Nonrivalry and the Economics of Data

Chad Jones and Chris Tonetti

FRBSF
23 August 2018
Examples of Data

- Google, Facebook
- Amazon
- Tesla, Uber, Waymo
- Medical and genetic data
- Location history
- Speech records
- Physical action data

Canonical example: data as input into machine learning algorithm. E.g. self-driving car.
What policies governing data use maximize welfare?

- European General Data Protection Regulation (GDPR)
  - Privacy vs. social gain from sharing
  - “The protection of natural persons in relation to the processing of personal data is a fundamental right”
  - “The right . . . must be considered in relation to its function in society and be balanced against other fundamental rights. . .”

- The California Consumer Privacy Act of 2018
  - Allows consumers to opt out of having their data sold
Data is Nonrival

- Growth literature: Ideas are nonrival
  - Unlike rival goods, ideas are not depleted by use

- Data is another nonrival good
  - Clearly not a blueprint / recipe ⇒ different from ideas
  - Ideas are production functions, data is a factor of production

- Multiple engineers/algorithms can use same data at same time (within and across firms)

- Data is nonrival ⇒ Important role for property rights/policies
In our economy, data ownership affects data use (Coase theorem fails)
  - Firms can’t commit to not using data it acquired from consumer
  - Consumers can’t commit to selling data to just one firm

Allocations differ when firms own data or consumers own data
Outline

• Economic environment

• Allocations:
  ○ Optimal allocation
  ○ Firms own data
  ○ Consumers own data
  ○ Extreme privacy protection: outlaw data sharing

• Theory results and a numerical example
The Economic Environment

Utility

\[ \int_0^\infty e^{-\rho t} L_t u(c_t, x_{it}, \tilde{x}_{it}) dt \]

Flow Utility

\[ u(c_t, x_{it}, \tilde{x}_{it}) = \log c_t - \frac{\kappa}{2} \frac{1}{N_t} \int_0^{N_t} x_{it}^2 di - \frac{\tilde{\kappa}}{2} \frac{1}{N_t} \int_0^{N_t} \tilde{x}_{it}^2 di \]

Consumption per person

\[ c_t = \left( \int_0^{N_t} c_{it}^{-\frac{\sigma}{\sigma-1}} di \right)^{\frac{\sigma}{\sigma-1}} \quad \text{with } \sigma > 1 \]

Data production

\[ J_{it} = c_{it} L_t \]

Variety resource constraint

\[ c_{it} = \frac{Y_{it}}{L_t} \]

Firm production

\[ Y_{it} = D_{it}^\eta L_{it}, \quad \eta \in (0, 1) \]

Data used by firm \( i \)

\[ D_{it} \leq \alpha x_{it} J_{it} + (1 - \alpha) B_t \quad \text{(nonrivalry)} \]

Data of firm \( i \) used by others

\[ D_{sit} \leq \tilde{x}_{it} J_{it} \]

Data bundle

\[ B_t = \left( N_t^{-\frac{1}{\epsilon}} \int_0^{N_t} D_{sit}^{\frac{\epsilon-1}{\epsilon}} di \right)^{\frac{\epsilon}{\epsilon-1}} \quad \text{with } \epsilon > 1 \]

Innovation (new varieties)

\[ \dot{N}_t = \frac{1}{\chi} \cdot L_{et} \]

Labor resource constraint

\[ L_{et} + \int_0^{N_t} L_{it} di = L_t \]

Population growth (exogenous)

\[ L_t = L_0 e^{g_L t} \]

Creative destruction

\[ \delta(\tilde{x}_{it}) = \frac{\delta_0}{2} \tilde{x}_{it}^2 \quad \text{(equilibrium)} \]
The Planner Problem (using symmetry of firms)

\[
\max_{\{L_{pt}, x_{it}, \tilde{x}_{it}\}} \int_0^\infty e^{-\tilde{\rho}t} L_0 u(c_t, x_{it}, \tilde{x}_{it}) \, dt, \quad \tilde{\rho} := \rho - g_L
\]

subject to

\[
c_t = Y_t / L_t
\]
\[
Y_t = N_t^{\frac{1}{\sigma - 1}} D_{it}^\eta L_{pt}
\]
\[
D_{it} = \alpha x_{it} Y_{it} + (1 - \alpha) N_t \tilde{x}_{it} Y_{it}
\]
\[
Y_{it} = D_{it}^\eta \cdot \frac{L_{pt}}{N_t}
\]
\[
\dot{N}_t = \frac{1}{\chi} (L_t - L_{pt})
\]
\[
L_t = L_0 e^{gL_t}
\]

