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

Discussant: Laura Veldkamp

Columbia Business School

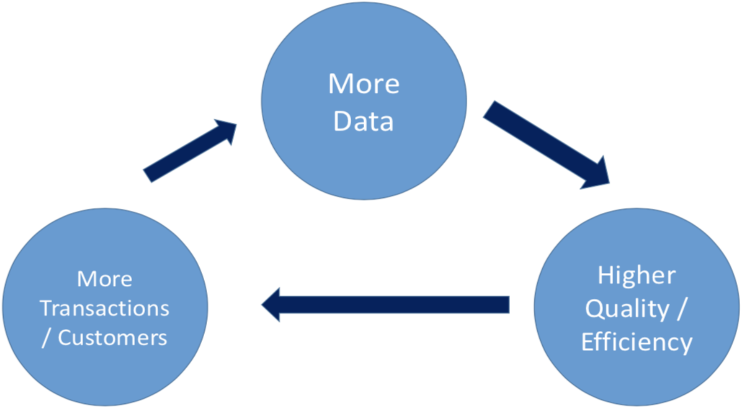
February 21 2020

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 $\rightarrow$ Knowledge Production
  - Data could be owned by individuals or even public.
- $\rightarrow$ Changes in Distribution of Income

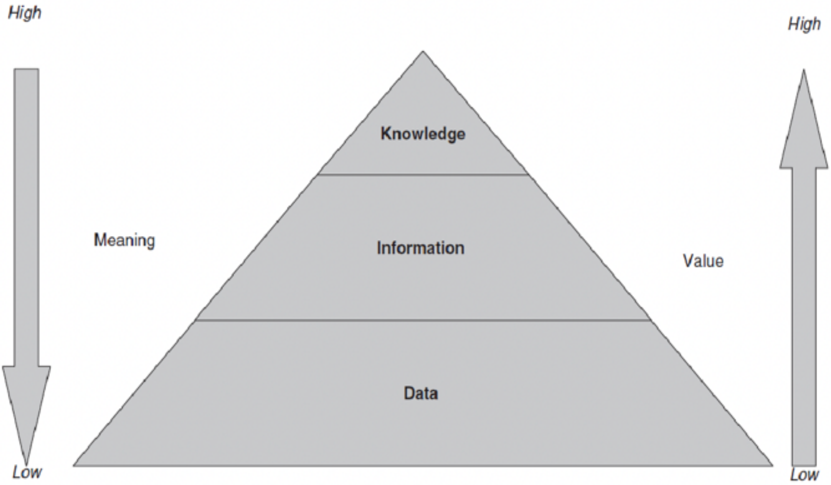
Data Feedback Loop: A Graphical Illustration



Data Feedback and Data Ownership

- Data feedback loop implies that bigger firms get bigger
  $\rightarrow$ less competition
  $\rightarrow$ higher markups
  $\rightarrow$ 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



Formalizing the Knowledge Pyramid

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

$$K_{it}^{AI} = D_{it}^{\alpha} L_{it}^{1-\alpha} \tag{1}$$

$$K_{it}^{OT} = D_{it}^{\gamma} L_{it}^{1-\gamma} \tag{2}$$

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

$$D_{i,t+1} = (1-\delta) D_{it} + \lambda_{it}^{1-\gamma} \tag{3}$$

- Firm Value Function:

$$v(D_{it}) = \max_{\lambda_{it}, L_{it}, l_{it}} P(D_{it}^{\alpha} L_{it}^{1-\alpha} + D_{it}^{\gamma} l_{it}^{1-\gamma}) - w_{L,t} L_{it} - w_{l,t} l_{it} - w_{\lambda,t} \lambda_{it} + \frac{1}{r} v(D_{i,t+1}) \tag{4}$$

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^{type}: separation rates by type-month (from BLS, match NAICS codes)
  - h_t^{type}: fraction of posted vacancies filled by type-month (BLS, same)
  - j_t^{type}: Burning Glass job postings rates by type-month

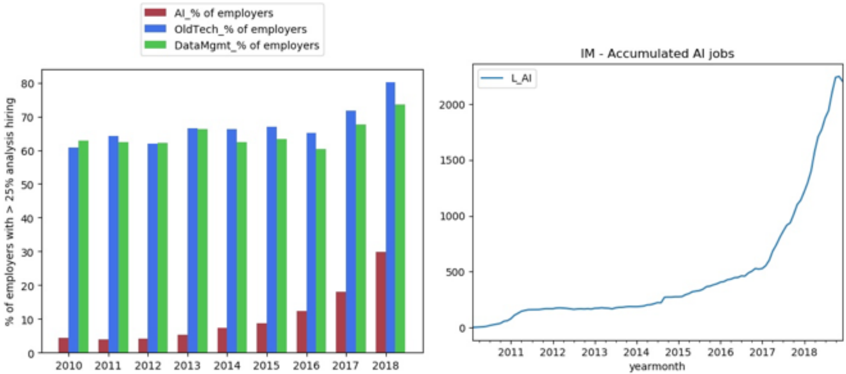
$$L_{it} = (1 - s_t^{AI})L_{i,t-1} + j_{it}^{AI} h_t^{AI} \tag{5}$$

$$l_{it} = (1 - s_t^{OT})l_{i,t-1} + j_{it}^{OT} h_t^{OT} \tag{6}$$

$$\lambda_{it} = (1 - s_t^{DM})\lambda_{i,t-1} + j_{it}^{DM} h_t^{DM} \tag{7}$$

- We estimate optimality conditions to obtain α, γ and ϕ.

Labor Stocks



Panel 1 shows the fraction of employers hiring in each category. Panel 2 shows the same AI jobs, measured as a number of jobs.

Results: Big Data Tech is Reducing Labor Share of Income

- What is happening to the labor share as knowledge production changes?
- Answer lies in the production function exponents on Labor:

Data Management	$1-\phi$	1.0
Traditional Analysis	$1-\alpha$	0.52
Machine Learning/AI Analysis	$1-\gamma$	0.45

- Like industrialization, big data technoloogy changes labor share of income: 1/7th fall in labor's share of income from traditional tech to big data tech/AI.

Conclusion

- Big data technology is changing the economic landscape.
- Like industralization, 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.