Competing with Big Data

Jens Prüfer* and Christoph Schottmüller**

*Tilburg University, Tilec
**University of Copenhagen

Outline

The Economist (Feb 6, 2016)

- "Today Alphabet is a giant advertising company with the potential to become a giant in other sectors as well—although exactly which ones, no one is yet sure. [. . .]
- As users spend more time with Google's services, the company learns more about them and sells more ads. Other firms have struggled to profit as much from users' engagement. On February 2nd Yahoo, a struggling rival, announced it was cutting 15% of its workforce and suggested it would consider selling its core internet business, which could put its boss, Marissa Mayer, out of a job. [. . .]
- The firm has started to look like a conglomerate, with interests in areas such as cars, health care, finance and space, as it tries to find the next big thing."

Some research questions

- Why is a firm like Google so successful whereas its competitors have more and more troubles?
- What is the role of "data" in this success story?
- How to explain Alphabet's strategy of turning a lean and focused company into a conglomerate?
- Which industries may be next on Alphabet's entry list?
- Is there a role for competition policy or regulation?

This paper

- defines "data-driven markets" and constructs a dynamic model of R&D competition in such markets
- shows that such markets tend to "tip"
- discusses how a dominant firm in one market can leverage its position to another ("connected") data-driven market ("domino effect")
- discusses some cases related to our story
- discusses a regulatory measure avoiding tipping

Basic observation

- User data is an input in R&D
- Bigger market share today —> cheaper R&D today —> bigger market share tomorrow —> ...

Is this like. . .

- . . . learning curve effects (dynamic economies of scale)?
  no, those are supply driven and in principle replicable by competitors
- . . . network effects?
  no, users do not care about other users
- . . . indirect network effects in computer consoles?
  no (though closer), you do not depend on third parties and new entrants in computer consoles can write software themselves

Model I

- unit mass of consumers each with unit demand in each period $t \in \{1, 2, \ldots, T\}$
- duopoly of firms ($i = 1, 2$) with quality $q_{i,t}$
- quality difference $\Delta_t = q_{1,t} - q_{2,t}$
- period t demand

$$D_1(\Delta_t) = \begin{cases} \frac{1+\Delta_t}{2} & \text{if } \Delta_t \in [-1, 1] \\ 1 & \text{if } \Delta_t > 1 \\ 0 & \text{if } \Delta_t < -1 \end{cases} \qquad D_2(\Delta_t) = \begin{cases} \frac{1-\Delta_t}{2} & \text{if } \Delta_t \in [-1, 1] \\ 0 & \text{if } \Delta_t > 1 \\ 1 & \text{if } \Delta_t < -1 \end{cases}$$

Model II

- firms choose innovation levels $x_{i,t}$: $q_{i,t} = q_{i,t-1} + x_{i,t}$
- cost of innovation $c(x, D_{i,t-1}) = \gamma x^2/2 + \alpha x(1 - D_{i,t-1})$, $\gamma > 0$, $\alpha \in (0, 1]$
- revenue proportional to demand
- firm 1 (2) can innovate in odd (even) periods
- discount factor $\delta \in [0, 1)$
- SPNE

Discussion of model assumptions I

- demand: Hotelling model
    - covered market
    - single homing users
- revenues: ad financed service, e.g. two-sided market
    - mass of multihoming advertisers
    - each consumer is receptive for ad of 1 specific advertiser each period
    - ad shown to receptive consumer generates expected revenue ν
    - platform i chooses ad price $\nu * D_i$ (or per click price ν)

Discussion of model assumptions II

- alternating moves
  - no pure strategy equilibria exist with simultaneous investment
  - alternating move duopoly not unknown (Cyert and de Groot 70, Maskin and Tirole 88)
- time horizon
  - finite T: managers with fixed term contracts or retirement
  - main interest: $T \to \infty$
    - equilibrium selection in infinitely repeated game

Data driven market

Definition (Data driven market)

A *data-driven market* is a market characterized by indirect network effects driven by machine-generated data about user preferences or characteristics, such that the *marginal costs of innovating are decreasing in demand*: $c_{x,D_i} < 0$.

