The future is private?

DIGITAL PRIVACY

Itay P. Fainmesser (Johns Hopkins University)
Andrea Galeotti (London Business School)
Ruslan Momot (HEC Paris)
Better products and services:
• better matching
• better ads and targeting
• safety
• ...

Risk of data being misused
• identity theft
• manipulating public opinion
• government overreach
• ...

User Data
4. Data, Privacy, and Security

How should we perceive and deal with digital data?
To what extent do we understand privacy as a tradeoff?
Are private sector incentives sufficient to provide adequate data security?
4. Data, Privacy, and Security

How should we perceive and deal with digital data?
To what extent do we understand privacy as a tradeoff?
Are private sector incentives sufficient to provide adequate data security?

“The problem with data protection laws is that it presumes the data collection was ok”
1. Do digital businesses have the right incentive to collect and protect users’ information?

   How do the incentives to collect and protect user data, and their welfare implications depend on the platform’s revenue model?

2. What is the scope for regulation of data collection and protection?

3. How much does data abuse actually hurts consumers?
Differential Privacy — design of algorithmic mechanisms for anonymizing individual-level data

Dwork and Roth 2014; Cummings et al. 2015; Ghosh and Roth 2015; Abowd and Schmutte 2019.

How users’ information is revealed through users’ actions and how firms acquire that data.


How revealed information can be used to design personalized pricing, targeted advertisement, shape norms and, in turn affects individuals’ behavior.


The Data Economy

Farboodi and Veldkamp 2019…
**MODEL OVERVIEW**

*Users*

How active to be?

*Business*

How much info to collect?
(e.g. encryption, transcription)

How much to invest in protection?
(firewall, API safeguards, ethical hackers)

*Adversaries*

To enter or not?
Actions and Timeline

Data Policy
- How much info to collect? $\xi \in [0,1]$ 
- How much to invest in protection? $C \geq 0$

How active to be? $a_i \geq 0$

To enter or not?

Business

Users

Adversary

Time
A continuum of mass 1 of users

\[ U_i(a_i) = a_i b_i - \frac{1}{2} a_i^2 + \beta a_i \bar{a} + \rho \cdot a_i \xi - \omega_i \cdot a_i \xi \]

\( \omega_i \) — expected number of adversarial against user \( i \)

\( a_i \xi \) — data collected on user \( i \)
A continuum of mass 1 of users

Direct benefit from usage (e.g., messaging, photos, etc.)

\[ U_i(a_i) = a_i b_i - \frac{1}{2} a_i^2 + \beta a_i \bar{a} + \rho \cdot a_i \xi - \omega_i \cdot a_i \xi \]

Positive network effects

Benefit from the business using user’s information (e.g., advertisements, better matches)

Expected loss from adversarial activity

\( \omega_i \) — expected number of adversarial against user \( i \)

\( a_i \xi \) — data collected on user \( i \)
Adversaries are heterogeneous in their ability to hack security systems $\gamma_j \sim U[0, K]$

If adversary $j$ chooses to enter, he pays a fixed cost $\gamma_j C$ and takes action against one user $i$ chosen u.a.r.

$\pi_{j \rightarrow i} = a_i \xi - \gamma_j C$

Benefit to adversary $j$ from action against user $i$  \hspace{1cm} Cost of hacking the system

$\Rightarrow$ Adversary participates if: $\gamma_j \leq \frac{a \xi}{C}$

(where $\bar{a} = average\{a_i\}_i$)
\[ \Pi = \Phi(\overline{a}, \overline{a}\xi) - \psi C \]

We assume that \( \Phi \):
1. Increases in both elements
2. Concave in both elements

Transaction-driven business \( \Phi'_y = 0 \)

Advertisement-driven business \( \Phi'_x = 0 \)
TIMELINE

Business

Data Policy
How much info to collect? $\xi \in [0,1]$ How much to invest in protection? $C \geq 0$

Users
How active to be? $a_i \geq 0$

Adversary
To enter or not?

Time
**EQUILIBRIUM**

**Proposition:** There exists a unique equilibrium. In equilibrium, average user activity is $\overline{a^*} = \frac{c(b + \rho \xi)}{c(1-\beta)+\xi^2}$ and $\omega_i^* = \frac{\overline{a^*}}{c}$ adversaries enter.

**Observation:** Consumer surplus is proportional to $\overline{a^*^2}$

**Proof:**

$CS = \frac{1}{2} \int a_i^*^2 = \frac{1}{2} \left( \sigma_b^2 + \overline{a^*^2} \right)$ where $\sigma_b^2$ is exogenous.
**Timeline**

**Business**
- Data Policy
  - How much info to collect? $\xi \in [0,1]$  
  - How much to invest in protection? $C \geq 0$

**Users**
- How active to be? $a_i \geq 0$

**Adversary**
- To enter or not?
Let $\xi_W$ and $\xi_{CS}$ be the welfare and consumer surplus maximizing levels of $\xi$ respectively.