- More sharing ⇒ negative utility cost but more consumption
- Balance labor across production and entry/innovation
Scale Effect from Sharing Data

\[ D_{it} = \alpha x_{it} J_{it} + (1 - \alpha) \left( N_t^{-\frac{1}{\epsilon}} \int_0^{N_t} (\tilde{x}_{it} J_{it})^\frac{\epsilon-1}{\epsilon} \, di \right)^\frac{\epsilon}{\epsilon-1} \]

\[ D_{it} = \alpha x_{it} Y_{it} + (1 - \alpha) N_t \tilde{x}_{it} Y_{it} \]

\[ = [\alpha x_{it} + (1 - \alpha) \tilde{x}_{it} N_t] Y_{it} \]

- No sharing versus sharing:
  - **No sharing**: Only the \( \alpha x_t \) term = no scale effect
  - **Sharing**: The \( (1 - \alpha) \tilde{x}_t N_t \) term = extra scale effect

Source of Scale Effect: \( N_t \) scales with \( L_t \)

- Plugging into production function:

\[ Y_{it} = \left( [\alpha x_t + (1 - \alpha) \tilde{x}_t N_t]^\eta L_{it} \right)^{\frac{1}{1-\eta}} \]
The Optimal Allocation on BGP (asymptotic)

\[ \tilde{x}_{it} = \tilde{x}_{sp} = \left( \frac{1}{\tilde{\kappa}} \cdot \frac{\eta}{1 - \eta} \right)^{1/2} \]  
(1)

\[ x_{it} = x_{sp} = \frac{\alpha}{1 - \alpha} \frac{\tilde{\kappa}}{\kappa} \left( \frac{1}{\tilde{\kappa}} \cdot \frac{\eta}{1 - \eta} \right)^{1/2} \]  
(2)

\[ L_{it}^{sp} = \chi \rho \cdot \frac{\sigma - 1}{1 - \eta} := \nu_{sp} \]  
(3)

\[ N_{t}^{sp} = \frac{L_{t}}{\chi (g_{L} + \nu_{sp})} := \psi_{sp} L_{t} \]  
(4)

\[ L_{pt}^{sp} = \nu_{sp} \psi_{sp} L_{t} \]  
(5)

\[ Y_{t}^{sp} = (\nu_{sp} (1 - \alpha)^{\eta} \tilde{x}_{sp}^{\eta}) \frac{1}{1 - \eta} (\nu_{sp} L_{t}) \frac{1}{\sigma - 1} + \frac{1}{1 - \eta} \]  
(6)

\[ c_{t}^{sp} = \frac{Y_{t}}{L_{t}} = (\nu_{sp} (1 - \alpha)^{\eta} \tilde{x}_{sp}^{\eta}) \frac{1}{1 - \eta} (\nu_{sp} L_{t}) \frac{1}{\sigma - 1} + \frac{\eta}{1 - \eta} \]  
(7)

\[ g_{c}^{sp} = \left( \frac{1}{\sigma - 1} + \frac{\eta}{1 - \eta} \right) g_{L} \]  
(8)

\[ D_{i}^{sp} = ((1 - \alpha)\tilde{x}_{sp} \nu_{sp} \psi_{sp} L_{t}) \frac{1}{1 - \eta} \]  
(9)

\[ D^{sp} = N D_{i} = ((1 - \alpha)\tilde{x}_{sp} \nu_{sp}) \frac{1}{1 - \eta} (\nu_{sp} L_{t}) \frac{1 + \frac{1}{1 - \eta}}{1 - \eta} \]  
(10)

\[ Y_{it}^{sp} = (\nu_{sp} (1 - \alpha)^{\eta} \tilde{x}_{sp}^{\eta}) \frac{1}{1 - \eta} (\nu_{sp} L_{t}) \frac{\eta}{1 - \eta} \]  
(11)
The Optimal Allocation: Data, Firm Size, Variety

\[ \tilde{x}_{sp} = \left( \frac{1}{\tilde{\kappa}} \cdot \frac{\eta}{1 - \eta} \right)^{1/2} \]

\[ L_{it}^{sp} = \chi \rho \cdot \frac{\sigma - 1}{1 - \eta} := \nu_{sp} \]

\[ N_{t}^{sp} = \frac{L_t}{\chi \delta L + \nu_{sp}} := \psi_{sp} L_t \]

