THE ECONOMICS OF DATA: AN INTEGRATED PERSPECTIVE

YAN CARRIÈRE-SWALLOW & VIKRAM HAKSAR

The views expressed herein are those of the authors and should not be attributed to the IMF, its Executive Board, or its management.

LHOFT WEBINAR

MAY 15, 2020

STRATEGY, POLICY AND REVIEW

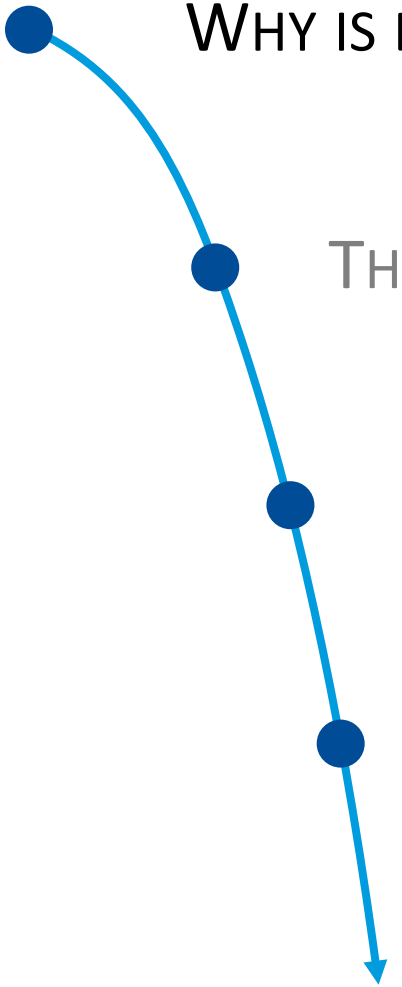
New paper

"The Economics and Implications of Data:
An Integrated Perspective"

IMF Departmental Paper 19/16

Released September 23, 2019

Why is data important now?



The economics of data

Implications for growth, equity & stability

Data policy frameworks

WHY DID WE WRITE THIS PAPER?

- Technological trends are driving data proliferation across the economy.
- Major changes to data policy in many member countries.
- Data raises **macro relevant** policy concerns across the Fund's mandate:

  Concentration and market power • Market opacity • Instability • Trade fragmentation

- Large literature on **individual** aspects—our think piece seeks to **integrate.**
- We argue key policy imperatives are to:
  - complete data markets by defining control/access rights;
  - prevent data hoarding;
  - adequately protect data;
  - avoid international fragmentation.

Effective data policy requires an integrated perspective

Some recent literature

Acemoglu, Daron, Ali Makhdoumi, Azarakhsh Malekian & Asuman Ozdaglar. 2019. "Too Much Data: Prices and Inefficiencies in Data Markets," NBER Working Paper (September).

Acquisti, Alessandro, Curtis Taylor & Liad Wagman, 2016. "The Economics of Privacy," *Journal of Economic Literature* 54(2): 442—92.

Arrieta-Ibarra, Imanol, Leonard Goff, Diego Jiménez-Hernández, Jaron Lanier & Glen Weyl, 2018. "Should We Treat Data as Labor? Moving Beyond 'Free'," *AEA Papers and Proceedings* 108: 38—42.

Begenau, Juliane, Maryam Farboodi & Laura Veldkamp, 2018. "Big Data in Finance and the Growth of Large Firms," *Journal of Monetary Economics* 97: 71—87.

Farboodi, Maryam, and Laura Veldkamp. 2019. "A Growth Model of the Data Economy." Working Paper, Columbia Business School (June).

Furman, Jason *et al.*, 2019. "Unlocking digital competition," Report of the Digital Competition Expert Panel, London, UK: H.M. Treasury.

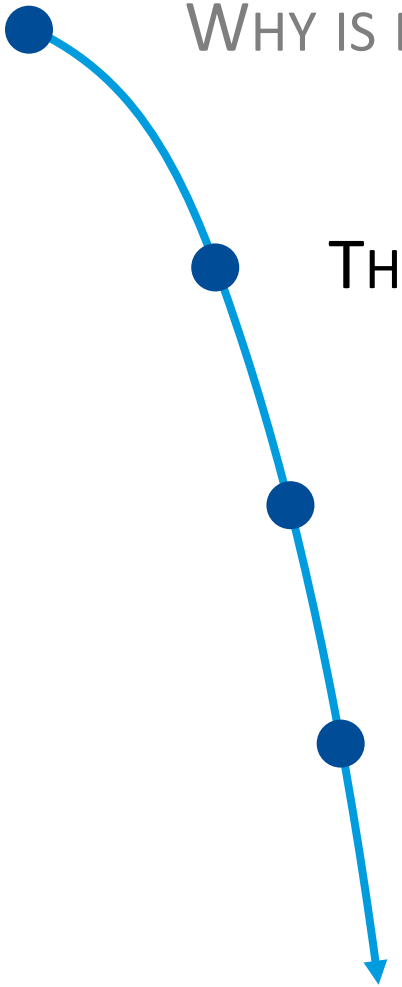
Goldfarb, Avi & Daniel Trefler, 2018. "AI and International Trade," NBER Working Paper (January).