Period T

- say T even, then

$$\max_{x\geq 0} D_2(\Delta_{T-1}-x)-\gamma x^2/2-\alpha x(1-D_2(\Delta_{T-1}))$$

- solution

$$x^T=\begin{cases}\frac{1}{2\gamma}-\frac{\alpha}{\gamma}(1-D_2(\Delta_{T-1})) & \text{if } \Delta_{T-1}\in\left[-\frac{2\gamma-1+\alpha}{2\gamma+\alpha},U_\alpha\right]\\ 1+\Delta_{T-1} & \text{if } -1<\Delta_{T-1}<-\frac{2\gamma-1+\alpha}{2\gamma+\alpha}\\ 0 & \text{else}\end{cases} \tag{1}$$

where

$$U_\alpha=\begin{cases}1/\alpha-1 & \text{if } \alpha\geq 1/2\\ 1+(1-4\alpha(1-\alpha))/(4\gamma) & \text{if } \alpha<1/2.\end{cases}$$

Period $t < T$ I

- similar maximization problem as in T

$$\max_{x \geq 0} D_1(\Delta_{t-1}+x) - \gamma x^2/2 - \alpha x(1-D_1(\Delta_{t-1})) + \delta V_1^{t+1}(\Delta_{t-1}+x)$$

- V_i^t is piecewise (linear-) quadratic
- standard monotonicity results
  - V_1^t increases in Δ_t
  - V_2^t decreases in Δ_t
  - higher Δ_t leads to higher Δ_{t+1}

Period $t < T$ II

- I^t be set of Δ_{t-1} such that no firm takes over complete market in any following period

Lemma

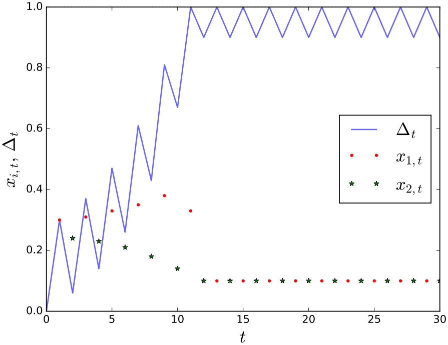
I^t is an interval.

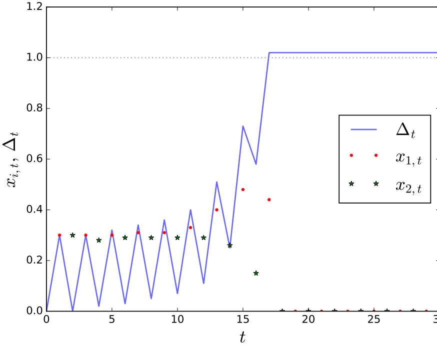
Theorem (Market tipping for $T \to \infty$)

The length of I^t is less than $\frac{2}{(1+\alpha/(2\gamma))^{\lceil T-t \rceil/2}}$. Consequently, the length of I^0 shrinks to zero at exponential speed for $T \to \infty$.

- If the time horizon is long, the market tips for (almost) any initial quality difference.

Period $t < T$ III





Extension: Entry into a traditional market I

- traditional incumbent with $c(x) = \gamma x^2/2 + \alpha x/2$
- entrant with data driven business model can enter at fixed cost F at quality level Δ_0

Corollary (Entry and tipping in a traditional market)

Let $T \to \infty$. For F very high, say $F > \bar{F}$, firm 1 does not enter. If F is below $\bar{F}$ but relatively high, then, whenever firm 1 enters, the market tips in favor of firm 1.

Extension: Entry into a traditional market II

Lemma (early indicators)

Take a stationary equilibrium of the game with infinite time horizon that is the limit of subgame-perfect equilibria in the T times repeated game, for $T \to \infty$.
If firm 1 enters and $\Delta_2 > \Delta_0$, then the market will eventually tip in favor of firm 1.

Discussion: Connected markets and Domino effect

Definition (Connected markets)

Markets A and B are *connected* for firm 1 if $c_{x_{1,A},D_{1,B}} < 0$ or $c_{x_{1,B},D_{1,A}} < 0$.

- identify connected markets and tip both
  - data from B helps to tip A (and/or vice versa)
- tip A, then B ("Domino effect")
  - data from A —> reduce MC in B sufficiently to profitably enter and tip

Theory of harm

- monopolization $\rightarrow$ low innovation rates
- pricing (?)
- lack of data as barrier to entry
- so far only minor antitrust cases against Google/Alphabet
  - lack of a theory of harm (?)
  - lack of understanding of market dynamics (?)
  - competition policy not applicable (?)
    - what is "abuse"?
    - can user data be viewed as essential facility?
- is this like *natural monopoly*?

