j$*, where* A^*_{jt} *is the equilibrium ad level and* m *is the attention level of homogenous users.*

Implication

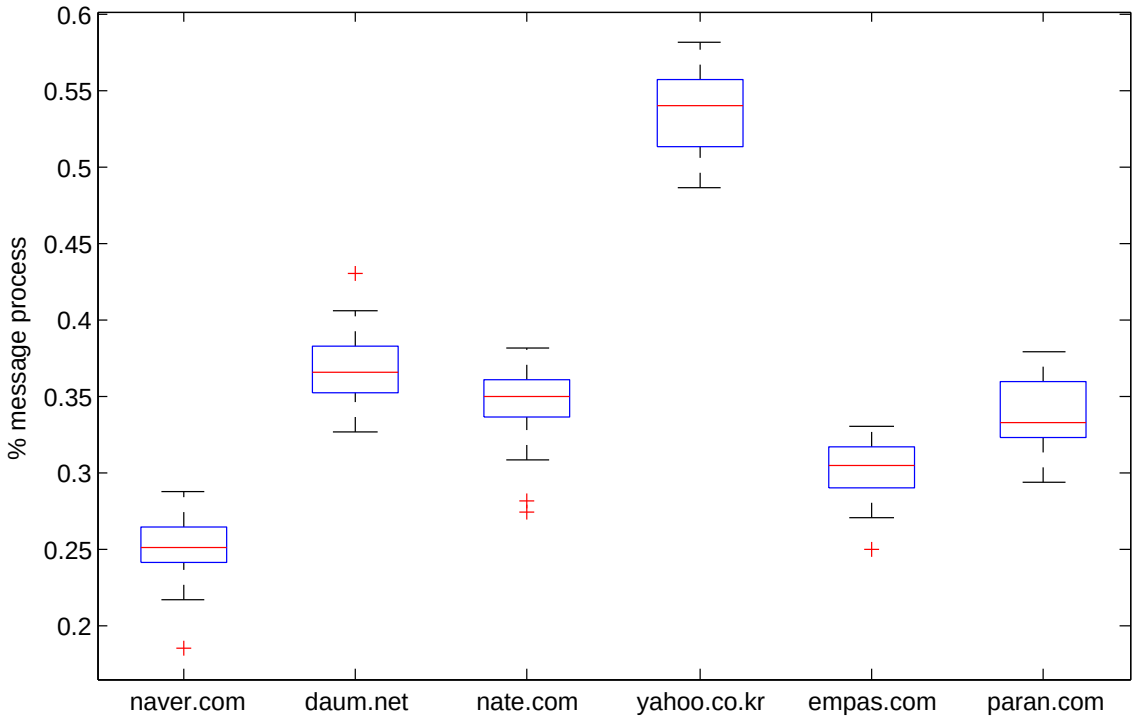
The condition for Proposition ($\hat{\alpha}^p + 1 < \hat{\beta}^p$) is satisfied.

- If information congestion is not priced out by the monopoly platform, it is due to the heterogenous attention spans of consumers.

Parameter Estimates in Attention Level Equation

- Attention level equation:
  - Male users have lower attention level: -1.9227 (.2399).
  - Attention span decreases in age squared: -1.0865 (.2815).
  - Education level is significantly positive: 2.0227 (.1062).

Message Processing Rates



- a significant degree of congestion in the market.
  - The message processing prob. ranges from 18% to 58%.
- The source of variation: ad level, user demographics, and website characteristics (fixed-effects).

User Demand (Mixed Logit)