**Proposition:** Let $C$ be fixed.

1. A business collects more data than the level that maximizes welfare and consumer surplus. That is $\xi^* \geq \xi_W, \xi_{CS}$.

2. A transaction-driven business collects the efficient amount of data given the level of data protection. That is, $\xi^* = \xi_W, \xi_{CS}$.
DATA COLLECTION (EXOGENOUS C)

Let $\xi_W$ and $\xi_{CS}$ be the welfare and consumer surplus maximizing levels of $\xi$ respectively.

Proposition: Let $C$ be fixed.

1. A business collects more data than the level that maximizes welfare and consumer surplus. That is $\xi^* \geq \xi_W, \xi_{CS}$.

2. A transaction-driven business collects the efficient amount of data given the level of data protection. That is, $\xi^* = \xi_W, \xi_{CS}$.

“there’s the important question of whether the advertising model encourages companies like ours [Facebook] to use and store more information that we otherwise would."
Currently in the US

**Regulation:** minimum level of data protection \((C)\).

**Enforcement:**
2. Rare: minimum level of data protection \((C)\).
Currently in the US

Regulation: minimum level of data protection ($C$).

Enforcement:
2. Rare: minimum level of data protection ($C$).

Proposition: Let $C$ be fixed.

Welfare can be maximized by fines on data breaches, equal to

$$L = \frac{\partial CS/\partial \xi}{\partial (\omega \bar{a} \xi)/\partial \xi} \bigg|_{\xi_W}$$

Consumer surplus can be maximized by fines on data breaches, equal to

$$L = \frac{2}{\rho} \Phi'_y \bigg|_{\xi_{CS}}$$
## Data Collection and Protection

<table>
<thead>
<tr>
<th></th>
<th>Data collection ((\xi))</th>
<th>Data collection and protection ((C, \xi))</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>General</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Direction of inefficiency</td>
<td>(+)</td>
<td></td>
</tr>
<tr>
<td>Effective regulation</td>
<td>Liability</td>
<td></td>
</tr>
<tr>
<td><strong>Transaction-driven business</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Direction of inefficiency</td>
<td>No inefficiency</td>
<td></td>
</tr>
<tr>
<td>Effective regulation</td>
<td>N\A</td>
<td></td>
</tr>
</tbody>
</table>
## Data Collection and Protection

<table>
<thead>
<tr>
<th></th>
<th>Data collection ($\xi$)</th>
<th>Data collection and protection ($C, \xi$)</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>General</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Direction of inefficiency</td>
<td>(+)</td>
<td>$(-, -); (-, +); (+, +)$</td>
</tr>
<tr>
<td>Effective regulation</td>
<td>Liability</td>
<td>$\underline{C} + \text{liability}$</td>
</tr>
<tr>
<td><strong>Transaction-driven business</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Direction of inefficiency</td>
<td>No inefficiency</td>
<td>$(-, -)$</td>
</tr>
<tr>
<td>Effective regulation</td>
<td>N\A</td>
<td>$\underline{C}$</td>
</tr>
</tbody>
</table>

**Observations:**
1. In general, digital businesses may collect and protect too much or too little.
2. In general, a lower bound on protection is insufficient and could have perverse effects if used by itself.
3. FTC’s enforcement may be efficiency enhancing.*
**The Loss Multiplier of Adversarial Activity**

The change in consumer surplus due to the presence of adversaries:

\[ U_i(a_i) = a_i b_i - \frac{1}{2} a_i^2 + \beta a_i \bar{a} + \rho \cdot a_i \xi - \omega \cdot a_i \xi \]

**The direct loss ignores:**

- The change in activity when users are fearful of adversaries
- A respective change in businesses’ data policies

Taking into account all these effects:

\[ \text{Loss Multiplier (LM)} = \frac{CS_{\text{without adversaries}} - CS_{\text{with adversaries}}}{\text{Direct loss}} \geq \frac{2}{1 - \beta} \]

The change in consumer surplus due to the presence of adversaries:

- is at least 2 times direct loss;
- quickly explodes with strength of network effects $\beta$
1. Do digital businesses have the right incentive to collect and protect users’ information?
   How do the incentives to collect and protect user data, and their welfare implications depend on the platform’s revenue model?

2. What is the scope for the regulation of data collection and protection?

3. How much does data abuse actually hurt consumers?
**DISCUSSION**

**Data protections policies (e.g., GDPR)**
- In equilibrium data protection and data collection are complements

**Potential regulatory / firm actions**
- Splitting up Facebook
- Pay for data
- Restricting data collection (also GDPR)

**Other questions**
- Competition
- Digital equity and data policies
  - an increase in $\xi$ can lead to an increase in $a_i\xi$ for some (high $b_i$) users and a decrease for others.
THANK YOU!