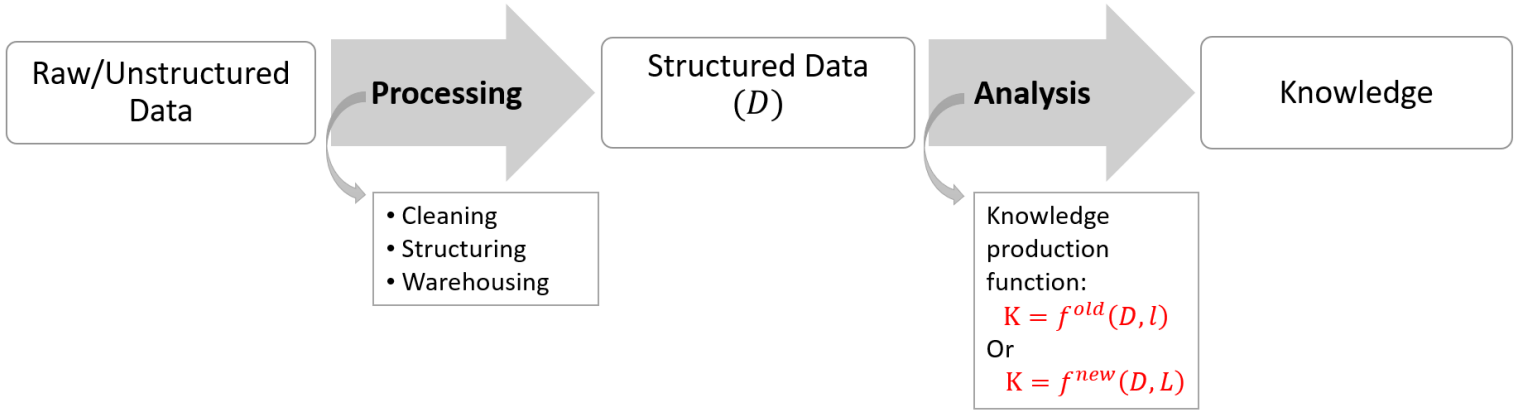
How to Value Data in a World with AI

Laura Veldkamp
Columbia University

based on joint work with Simona Abis

Motivation

- How valuable is data?
  - Many companies are valued based on their data. This value is exploding. Why?
- Key feature of new data technologies: They change the relative intensity of labor and data.
  - How much is AI changing the labor intensity of knowledge production?
  - This matters for employment / labor income share / firm size and competition...
- Investment Management is a good lab because it's a knowledge industry.



Outline

A Model for Measurement

- Knowledge is produced using either the old technology or big data tech (AI). Same data can be used for both. Technologies have different rates of diminishing returns and use differently-skilled labor:

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

$$K_{it}^{OT} = A_t^{OT} a_i^{OT} D_{it}^{\gamma} I_{it}^{1-\gamma}. \tag{2}$$

  A large $(\alpha - \gamma)$ = big revolution

- Data inputs are not raw data. They need to be structured, cleaned and machine-readable. This requires labor (λ) with diminishing marginal returns.
- New structured data is added to the existing stock of structured data. But data also depreciates at rate δ:

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

Maximization

- Firms maximize the value of the firm:

$$v(D_{it}) = \max_{\lambda_{it}, L_{it}, I_{it}} A_t^{AI} a_i^{AI} D_{it}^{\alpha} L_{it}^{1-\alpha} + A_t^{OT} a_i^{OT} D_{it}^{\gamma} I_{it}^{1-\gamma} - w_{L,t} L_{it} - w_{I,t} I_{it} - w_{\lambda,t} \lambda_{it} + \frac{1}{r} v(D_{i(t+1)}) \quad (4)$$

  where (3) holds.

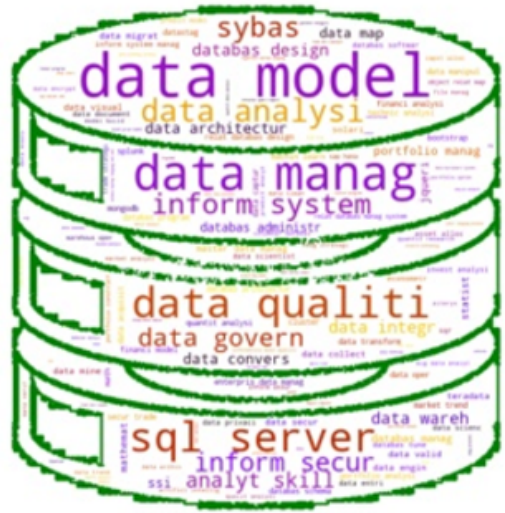
- Optimality conditions equate marginal benefit and marginal cost of workers.

