Competing with Big Data

Jens Prüfer (Tilburg University)

Based on work with Christoph Schottmüller (Copenhagen)

Competition policy in the digital economy
CEPS seminar, Brussels, June 2016

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**. [...]

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**.

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 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."

Research questions

- Why is a firm like Google so successful ("world's largest listed company by market value"), whereas its competitors have more and more troubles?
- What is the role of "data" in this success story ("the company learns more about them")?
- How to explain the apparent oddity in Alphabet's strategy: being most successful, on the one side, while turning a lean and focused company into a conglomerate, on the other side?
- Which industries may be next on Alphabet's entry list?

Research questions

- Why is a firm like Google so successful ("world's largest listed company by market value"), whereas its competitors have more and more troubles?
- What is the role of "data" in this success story ("the company learns more about them")?
- How to explain the apparent oddity in Alphabet's strategy: being most successful, on the one side, while turning a lean and focused company into a conglomerate, on the other side?
- Which industries may be next on Alphabet's entry list?
- $\Rightarrow$ These questions have wide-ranging implications for many industries (suppliers, buyers, competition policy and regulation)!

This paper

- Defines "data-driven markets" and constructs a dynamic model of R&D competition in such markets.
- Shows that such markets tip under very mild conditions, moving towards monopoly.
- Shows how a dominant firm in one market can leverage its position to another data-driven market, thereby initiating a *domino effect*.
- Applies the model to Google's case and tries to explain the logic behind their strategy (and success).
- Analyzes a regulatory measure how to mitigate the negative effects of a fundamental characteristic of data-driven markets for innovation.

This paper

- Defines "data-driven markets" and constructs a dynamic model of R&D competition in such markets.
- Shows that such markets tip under very mild conditions, moving towards monopoly.
- Shows how a dominant firm in one market can leverage its position to another data-driven market, thereby initiating a *domino effect*.
- Applies the model to Google's case and tries to explain the logic behind their strategy (and success).
- Analyzes a regulatory measure how to mitigate the negative effects of a fundamental characteristic of data-driven markets for innovation.
- **Today**: overview of results + problem of delineation of data-driven markets (via "market connectedness") + theory of harm (via indirect network effects)

Important distinctions: three economic mechanisms

1. *Direct network effects*: consumption utility of one consumer increases in amount of other consumers on the same network: $u_j'(D_i) > 0$
   - E.g. telecoms $\Rightarrow$ demand-driven, supplier's cost unaffected

Important distinctions: three economic mechanisms

1. *Direct network effects*: consumption utility of one consumer increases in amount of other consumers on the same network: $u_j'(D_i) > 0$
   - E.g. telecoms $\Rightarrow$ demand-driven, supplier's cost unaffected
2. *Dynamic economies of scale* (learning-curve effects): Marginal cost of manufacturing decreases in the number of goods produced: $c_{x_i,s_i} < 0$
   - E.g. airplanes $\Rightarrow$ supply-driven, no users/demand involved

Important distinctions: three economic mechanisms

1. *Direct network effects*: consumption utility of one consumer increases in amount of other consumers on the same network: $u_j'(D_i) > 0$
   - E.g. telecoms $\Rightarrow$ demand-driven, supplier's cost unaffected
2. *Dynamic economies of scale* (learning-curve effects): Marginal cost of manufacturing decreases in the number of goods produced: $c_{x_i,s_i} < 0$
   - E.g. airplanes $\Rightarrow$ supply-driven, no users/demand involved
3. *Indirect network effects*: supplier's marginal cost of innovation decreases in amount of user information, which increases with demand: $c_{x_i,D_i} < 0$
   - Data-driven markets $\Rightarrow$ supply and demand interact
   - $\Rightarrow$ not another *Microsoft* case (where next innovation offsets incumbent's strength)!

Data-driven markets

Definition 1

A *data-driven market* is a market characterized by indirect network effects driven by machine-generated data about user preferences or characteristics, s.t. the instantaneous costs of innovating, $c(x_i, D_i)$, are decreasing in demand: $c_{x_i,D_i} < 0$.