- Data shared increasing in data production elasticity and decreasing in privacy cost.
- Firm size constant on BGP. \( N \) has opposite comparative statics.
- Higher entry cost, time preference, population growth, and elasticity of substitution raise firm size and reduce varieties.
- Higher \( \eta \) raises firm size and reduces varieties. Entry does not create data.
Firms Own Data
Firms Own Data: Consumer Problem

- Firms own data, so consumers just choose consumption:

\[ U_0 = \max_{\{c_{it}\}} \int_0^\infty e^{-\tilde{\rho}t} L_0 u(c_t, x_{it}, \tilde{x}_{it}) dt \]

s.t. \[ c_t = \left( \int_0^{N_t} c_{it}^{\frac{\sigma-1}{\sigma}} di \right)^{\frac{\sigma}{\sigma-1}} \]

\[ \dot{a}_t = (r_t - g_L)a_t + \omega_t - \int_0^{N_t} p_{it} c_{it} di \]
Firms own the Data: Incumbent Firm Problem

- Monopolistically competitive firm takes demand for variety as given (from FOC of consumer problem):  
  \[ p_{it} = \left( \frac{c_t}{c_{it}} \right)^{\frac{1}{\sigma}} = \left( \frac{Y_t}{Y_{it}} \right)^{\frac{1}{\sigma}} \]

\[ r_t V_{it} = \max_{L_{it}, D_{bit}, x_{it}, \tilde{x}_{it}, \bar{x}_{it}} \left( \frac{Y_t}{Y_{it}} \right)^{\frac{1}{\sigma}} Y_{it} - w_t L_{it} - p_{bt} D_{bit} + p_{st} \bar{x}_{it} Y_{it} + \dot{V}_{it} - \delta(\tilde{x}_{it}) V_{it} \]

s.t.  
\[ Y_{it} = D_{it}^{\eta} L_{it} \]

\[ D_{it} = \alpha x_{it} Y_{it} + (1 - \alpha) D_{bit} \]

\[ x_{it} \in [0, 1], \tilde{x}_{it} \in [0, 1] \]

\[ p_{sit} = \lambda DI N_t^{-\frac{1}{\epsilon}} \left( \frac{B_t}{\tilde{x}_{it} Y_{it}} \right)^{\frac{1}{\epsilon}} \]

- Data Intermediary \((p_{bt}, p_{st}, D_{bit})\) and Free Entry complete eqm.
Firms own the Data: Data Intermediary Problem

- A monopolist takes data purchase price as given and sees the downward sloping demand curve for data $p_{bt}(D_{bit})$:

$$\max_{p_{bt}, D_{sit}} \quad p_{bt} \int_0^{N_t} D_{bit} \, di - p_{st} \int_0^{N_t} D_{sit} \, di$$

subject to:

$$D_{bit} \leq B_t = \left( N_t^{-\frac{1}{\epsilon}} \int_0^{N_t} D_{sit}^{\frac{\epsilon-1}{\epsilon}} \, di \right)^{\frac{\epsilon}{\epsilon - 1}}$$

- Free entry at zero cost $\Rightarrow$ zero profits

- Problem incorporates data nonrivalry
  - Buys data once from each firm
  - But can sell the same bundle multiple times
Entry: Innovation Creates a New Variety

- $\chi$ units of labor needed to create an additional variety

- Free entry condition:

  $$\chi \omega_t = V_{it} + \int_0^{N_t} \frac{\delta(\tilde{x}_{it}) V_{it}}{\dot{N}_t} \, di$$

- The value of a new variety and the per-entrant share of business stealing from creative destruction
Firms Own the Data: Definition of Equilibrium