- Taking these conditions to data allows us to identify α, γ and ϕ.

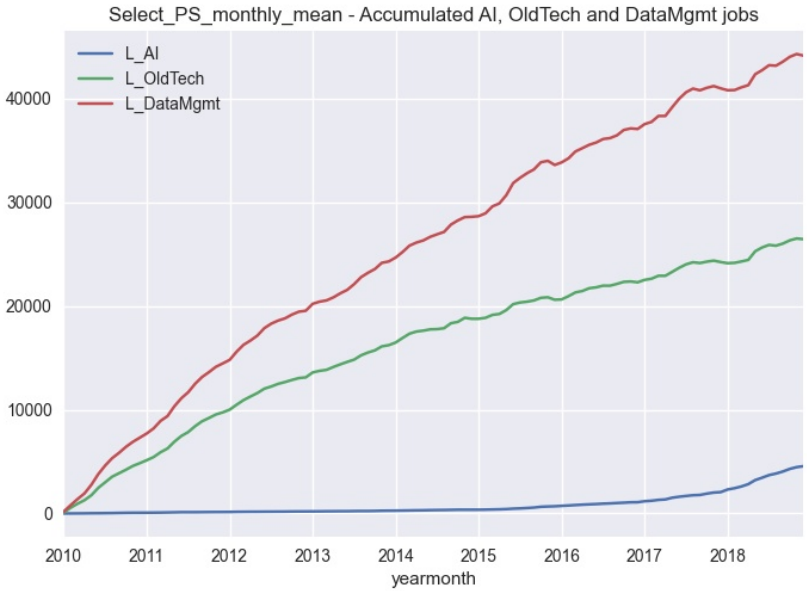
Outline

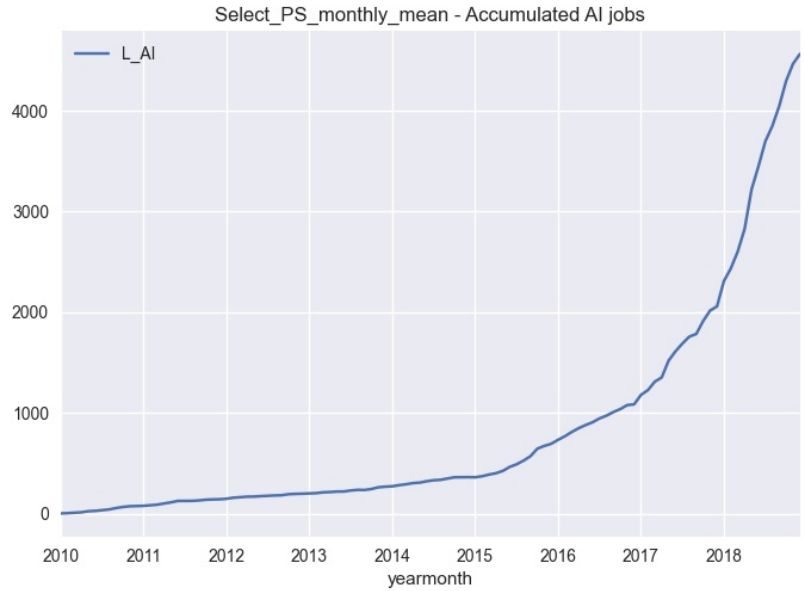
Labor Demand: Job Postings Categorization

- Job postings sample: Burning Glass Technologies (BGT), 2010 – 2018.
- Identify Data Management, AI or OldTech jobs:
  - Develop dictionaries of words and short phrases indicating data management or analysis (AI or OldTech) skills.



Labor Stocks: Growth in AI employment





- AI employment rose 13 times in 2015-2018 (from 350 to 4537 jobs).
- We adjust for job separation rates and fraction of vacancies filled (BLS).
- We don't observe firms' data stock (D_{it}).
  But we can infer it from observing the data managers they hire.
- Our sample contains 308,600 job postings, 33,392 employer-months for 812 unique companies.

Wages: PayScale

- Crowd-sourced salary data from PayScale salary surveys.
- AI-skilled workers earn US$26,333 more per year.

