Governing Data as a Flow: The Economics of Data and Privacy Protection

Discussant: Laura Veldkamp

Columbia Business School

February 21 2020
Focus of the Discussion

Two separated discussions in the report:
Data ownership and aggregate effects of data.
Do they interact?

- Data Feedback Loop
  - Implies firm ownership of data
- Data → Knowledge Production
  - Data could be owned by individuals or even public.

→ Changes in Distribution of Income
Data Feedback Loop: A Graphical Illustration

More Data

More Transactions / Customers

Higher Quality / Efficiency
Data feedback loop implies that bigger firms get bigger
→ less competition
→ higher markups
→ more inequality.

In the data feedback loop, firms own transactions data.

What if consumers have rights to their data? Or if data is available to all?

There is still a valuable asset in structured data/knowledge.

What are the income inequality implications?
The Knowledge Pyramid

Knowledge

Information

Data

Value

Meaning

High

Low
Formalizing the Knowledge Pyramid

- Knowledge produced via old technology or with big data tech (AI).

\[ K_{it}^{Al} = D_{it}^{\alpha} L_{it}^{1-\alpha} \]  
\[ K_{it}^{OT} = D_{it}^{\gamma} L_{it}^{1-\gamma} \]  

- Information (Structured Data) produced with data management labor and free or abundant raw data:

\[ D_{i,t+1} = (1 - \delta)D_{it} + \lambda_{it}^{1-\gamma} \]  

- Firm Value Function:

\[ v(D_{it}) = \max_{\lambda_{it}, L_{it}, I_{it}} P(D_{it}^{\alpha} L_{it}^{1-\alpha} + D_{it}^{\gamma} I_{it}^{1-\gamma}) - w_{L,t} L_{it} - w_{I,t} I_{it} - w_{\lambda,t} \lambda_{it} + \frac{1}{r} v(D_{i,t+1}) \]
Estimating Knowledge Production

- Data (Abis and Veldkamp (2020))
  - Data is job postings and salaries from Burning Glass. 2007 and January 2010 through March 2019.
  - Job separation and filling rates from BLS data.
- Estimation – cumulating postings to labor stock:
  - $s_t^{\text{type}}$: separation rates by type-month (from BLS, match NAICS codes)
  - $h_t^{\text{type}}$: fraction of posted vacancies filled by type-month (BLS, same)
  - $j_t^{\text{type}}$: Burning Glass job postings rates by type-month

\[
L_{it} = (1 - s_t^{\text{AI}})L_{i,t-1} + j_{it}^{\text{AI}} h_t^{\text{AI}} \quad (5)
\]

\[
l_{it} = (1 - s_t^{\text{OT}})l_{i,t-1} + j_{it}^{\text{OT}} h_t^{\text{OT}} \quad (6)
\]

\[
\lambda_{it} = (1 - s_t^{\text{DM}})\lambda_{i,t-1} + j_{it}^{\text{DM}} h_t^{\text{DM}} \quad (7)
\]

- We estimate optimality conditions to obtain $\alpha$, $\gamma$ and $\phi$.  

Discussant: Laura Veldkamp (Columbia Business School)
Panel 1 shows the fraction of employers hiring in each category. Panel 2 shows the same AI jobs, measured as a number of jobs.
What is happening to the labor share as knowledge production changes?

Answer lies in the production function exponents on Labor:

<table>
<thead>
<tr>
<th></th>
<th>$1 - \phi$</th>
<th>1.0</th>
</tr>
</thead>
<tbody>
<tr>
<td>Data Management</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Traditional Analysis</td>
<td>$1 - \alpha$</td>
<td>0.52</td>
</tr>
<tr>
<td>Machine Learning/AI Analysis</td>
<td>$1 - \gamma$</td>
<td>0.45</td>
</tr>
</tbody>
</table>

Like industrialization, big data technology changes labor share of income: 1/7th fall in labor’s share of income from traditional tech to big data tech/AI.
Big data technology is changing the economic landscape.
Like industrialization, factor intensity and income shares are changing.
Firm ownership of data benefits big firms with many customers.
But even when data is free, there are still winners and losers.