The equilibrium in which firms own the data consists of quantities \( \{c_t, Y_t, c_{it}, x_{it}, \tilde{x}_{it}, a_t, Y_{it}, L_{it}, D_{it}, D_{bit}, B_t, D_{sit}, N_t, L_{pt}, L_{et}\} \) and prices \( \{p_{it}, p_{bt}, p_{sit}, w_t, r_t, V_{it}\} \) such that

1. \( \{c_t, c_{it}, a_t\} \) solve the Household Problem.
2. \( \{L_{it}, Y_{it}, p_{it}, p_{sit}, D_{bit}, D_{it}, x_{it}, \tilde{x}_{it}, V_{it}\} \) solve the Firm Problem.
3. \( (D_{sit}, B_t) \) Data markets clear: \( D_{bit} = B_t \) and \( D_{sit} = \tilde{x}_{it} Y_{it} \)
4. \( (p_{bt}) \) Free entry into data intermediation gives zero profits there (constrains \( p_b \) as a function of \( p_s \))
5. \( (L_{et}) \) Free entry into producing a new variety leads to zero profits:
   \[ \chi w_t = V_{it} + \int_0^{N_t} \delta(\tilde{x}_{it})V_{it} \, di / \bar{N}_t \]
6. Definition of \( L_{pt} \): \( L_{pt} = \int_0^{N_t} L_{it} \, di \)
7. \( w_t \) clears the labor market: \( L_{pt} + L_{et} = L_t \)
8. \( r_t \) clears the asset market: \( a_t = \int_0^{N_t} V_{it} \, di / L_t \)
9. \( N_t \) follows its law of motion: \( \dot{N}_t = \frac{1}{\chi} (L_t - L_{pt}) \)
10. \( Y_t := c_t L_t \) denotes aggregate output.
Firms Own Data: A “No Trade” Law

- What if the government, in an attempt to protect consumers privacy, makes data sharing illegal?

- Government chooses
  - $x_{it} \in (0, 1]$
  - $\tilde{x}_{it} = 0$

- We call this the “Outlaw Sharing” allocation
Consumers Own Data
Consumers own Data: Consumer Problem

- Consumers own data, so now choose how much to share $(x_{it}, \tilde{x}_{it})$:

\[
U_0 = \max \left\{ c_{it}, x_{it}, \tilde{x}_{it} \right\} \int_0^\infty e^{-\tilde{\rho} t} L_0 u(c_t, x_{it}, \tilde{x}_{it}) dt
\]

s.t. \[
c_t = \left( \int_0^{N_t} \frac{\sigma-1}{\sigma} c_{it}^{\sigma-1} di \right)^{\frac{\sigma}{\sigma-1}}
\]

\[
\dot{a}_t = (r_t - g_L)a_t + \omega_t - \int_0^{N_t} p_{it}c_{it} di + \int_0^{N_t} x_{it}p_{st}^a c_{it} di + \int_0^{N_t} \tilde{x}_{it}p_{st}^b c_{it} di
\]

- Firm problem similar to before, but now takes $x, \tilde{x}$ as given, can’t sell data, and has to buy “own” data
Consumers own the Data: Incumbent Firm Problem

- Monopolistically competitive firm takes demand for variety as given (from FOC of consumer problem):
  \[ q_{it} = \left( \frac{c_{it}}{c_{it}} \right) ^ {\frac{1}{\sigma}} = \left( \frac{Y_t}{Y_{it}} \right) ^ {\frac{1}{\sigma}} = p_{it} - x_{it}p_{st}a - \tilde{x}_{it}p_{st}b \]

- Firm buys data on its own variety \((D_{ait})\) and data on other firms varieties \((D_{bit})\)

  \[
  r_t V_{it} = \max_{L_{it}, D_{ait}, D_{bit}} \left[ \left( \frac{Y_t}{Y_{it}} \right) ^ {\frac{1}{\sigma}} + p_{st}^a x_{it} + p_{st}^b \tilde{x}_{it} \right] Y_{it} - w_t L_{it} \\
  \quad - p_{at} D_{ait} - p_{bt} D_{bit} + \dot{V}_{it} - \delta(\tilde{x}_t) V_{it} \\
  \text{s.t. } Y_{it} = D_{ii}^\eta L_{it} \\
  D_{it} = \alpha D_{ait} + (1 - \alpha) D_{bit} \\
  D_{ait} \geq 0, \; D_{bit} \geq 0
  \]
Consumers own the Data: Data Intermediary Problem