- Examples: search engines, online social networks, platform sites for dating, accomodation, car sharing, etc.
  - More examples to come below...
- Henceforth: "machine-generated data about user preferences or characteristics" $\equiv$ *user information*

Helpful definition

Definition 2

(Market Tipping): A market is *tipping* if, in the long run, one firm emerges with full demand and the other firm with zero demand.

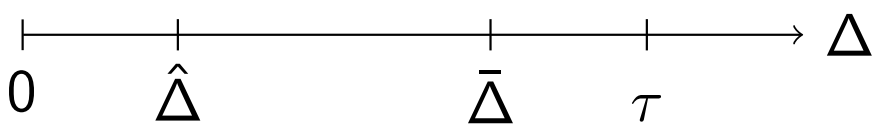
- Important: this is technical definition for model.
- Reality is more nuanced: *tipping* if the largest firm (by market share) gets larger, smaller firms get smaller (and some exit)

Main result

Proposition 1

If initial quality difference between competitors, Δ, sufficiently high, the weaker firm will eventually exit the market (market tipping).

- Details:
  - If $\Delta > \tau$, firm 1 has full demand and firm 2 is effectively out.
  - If $\Delta > \bar{\Delta}$, firm 2 stops innovating, i.e. $x_2(\Delta) = 0$.
  - If $\Delta > \hat{\Delta}$, firm 1 will eventually have full demand.



Tipping a traditional market

- Consider a mature "traditional" market, where several firms compete with (relatively) homogenous products (e.g. geographical maps in 1990s).

Tipping a traditional market

- Consider a mature "traditional" market, where several firms compete with (relatively) homogenous products (e.g. geographical maps in 1990s).

Proposition 2

If the market entry cost are not prohibitive, a firm that manages to find a "data-driven" business model can tip/dominate any market in the long term.

- Because of *first-mover advantage* in data-driven markets, the first "data-driven firm" (=with a data-driven business model) to enter is likely to dominate the industry.
  - Any second-mover copying the business model would face serious problem.

The Domino Effect

Definition 3

Markets A and B are *connected* if, for all $i \in \{1, 2\}$, $c_{x_{i,A}, D_{i,B}} < 0$ or $c_{x_{i,B}, D_{i,A}} < 0$.

- *Domino effect*: If firm 1 is dominant in data-driven market A and it can identify a *connected* market B, where entry is not prohibitive ($F \leq \bar{F}(\tilde{q})$), market B will also tip.
  - $\Rightarrow$ Firm 1 can *leverage* its dominant position from market A to market B.

Implications

- What could firm 2 do against adverse tipping?

Implications

- What could firm 2 do against adverse tipping?
  Firm 2 can invest in (non-data-related) quality *before* firm 1 entered market B $\Rightarrow \bar{F}(\tilde{q}\uparrow)\downarrow\Rightarrow$ profitable entry of firm 1 less likely
- But once firm 1 has entered a market (and found a data-driven business model), tipping seems inevitable

Implications

- What could firm 2 do against adverse tipping?
  Firm 2 can invest in (non-data-related) quality *before* firm 1 entered market B $\Rightarrow \bar{F}(\tilde{q}\uparrow)\downarrow\Rightarrow$ profitable entry of firm 1 less likely
- But once firm 1 has entered a market (and found a data-driven business model), tipping seems inevitable
- *Prediction*: race of *data companies* to identify data-driven business models utilizing their existing data stocks, and *traditional companies* trying to increase data-independent product quality (or to develop large enough stock of user data to deter market entry)

Robustness

Innovation in the long(er) run

- Profits follow demand: **dominant firm makes huge profits**. **Profits of competitors decline**.

Innovation in the long(er) run

- Profits follow demand: **dominant firm makes huge profits**. **Profits of competitors decline**.
- A tipped market provides *no incentives for the dominant firm to innovate further*.
- $\Rightarrow$ tipped market **negative for consumers**.