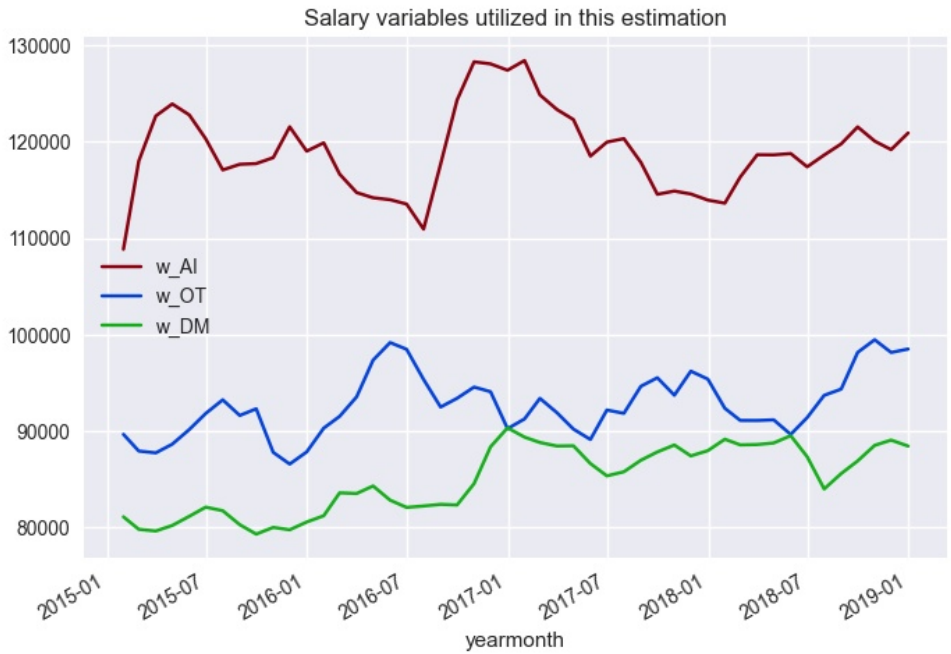


Figure: Median monthly total compensation (salary + bonus) for AI, OldTech and DataMgmt workers. PayScale, 2015-2018.

Outline

Results: AI raises data's income share by 5%.

AI Analysis	α	0.867
		(0.0015)
Old Technology Analysis	γ	0.816
		(0.0020)
Data Management	ϕ	0.453
		(0.0058)
Change in Labor Share	$\gamma - \alpha$	-5.1%

Table: α and γ are Diminishing Returns to Data in AI and Old Tech.

- AI has significantly raised the productivity of analyzing larger data sets: $\alpha > \gamma$
- These exponents also represent the fraction of firm revenue paid to data.
  - Owners of data earn a larger slice of the pie, when firms use AI.
  - Labor share fell from 18% to 13%.

Results: Valuing Data

- Substitute estimated production parameters and data stocks into our value function.
- Data value rose 25% in 4 years (320 to 400 bn).