- The DI chooses the price at which it sells a firm its own data and the price of other firms data, given its purchase price

\[
\max_{p_{ait}, p_{bit}, D_{cit}^a, D_{cit}^b} \int_0^{N_t} (p_{ait}D_{ait} + p_{bit}D_{bit})di - \int_0^{N_t} (p_{st}^aD_{cit}^a + p_{st}^bD_{cit}^b)di
\]

s.t.

\[
D_{ait} \leq D_{cit}^a \quad \forall i
\]

\[
D_{bit} \leq B_t = \left( N_t^{-\frac{1}{\epsilon}} \int_0^{N_t} \left( D_{cit}^b \right)^{\frac{\epsilon - 1}{\epsilon}} di \right)^{\frac{\epsilon}{\epsilon - 1}} \quad \forall i
\]

- Can not sell more data on firm \( i \) than it buys from consumers

- Can sell all data purchased as “type-b” data to each firm (nonrivalry)
Consumers own the Data: Equilibrium

An equilibrium in which consumers own data consists of quantities \( \{c_t, Y_t, c_{it}, x_{it}, \tilde{x}_{it}, a_t, Y_{it}, L_{it}, D_{ait}, D_{bit}, D_{cit}, D_{cit}, B_t, N_t, L_{pt}, L_{et}\} \) and prices \( \{q_{it}, p_{it}, p_{ait}, p_{bit}, p_{st}^a, p_{st}^b, w_t, r_t, V_{it}\} \) such that

1. \( \{c_t, c_{it}, x_{it}, \tilde{x}_{it}, a_t\} \) solve the Household Problem.
2. \( \{L_{it}, Y_{it}, p_{it}, D_{ait}, D_{bit}, D_{it}, V_{it}\} \) solve the Firm Problem.
3. \( (q_{it}) \) The effective consumer price is \( q_{it} = p_{it} - x_{it}p_{st}^a - \tilde{x}_{it}p_{st}^b \)
4. \( D_{cit}^a, D_{cit}^b, B_t, p_{ait}, \) and \( p_{bit} \) solve the Data Intermediary Problem (with zero profits).
5. \( p_{st}^a \) clears the data market: \( D_{cit}^a = x_{it}c_{it}L_t \).
6. \( p_{st}^b \) clears the data market: \( D_{cit}^b = \tilde{x}_{it}c_{it}L_t \).
7. \( (L_{et}) \) Free entry into new varieties leads to zero profits:
   \[
   \chi w_t = V_{it} + \int_0^{N_t} \delta(\tilde{x}_{it})V_{it} \, di \]
8. Definition of \( L_{pt}: L_{pt} = \int_0^{N_t} L_{it} \, di \)
9. \( w_t \) clears the labor market: \( L_{pt} + L_{et} = L_t \)
10. \( r_t \) clears the asset market: \( a_t = \int_0^{N_t} V_{it} \, di / L_t \)
11. \( N_t \) follows its law of motion: \( \dot{N}_t = \frac{1}{\chi} (L_t - L_{pt}) \)
12. \( Y_t := c_t L_t \) denotes aggregate GDP.
Key Forces: Consumers vs. Firms vs. Outlaw Sharing

- **Firms**
  - use all data on own variety, ignoring consumer privacy
  - restrict data sharing because of creative destruction

- **Consumers**
  - respect their own privacy concerns
  - sell data broadly, ignoring creative destruction

- **Outlaw sharing**
  - maximizes privacy gains
  - no scale effect leads to reduced consumption
Results: Comparing Allocations

1. Planner Problem
2. Firms Own Data
3. Outlaw Data Sharing
4. Consumers Own Data
## Data Sharing

### Own Firm Data

\[ x_{sp} = \frac{\alpha}{1-\alpha} \frac{\tilde{\kappa}}{\kappa} \left( \frac{1}{\tilde{\kappa}} \cdot \frac{\eta}{1-\eta} \right)^{1/2} \]

\[ x_f = 1 \]

\[ x_{os} \in (0, 1] \]

\[ x_c = \frac{\alpha}{1-\alpha} \frac{\tilde{\kappa}}{\kappa} \left( \frac{1}{\tilde{\kappa}} \cdot \frac{\eta}{1-\eta} \cdot \frac{\sigma-1}{\sigma} \right)^{1/2} \]