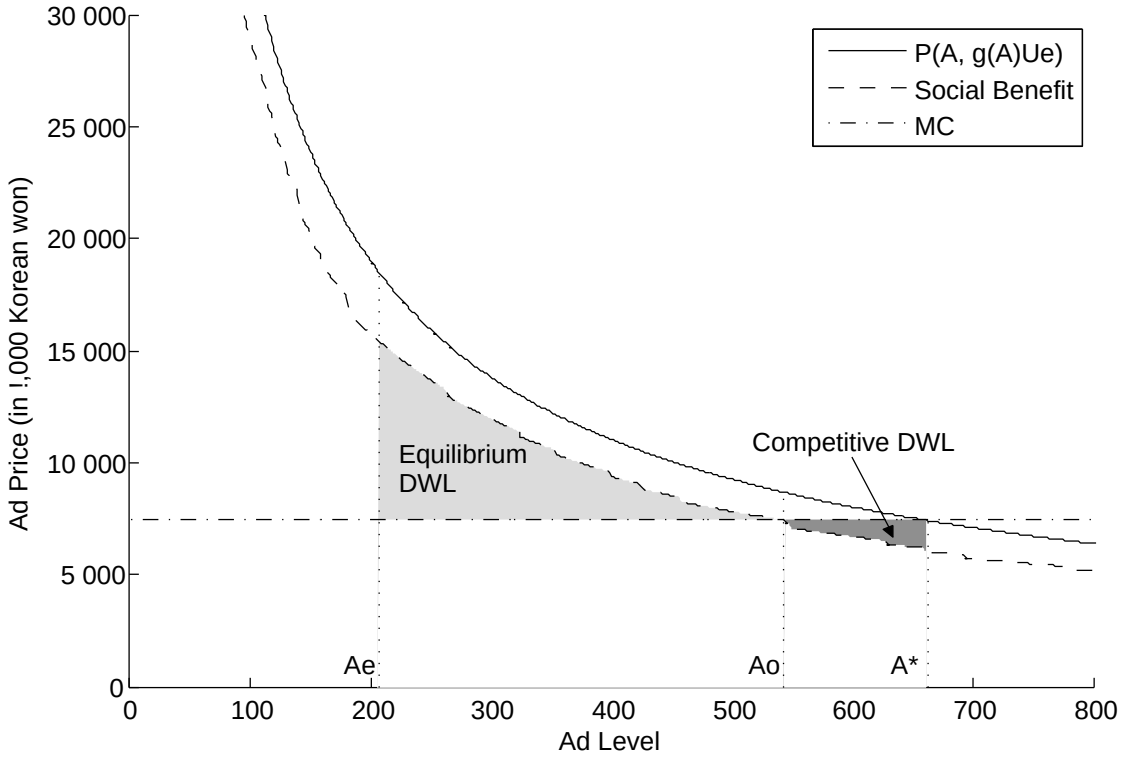
- Parameter distribution for the ad quantity:
  - Estimated mean: -.0038 with the standard deviation: .0003.
  - 92.17% of users have negative effect, *i.e.* nuisance, for the increase of advertising.

Elasticities

	MC	Price Elasticity (ε_A^P)	User Elas (ε_A^U)	Att Elas (ε_A^g)
Pooled Average	3,963.9	-0.7789	-0.0039	-0.0606