Innovation in the long(er) run

- Profits follow demand: **dominant firm makes huge profits**. **Profits of competitors decline**.
- A tipped market provides *no incentives for the dominant firm to innovate further*.
- $\Rightarrow$ tipped market **negative for consumers**.
- Because $x = 0 < x_{efficient}$, (where $x = c_x$), tipped market also inefficient
  - Edelman (2016): "Google dulls the incentive to enter affected sectors. Leaders of TripAdvisor and Yelp, among others, report that they would not have started their companies had Google engaged in behaviors that later became commonplace [= tying a new service to its search engine/leveraging data-leadership in SEs to new sectors]."

Policy Proposal: regulation, not competition policy!

- *Theory of harm*: Even without any abusive behavior, data-driven market will tip $\Rightarrow$ dominant firm (and everybody else) will have very low incentives to innovate.

Policy Proposal: regulation, not competition policy!

- *Theory of harm*: Even without any abusive behavior, data-driven market will tip $\Rightarrow$ dominant firm (and everybody else) will have very low incentives to innovate.
- Argenton/Prüfer (2012) propose that search engines should be *required to share search log data amongst each other*.
- Here: analyze effects of this proposal extended to all data-driven markets.
- **What if firms with data-driven business models are legally required to share their (anonymized) data about user preferences?**

Policy Proposal: regulation, not competition policy!

- *Theory of harm*: Even without any abusive behavior, data-driven market will tip $\Rightarrow$ dominant firm (and everybody else) will have very low incentives to innovate.
- Argenton/Prüfer (2012) propose that search engines should be *required to share search log data amongst each other*.
- Here: analyze effects of this proposal extended to all data-driven markets.
- **What if firms with data-driven business models are legally required to share their (anonymized) data about user preferences?**
- Question how legal and how technically feasible such sharing is deferred to A/P(2012) and future research.

A solution to market tipping?!

Proposition 3

Data sharing can avoid market tipping.

- Main insight: Previous results on tipping do not hold if data shared. Both firms keep investing in innovation.
  - NB: in reality, perfect data sharing may be impossible $\Rightarrow$ market tipping still occurs but slower
  - Other caveats apply, e.g. possibility that advertisers of large platform willing to pay more at the margin ($p_1 > p_2$ if $D_1 > D_2$)

Political support

Conclusion

- In *data-driven markets*, costs of innovation decrease in amount of *data on user preferences or characteristics*.

Conclusion

- In *data-driven markets*, costs of innovation decrease in amount of *data on user preferences or characteristics*.
- If these data property of firm collecting the data (as byproduct of serving consumers), market tips under very mild conditions.

Conclusion

- In *data-driven markets*, costs of innovation decrease in amount of *data on user preferences or characteristics*.
- If these data property of firm collecting the data (as byproduct of serving consumers), market tips under very mild conditions.
- Dominant position in one data-driven market can be leveraged into other market if markets are *connected*.

Conclusion

- In *data-driven markets*, costs of innovation decrease in amount of *data on user preferences or characteristics*.
- If these data property of firm collecting the data (as byproduct of serving consumers), market tips under very mild conditions.
- Dominant position in one data-driven market can be leveraged into other market if markets are *connected*.
- Google provides blueprint how data-driven dominance in one market can lead to dominance in other markets (even without abuse of dominant position!):
  - Search engines $\Rightarrow$ maps $\Rightarrow$ self-driving cars

Conclusion

- In *data-driven markets*, costs of innovation decrease in amount of *data on user preferences or characteristics*.
- If these data property of firm collecting the data (as byproduct of serving consumers), market tips under very mild conditions.
- Dominant position in one data-driven market can be leveraged into other market if markets are *connected*.
- Google provides blueprint how data-driven dominance in one market can lead to dominance in other markets (even without abuse of dominant position!):
  - Search engines $\Rightarrow$ maps $\Rightarrow$ self-driving cars
- Market tipping is negative for innovation and, hence, consumers.