### Sharing with Other Firms

\[ \tilde{x}_{sp} = \left( \frac{1}{\tilde{\kappa}} \cdot \frac{\eta}{1-\eta} \right)^{1/2} \]

\[ \tilde{x}_f = \left( \frac{\Gamma \rho}{(2-\Gamma)\delta_0} \right)^{1/2}, \quad \Gamma := \frac{\eta(\sigma-1)}{e^{\frac{1}{\sigma-1}} - \sigma \eta} \]

\[ \tilde{x}_{os} = 0 \]

\[ \tilde{x}_c = \left( \frac{1}{\tilde{\kappa}} \cdot \frac{\eta}{1-\eta} \cdot \frac{\sigma-1}{\sigma} \right)^{1/2} \]

- Firms fear creative destruction and share less than planner (\( \delta_0 \))
- Consumers share less than planner because of mark up
- No sharing law restricts data even more
- Firms use more own-variety data compared to consumer/planner
• For $alloc \in \{sp, c, f\}$:

$$Y^alloc_t = \left[ \nu_{alloc} (1 - \alpha) \eta \tilde{x}^\eta_{alloc} \right] \frac{1}{1-\eta} \left( \psi_{alloc} L_t \right)^{1+\frac{1}{\sigma-1} + \frac{\eta}{1-\eta}}$$

• For Outlaw Sharing:

$$Y^{os}_t = \left[ \nu_{os} \alpha \eta x^\alpha_{os} \right] \frac{1}{1-\eta} \left( \psi_{os} L_t \right)^{1+\frac{1}{\sigma-1}}$$

• Two source of increasing returns to scale:
  
  o Standard variety effect: $\frac{\sigma}{\sigma-1}$
  
  o Data sharing: $\frac{\eta}{1-\eta}$

• Recall $\tilde{x}_t > 0$ from data sharing $\Rightarrow$ scale effect
Consumption per person and Growth

- For $\text{alloc} \in \{sp, c, f\}$:
  \[
  c_t^{\text{alloc}} = \text{Const}_{\text{alloc}} \cdot L_t^{\frac{1}{\sigma-1} + \frac{\eta}{1-\eta}}
  \]

- For Outlaw Sharing:
  \[
  c_t^{\text{os}} = \text{Const}_{\text{os}} \cdot L_t^{\frac{1}{\sigma-1}}
  \]

- Which implies per capita growth of
  \[
  g_{sp}^c = g_{f}^c = g_{c}^c = \left(\frac{1}{\sigma - 1} + \frac{\eta}{1 - \eta}\right) g_L
  \]
  \[
  g_{os}^c = \left(\frac{1}{\sigma - 1}\right) g_L
  \]
Numerical Example: How large is $\eta$?

Error rate is proportional to $M^{-\eta}$. Productivity $= \frac{1}{\text{error rate}}$

- Average $\eta = 0.08$. Double data, 6% reduction in error rate
- $\frac{\eta}{1-\eta} \cdot g_L \rightarrow g_y$ is 0.17 pp higher b/c of data sharing
### Numerical Example: Other Parameters

<table>
<thead>
<tr>
<th>Description</th>
<th>Parameter</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Importance of data</td>
<td>$\eta$</td>
<td>0.08</td>
</tr>
<tr>
<td>Elasticity of substitution</td>
<td>$\sigma$</td>
<td>5</td>
</tr>
<tr>
<td>Weight on privacy</td>
<td>$\kappa = \tilde{\kappa}$</td>
<td>0.20</td>
</tr>
<tr>
<td>Population level</td>
<td>$L_0$</td>
<td>100</td>
</tr>
<tr>
<td>Population growth rate</td>
<td>$g_L$</td>
<td>0.02</td>
</tr>
<tr>
<td>Rate of time preference</td>
<td>$\rho$</td>
<td>0.03</td>
</tr>
<tr>
<td>Labor cost of entry</td>
<td>$\chi$</td>
<td>0.01</td>
</tr>
<tr>
<td>Creative destruction</td>
<td>$\delta_0$</td>
<td>0.4</td>
</tr>
<tr>
<td>Weight on own data</td>
<td>$\alpha$</td>
<td>1/2</td>
</tr>
<tr>
<td>Use of own data in NS</td>
<td>$\bar{x}$</td>
<td>1</td>
</tr>
</tbody>
</table>
## Allocations