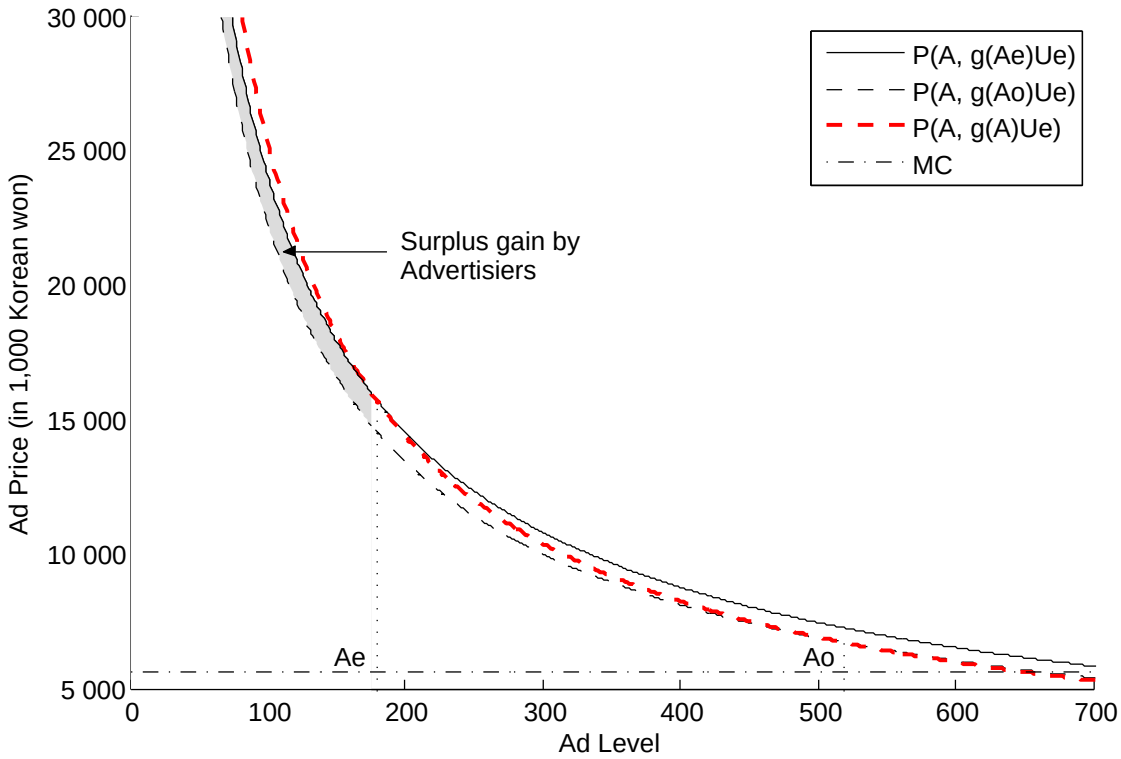
- The price elasticity wrt ad level is a negative value of Lerner Index (a markup or a market power).
  - Overall, these markups are shown to be in (0, 1) to meet the general condition of pricing by monopoly platforms.
- On average, .0039% users are shown to switch to other options when the ad level rises in 1%.
- The markup is explained more by the scarcity of attention than by the nuisance effect in the user demand.

Market Efficiency



- The choice of A^e generates a DWL (equilibrium DWL).
  - Both monopoly market power and information congestion.
  - About 2 billion won per website per month.
- When market moves to competitive level (A^*) (comp. DWL).
  - Negative and is created solely by information congestion.

Market Efficiency



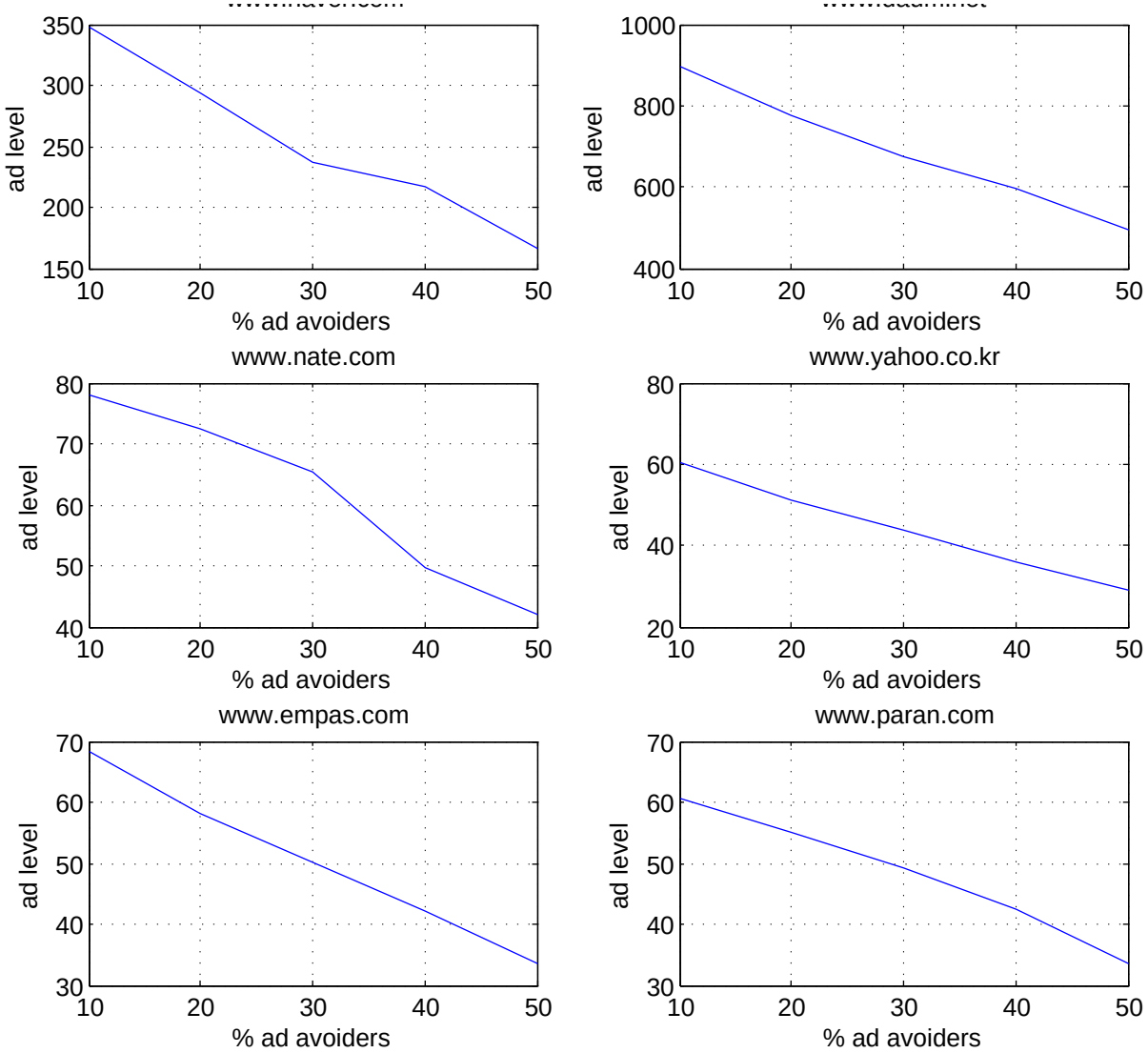
- The willingness-to-pay curve shifts up from A^o to A^e.
- A surplus gain for advertisers by information congestion.
- About 816 million Korean won.