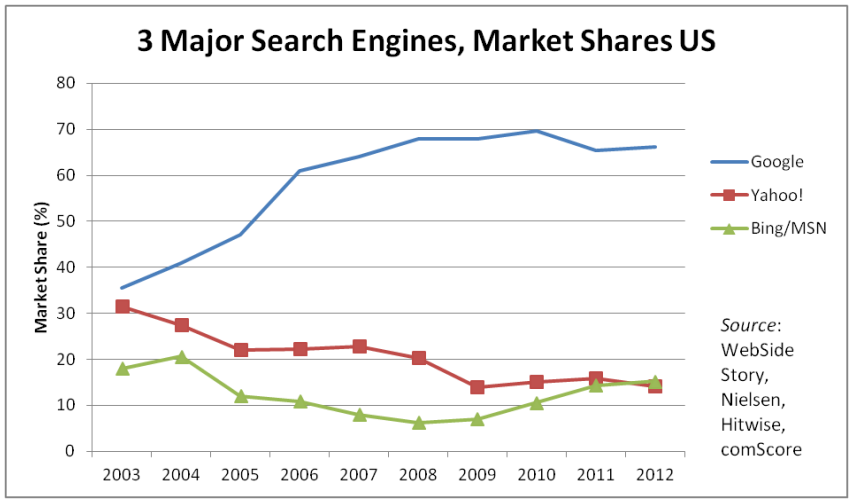
Conclusion

- In *data-driven markets*, costs of innovation decrease in amount of *data on user preferences or characteristics*.
- If these data property of firm collecting the data (as byproduct of serving consumers), market tips under very mild conditions.
- Dominant position in one data-driven market can be leveraged into other market if markets are *connected*.
- Google provides blueprint how data-driven dominance in one market can lead to dominance in other markets (even without abuse of dominant position!):
  - Search engines $\Rightarrow$ maps $\Rightarrow$ self-driving cars
- Market tipping is negative for innovation and, hence, consumers.
- Policy proposal: mandatory sharing of data on user preferences can mitigate market tipping.

Thank you very much for your attention!

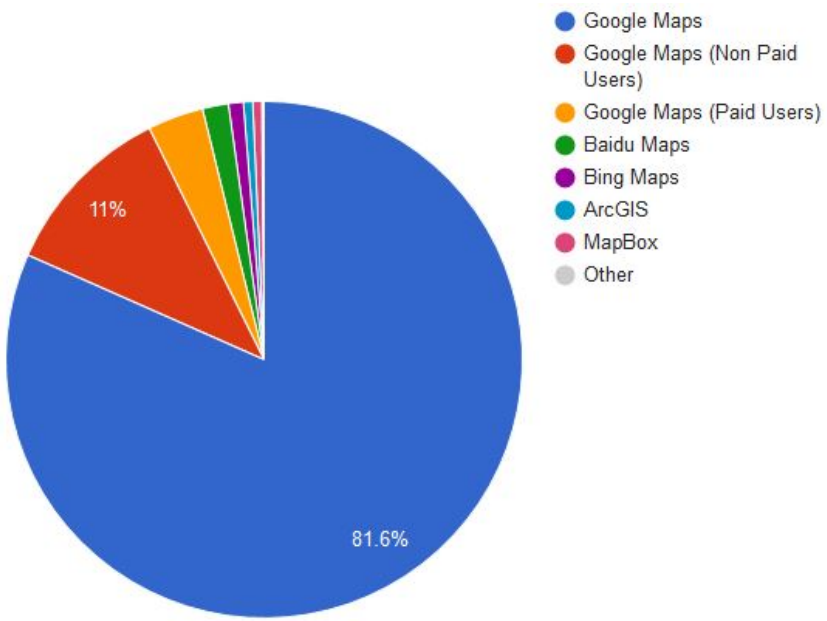
Applying the model to a real-world case: Google