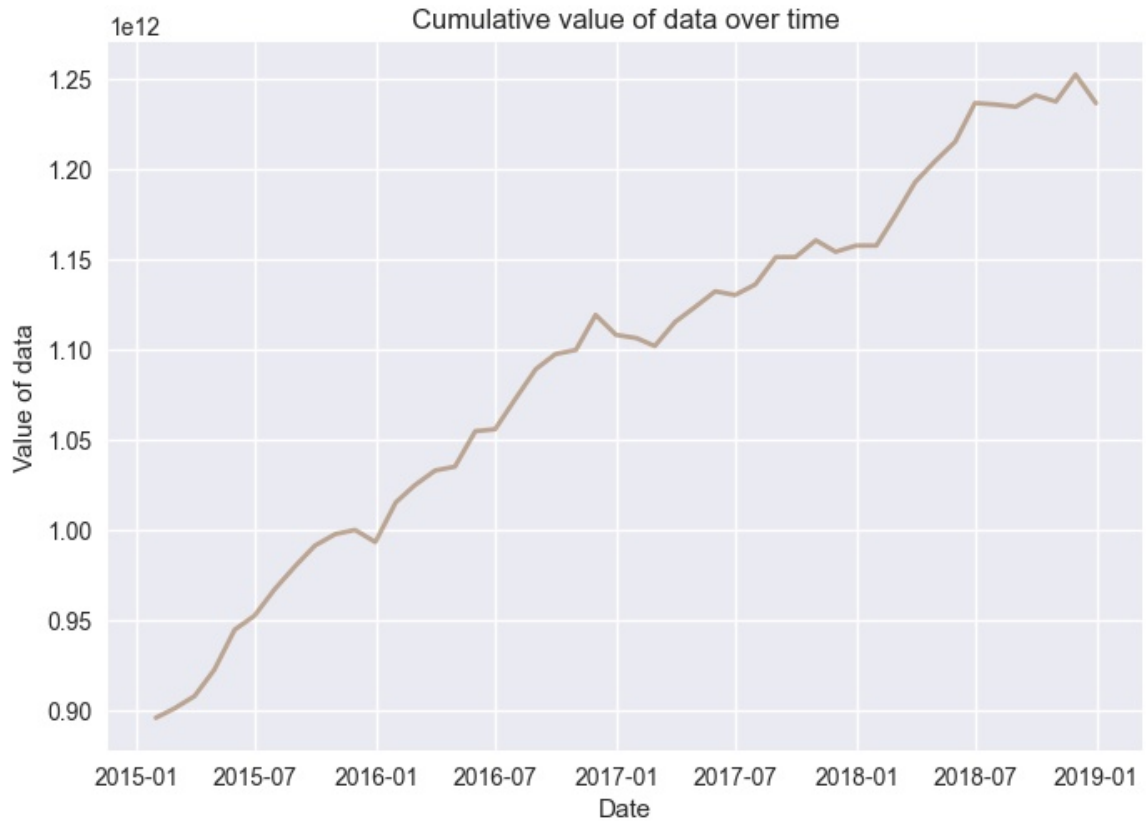
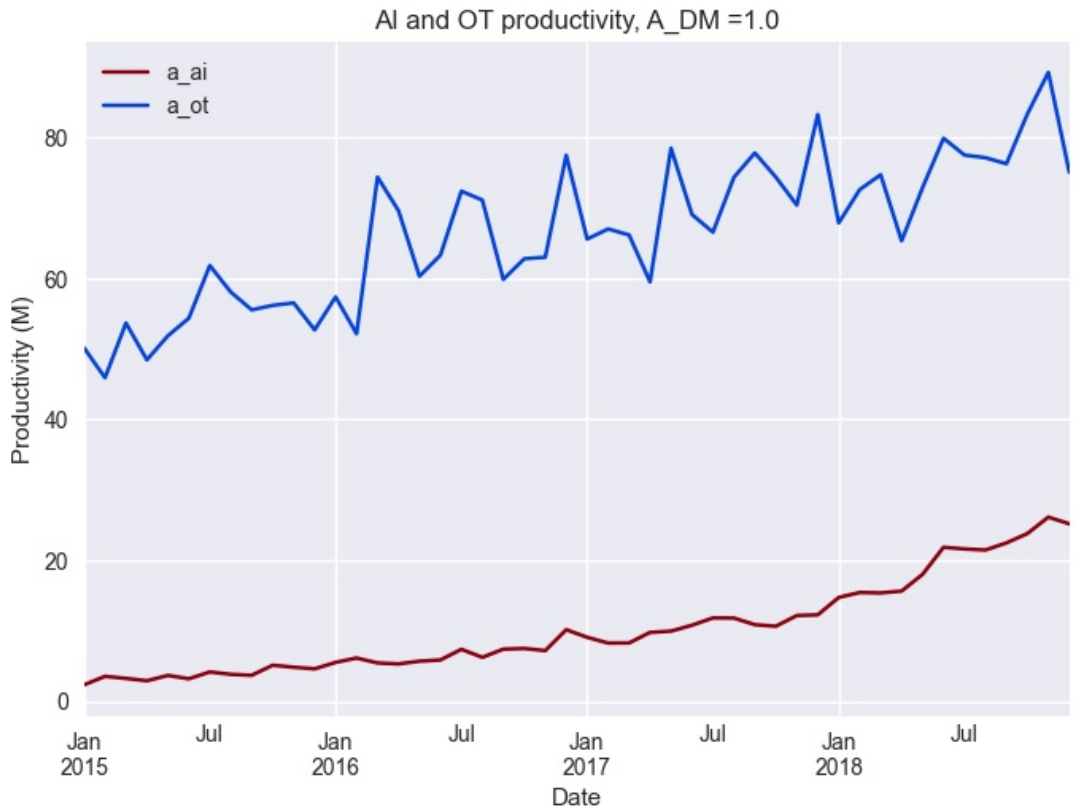


Figure: Estimated Value of the Aggregate Stock of Data, in hundreds of billions of current U.S. dollars, 2015-2018.

AI is Raising the Value of Data in 3 Ways

1. A larger data stock determines a higher cumulative value of data
2. More analysis workers make each data point more valuable
3. Firms are becoming more productive at using data:



Outline

Take-Aways: Data Value is Exploding

- Data is tough to measure. Data management hiring can be an important clue.
- Not only is there more data, the data that is there creates more value for the data owners when it is paired with productive technologies like AI.
- The magnitude of the technological change looks like the industrial revolution – but for knowledge production.
- The change in diminishing returns matters for the value of data as an asset, for inequality and for firm size/competition.