Counterfactual: Paying to avoid ads

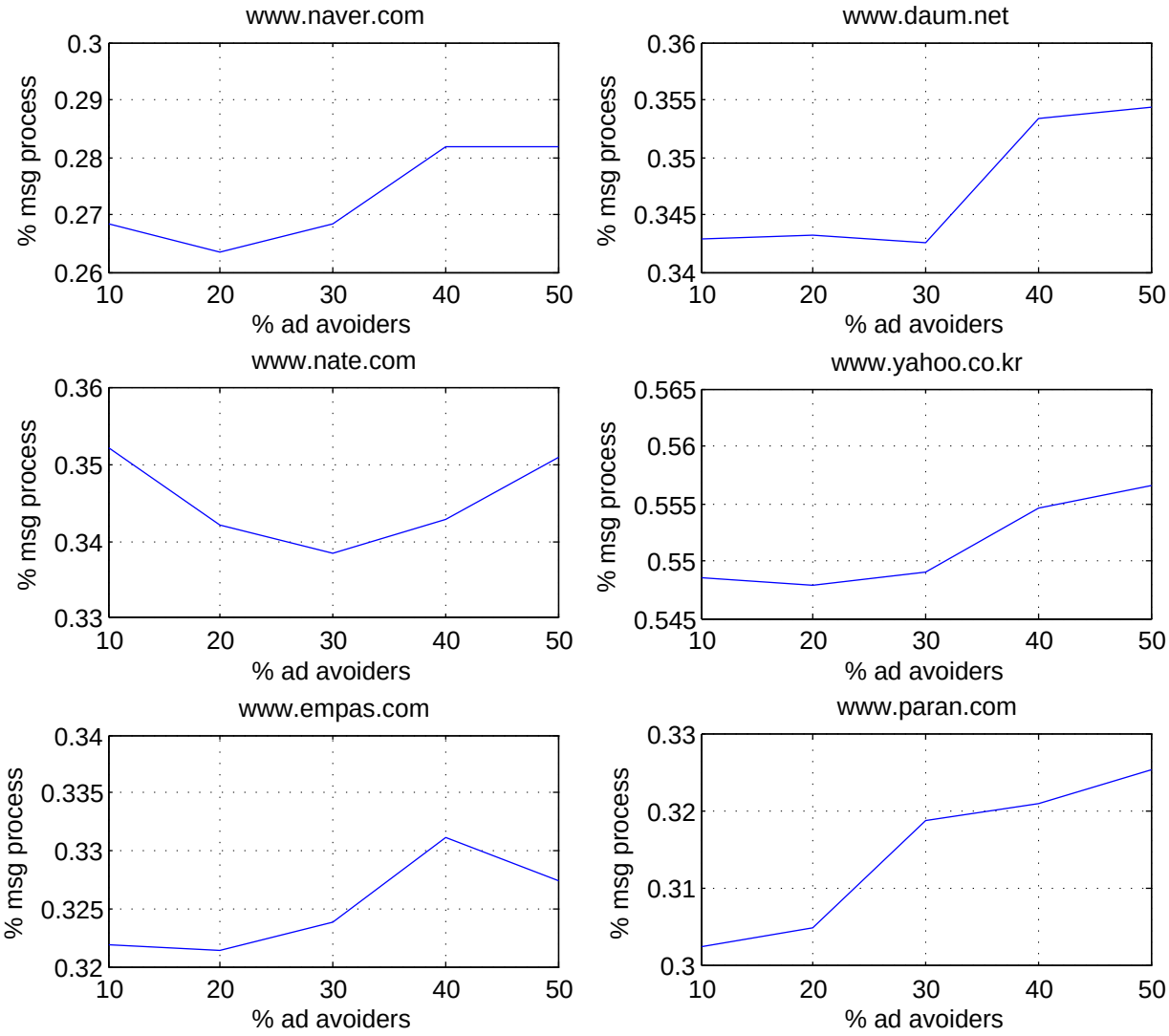
When allow users to avoid ads by paying a subscription fee:

- Ad levels decrease in the share of ad-avoiders.
  - The value of ad-viewers is augmenting
  - The average attention level grows as the share of ad-avoiders increases.
- Opposite to the result of Wilbur (2008) where the ad level rises in the ad-avoiders.
  - Ad-viewers are becoming less sensitive to the ad level and this effect dominates the value of ad-viewers

Counterfactual: Ad-Levels



Counterfactual: MSG processing rates



Summary

- Probabilities of processing messages by users.
  - Range from .18 to nearly .58 depending on websites.
  - Source of differentiation: the advertising level, user demographics, and website characteristics
- Male, the aged, and the less educated have lower attention span.
- 92.17% of users are shown to have ad-nuisance.
- The attention-advertising elasticity, ε_A^g, is higher than the users-advertising elasticity ε_A^U on average.

Summary

- Equilibrium ad level is lower than the social optimal level, generating DWL as 2 billion Korean won on average.
- The average surplus gain of advertisers is 816 million Korean won on average.
- Ad-avoid experiment shows that the value of ad-viewer increases so that the platform reduces ad-level.

Discussion

- Interpretation of platform's market power:
  - information congestion makes ad-demand curve less elastic.
  - monopolist's DWL includes the congestion effect.
- Problem of market definition:
  - competitive bottleneck if single-homing users.
  - user share is meaningless.
- Merger issue:
  - standard theory (Anderson and Coate, 2005): merger – price up.
  - in reality, merger – price down (merged firm internalizes the congestion effect).

Discussion

- More channels: new media such as mobile phone, e-book, digital watch, electronic car, IoT, smart media ...
- Will be more severe competition for your attention.

- I show that congestion is generated even with a platform.
- Individual point of view: always excessive messages to be delivered.
- Advertiser point of view: platforms exploit market power and reduce the opportunity.

- Question: more opportunities to advertisers? or protect consumers' right to be silent?

Limitation and Further Works

- Relaxing single-homing users assumption.
- Relaxing homogenous ad messages (introducing matching between ads and consumers).
- Further counterfactual experiments:
  - Entry simulation.
  - Introduction of targeting tech.
- The effect of interaction w/ keyword search ads.
- Economy of attention considering other media channels.

Thank you!