- Argenton/Prüfer (2012) document:
  1. market tipping in search engine market since Google became market leader in 2003
  2. strong profit growth of Google / increasing losses of competitors (Microsoft, Yahoo)
  3. market exit of one follower (Yahoo selling its search/ad business to MS in 2010)



From search engines to maps (backward-looking)

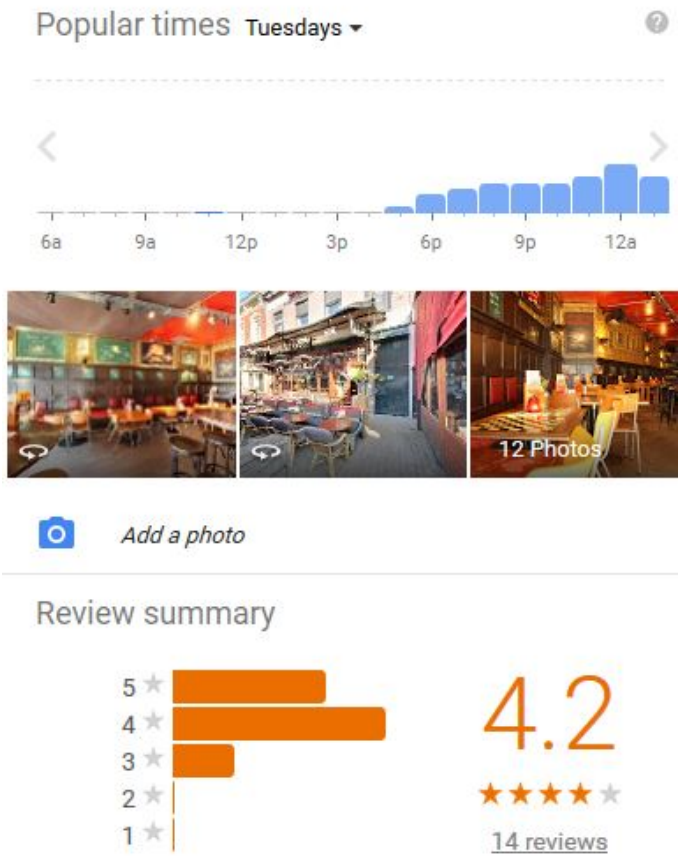
- *Connection*: Many queries in search engine are geography-related ⇒ SE's get "free" stock of users' geo-preferences
- *Investing F* : buying/producing geographical data
- *Investing x* : adding features



$\frac{\#\text{ Websites using given technology}}{\text{Total }\#\text{ websites using any of the technologies}}$, 05/2016, Source: `https://www.datanyze.com/market-share/maps/`

Machine and user-generated data in GoogleMaps

- Screenshot Cafe Stoffel, Tilburg, @GoogleMaps (05/17/2016)
- Features based on user-preferences (machine-generated and user-generated)
- Zuboff (2016): "Google recently announced that its maps will not only provide the route you search but will also suggest a destination."



From maps to self-driving cars (forward-looking)

- *Connection:* Geodata & data on user characteristics (e.g. where are traffic jams) important inputs into SD cars
- "The car **processes both map and sensor information** to determine where it is in the world. [...] The **software predicts** what all the objects around us might do next. It predicts that the cyclist will ride by and the pedestrian will cross the street."