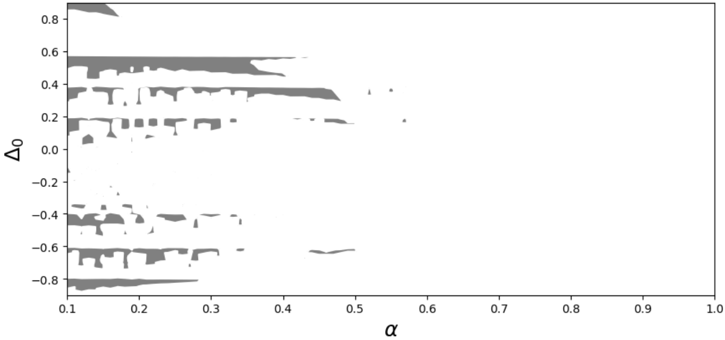
Regulation: Data sharing I

- suppose all usage data had to be shared among competing firms
- cost function as if full demand last period $c(x) = \gamma x^2/2$
- we show: length of I^t does not necessarily go to zero as $T \to \infty$

Regulation: Data sharing II

welfare effects:

- lower cost
- higher quality (low mc + no tipping)
- more consumers buy from their preferred firm (no tipping)
- higher investment



Data sharing (compared to *no data sharing*) reduces welfare in shaded regions and improves welfare in white regions. Parameters: $T = 30$, γ=2, δ=0.3, α-grid: 0.1,0.11,...,1.0, Δ_0-grid: -0.9,-0.91,...,0.9.

Illustrating cases

- internet search
- maps
- music streaming
    - Pandora's "music genome" vs. last.fm
- self-driving cars

Conclusion

- higher market share $\rightarrow$ more usage data $\rightarrow$ easier innovation $\rightarrow$ higher market share...
- data driven markets tend to tip (natural monopoly)
- monopolization can spill over to connected markets (Domino effect)
- potential long term harm to consumers possible
- data sharing?, other regulatory measures?

Period T

$$V_1^T(\Delta_{T-1}) = \begin{cases} \frac{2\gamma+\alpha-1+(2\gamma+\alpha)\Delta_{T-1}}{4\gamma} & \text{if } \Delta_{T-1} \in \left[-\frac{2\gamma-1+\alpha}{2\gamma+\alpha}, U_\alpha\right] \\ 0 & \text{if } \Delta_{T-1} < -\frac{2\gamma-1+\alpha}{2\gamma+\alpha} \\ D_1(\Delta_{T-1}) & \text{else} \end{cases}$$

$$V_2^T(\Delta_{T-1}) = \begin{cases} \frac{4\gamma+1-2\alpha+\alpha^2}{8\gamma} - \frac{2\gamma+\alpha-\alpha^2}{4\gamma}\Delta_{T-1} + \frac{\alpha^2}{8\gamma}\Delta_{T-1}^2 & \text{if } \Delta_{T-1} \in \left[-\frac{2\gamma-1+\alpha}{2\gamma+\alpha}, U_\alpha\right] \\ \frac{2-\alpha-\gamma}{2} - (\alpha+\gamma)\Delta_{T-1} - \frac{\alpha+\gamma}{2}\Delta_{T-1}^2 & \text{if } -1 < \Delta_{T-1} < -\frac{2\gamma-1+\alpha}{2\gamma+\alpha} \\ D_2(\Delta_{T-1}) & \text{else.} \end{cases}$$

- profits piecewise linear/quadratic in Δ_{t-1}

Period $t < T$ II

Lemma (Persistent dominance: finite)

If firm i has full demand in period $t < T - 1$, then firm i will have again full demand in a later period and firm j will not have full demand in any following period.

Lemma (Persistent dominance: infinite)

Take a stationary equilibrium that is limit of finite time horizon equilibria as $T \to \infty$.
If firm i has full demand in period t but not in period $t - 2$, then firm i will have full demand in all periods $t + 2n$ for $n \in \mathbb{N}$. Furthermore, firm j will have less demand in all consecutive periods than in $t - 1$.

Robustness: Infinite time horizon I

- stationary Markov equilibria: strategy depends only on Δ

Proposition (stable steady states)

Quality differences 1 and -1 are stable steady states.

- if market is sufficiently lopsided, it will tip completely

Proposition (tipping)

For every $\varepsilon > 0$, there exists a $\bar{\delta} > 0$ such that the market tips for all initial quality levels apart from a set of measure less than ε if $\delta < \bar{\delta}$.

- market tips if firms are sufficiently impatient

Robustness: Infinite time horizon II

- slight change: consider finite grid of qualities
- numerically solving for all Markov equilibria
- there are many Markov equilibria

# grid points	2	3	4	5	6	7	8	9	10
# eq.	2	4	8	16	38	96	113	113	113

- 2 selections
    - steady state selection: choose $x_1 = x_2 = 0$ wherever possible
    - invest selection: choose $x_1 > 0$, $x_2 > 0$ wherever possible
- numerical result: If data driven indirect network effects (α) are strong enough, the market tips.

Robustness: Quality decay

- assume $q_{i,t} = \mu q_{i,t-1} + x_{i,t}$ with $\mu \in (0, 1]$
- quality decay affects high quality firm more
- our results hold if μ not too low
  Proposition 1 holds if $\mu(\mu + \alpha/(2\gamma)) > 1$