<table>
<thead>
<tr>
<th>Allocation</th>
<th>Data Sharing</th>
<th>Firm Size</th>
<th>Variety</th>
<th>Consumption</th>
<th>Growth</th>
<th>Creative Destruct.</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>“own” $x$</td>
<td>“others” $\tilde{x}$</td>
<td>$\nu$</td>
<td>$N/L = \psi$</td>
<td>$c$</td>
<td>$g$</td>
</tr>
<tr>
<td>Social Planner</td>
<td>0.66</td>
<td>0.66</td>
<td>1304</td>
<td>665</td>
<td>18.6</td>
<td>0.67%</td>
</tr>
<tr>
<td>Consumers Own Data</td>
<td>0.59</td>
<td>0.59</td>
<td>1482</td>
<td>594</td>
<td>18.3</td>
<td>0.67%</td>
</tr>
<tr>
<td>Firms Own Data</td>
<td>1</td>
<td>0.16</td>
<td>1838</td>
<td>491</td>
<td>16.0</td>
<td>0.67%</td>
</tr>
<tr>
<td>Outlaw Sharing</td>
<td>1</td>
<td>0</td>
<td>2000</td>
<td>455</td>
<td>7.3</td>
<td>0.50%</td>
</tr>
</tbody>
</table>

- Firms overuse their own data and undershare with others
- Consumers share less data than planner, but not by much
- Growth rate scale effect is modest, level differences are large
## Consumption Equivalent Welfare

<table>
<thead>
<tr>
<th>Allocation</th>
<th>Welfare</th>
<th>Level term</th>
<th>Privacy term</th>
<th>Growth term</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>$\lambda$</td>
<td>$\log \lambda$</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Optimal Allocation</td>
<td>1</td>
<td>0</td>
<td>..</td>
<td>..</td>
</tr>
<tr>
<td>Consumers Own Data</td>
<td>0.9886</td>
<td>-0.0115</td>
<td>-0.0202</td>
<td>0.0087</td>
</tr>
<tr>
<td>Firms Own Data</td>
<td>0.8917</td>
<td>-0.1146</td>
<td>-0.1555</td>
<td>0.0409</td>
</tr>
<tr>
<td>Outlaw Sharing</td>
<td>0.3429</td>
<td>-1.0703</td>
<td>-0.9399</td>
<td>0.0435</td>
</tr>
</tbody>
</table>

- **Outlaw sharing**: particularly harmful law (66 percent worse!)
- **Firms own data**: substantially lower welfare (11 percent worse)
- **Consumers own data**: nearly optimal (1 or 2 percent worse)
Implications for IO

- Firms that use data might grow fast compared to those that don’t
- Firms would like to merge into one single economy-wide firm
  - Implications for antitrust
- Targeted mandatory sharing?
  - E.g., airplane safety (after a crash)
- What are the costs of forced sharing?
  - Disincentive to create data
    (in MS, data is aggregate, good approx?)
  - Data as a barrier to entry
    (extension to quality ladder model)
Data versus Ideas: Excludability

• Maybe technologically easier to transmit data than idea (usb key vs. education) . . .

• But data can be encrypted and monitored

• Data seems highly excludable
  - Idea: use machine learning to train self-driving car algorithm
  - ML needs lots of data. Each firm gathering own data
How does data diffuse across firms and countries?
  - Ideas eventually diffuse across firms or countries, so no country scale effect (e.g. HK vs China)
  - What about data?
    - Intellectual property rights

Scale effects and country size
  - Larger countries may have an important advantage as data grows in importance

Scale effects and institutions
  - What if China mandates data sharing across Chinese firms and U.S. has no such policy
Conclusion

- Data property rights likely to affect data use
- Nonrival data ⇒ large social gain from sharing data
- If firms own data, they may:
  - privately use more data than consumers/planner would
  - share less data across firms than consumers/planner would
- Nonrivalry ⇒ Laws that outlaw sharing could be very harmful
- Consumers owning data good at balancing privacy and sharing