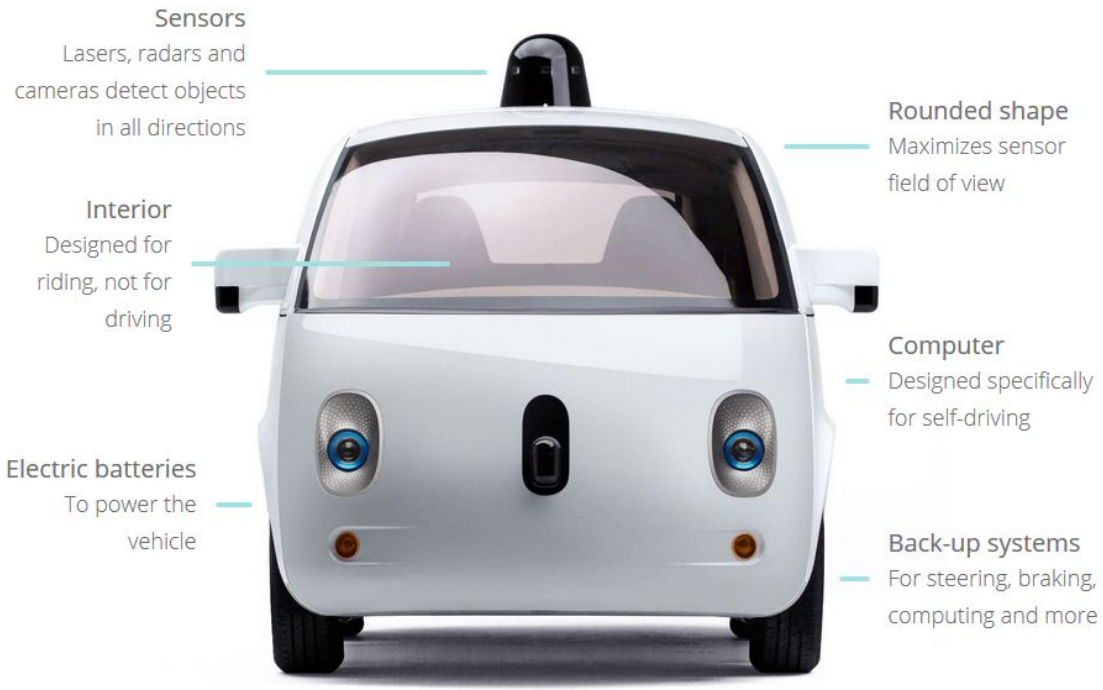
  (`https://www.google.com/selfdrivingcar/how/`, 2016/05/17)

From maps to self-driving cars (forward-looking)

- *Connection:* Geodata & data on user characteristics (e.g. where are traffic jams) important inputs into SD cars
- "The car **processes both map and sensor information** to determine where it is in the world. [...] The **software predicts** what all the objects around us might do next. It predicts that the cyclist will ride by and the pedestrian will cross the street."

  (`https://www.google.com/selfdrivingcar/how/`, 2016/05/17)
- *Investing F* : engineering self-driving car with public approval
- *Investing x* : adding features utilizing data by *other* cars
  - "A high-end car [...] has the digital horsepower of 20 personal computers and generates 25 gigabytes of data per hour of driving [says consultancy Accenture]" (Economist, 2015/11/21)

Google's self-driving car



Source: `https://www.google.com/selfdrivingcar/how/`, 2016/05/17.

- "US vehicle safety regulators have said the artificial intelligence system piloting a self-driving Google car could be considered the driver under federal law" (NY Times, 2016/02/16).

Robustness

- In the paper, we provide extensions and robustness checks:
  - *Two-sided markets*: endogenize prices of platforms to advertisers (who are distributed uniformly on Hotelling line)
  - *Quality decay*: $\dot{q}_i = x_i - \delta q_i \Rightarrow \delta \geq 0$ is rate of quality decay
  - Entry into traditional market, where *incumbent* has market power/*can set price strategically* (not $p = \tilde{c}$)
  - *Competition* among several data-driven firms entering traditional market
  - Case where traditional incumbent has captured a share of market (via selling *durable products*)
- $\Rightarrow$ **Model results qualitatively robust, market still tipping**

Back

We are not alone ...

- Zuboff (2016): "We need new interventions that interrupt, outlaw, or regulate
  1. the initial capture of behavioral surplus,
  2. the use of behavioral surplus as free raw material,
  3. excessive and exclusive concentrations of the new means of production,
  4. the manufacture of prediction products,
  5. the sale of prediction products,
  6. the use of prediction products for third-order operations of modification, influence, and control, and
  7. the monetization of the results of these operations."

Back