Goldfarb, Avi & Catherine Tucker, 2019. "Digital Economics," *Journal of Economic Literature* 57(1): 3—43.

Jones, Charles & Chris Tonetti, 2019. "Nonrivalry and the Economics of Data," NBER Working Paper (August).

Kashyap, Anil & Anne Wetherilt, 2019. "Some Principles for Regulating Cyber Risk," *AEA Papers and Proceedings* 109: 484—87.

Varian, Hal, 2018. "Artificial Intelligence, Economics, and Industrial Organization," in *The Economics of Artificial Intelligence: An Agenda*, National Bureau of Economic Research.



WHY IS DATA IMPORTANT NOW?

THE ECONOMICS OF DATA

IMPLICATIONS FOR GROWTH, EQUITY & STABILITY

DATA POLICY FRAMEWORKS

What does data do in the economy?

Two broad approaches in the literature:

1. **A factor of production**
   - Data modeled as an input in the production function (Jones & Tonetti, 2019; Farboodi & Veldkamp, 2019).
   - Most salient function of data used in machine learning/artificial intelligence.

2. **Creates information and shifts it across agents**
   - Data collection, sharing and processing affect information asymmetries and frictions (e.g. Begenau, Farboodi & Veldkamp, 2018).
   - Most salient function of individual data, including in finance.

Characteristic 1: Non-Rivalry

One agent's use of data does not preclude its simultaneous use by others

- A potential source of economies of scale and scope:
  - *Scale*: Each unit of data can be used by all units of other factors at the same time.
  - *Scope*: Data acquired for one purpose can be used simultaneously to produce other goods or services.
  - *Bigger is more productive*: Economies/firms with more (complementary) labor and capital will be able to better utilize data.
- At the economy level, socially optimal to make data widely shared.

Characteristic 2: Privacy Externalities

Agents' decisions to create, trade or process individual data impose economic costs on others who may not be compensated

Externality:

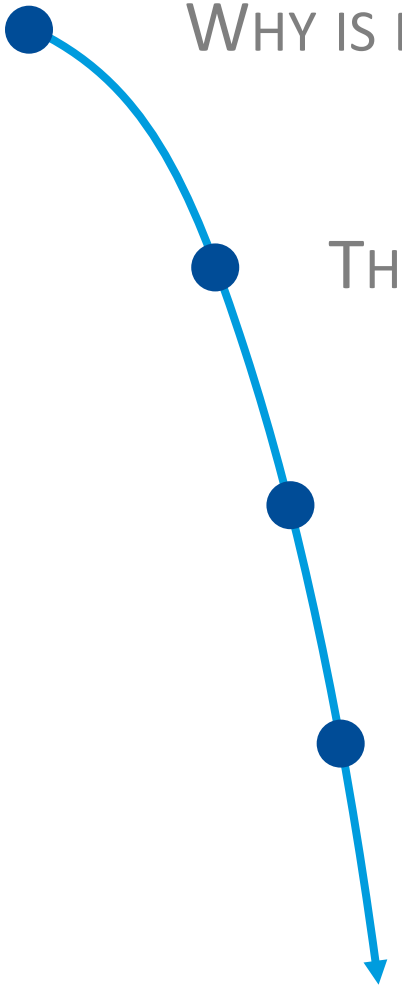
- *Direct costs*: Preference for privacy reflects individuals' desire to keep certain information confidential (Acemoglu *et al.* 2019).
- *Strategic costs*: If individual data is shared (say, their income or their medical condition) with other agents, information asymmetries will shift, and may impose present and future economic costs on that person.

Privacy: The degree of control users have over their data; not how limited is access to it (i.e. more privacy ≠ less sharing)
– Acquisti, Taylor & Wagman (2016)

Characteristic 3: Partial Excludability

It is expensive to control who has access to data

- Firms must decide how much to spend on protecting the data they have collected, making it more excludable and lowering the probability of a cyber event.
- *Externality*: Are there adequate private incentives to pay for cyber security? In the event of a data breach, costs are borne not only by the data controller (i.e. reputation), but also by others through direct (i.e. identity theft) and systemic channels (i.e. trust in the technology or system). (Kashyap & Wetherilt, 2019)



WHY IS DATA IMPORTANT NOW?

THE ECONOMICS OF DATA

IMPLICATIONS FOR GROWTH, EQUITY & STABILITY

DATA POLICY FRAMEWORKS

Implications: Growth and Welfare

Data contributes to growth and welfare:

- Needed for production and innovation in many sectors.
- A source of increasing returns from non-rivalry (Goldfarb & Trefler, 2018).
- Facilitates learning by doing in firms (Varian, 2018).
- Lower information asymmetries can reduce funding costs (Begenau *et al.,* 2018).
- Better matching of products to consumers (including for savings/investment).

But data can increase concentration and create barriers to entry that hurt competition

- Economies of scale from large fixed costs of collecting and managing data
- Increasing returns to scale:
  - Network externalities: On platforms, average utility increases in number of users.
  - Non-linear returns in machine learning: In more complex AI/ML applications, datasets must become very large before they offer any meaningful returns (Agrawal *et al.,* 2018).

And can undermine privacy without compensation, lowering welfare.

Implications: Equity

Data can promote inclusion:

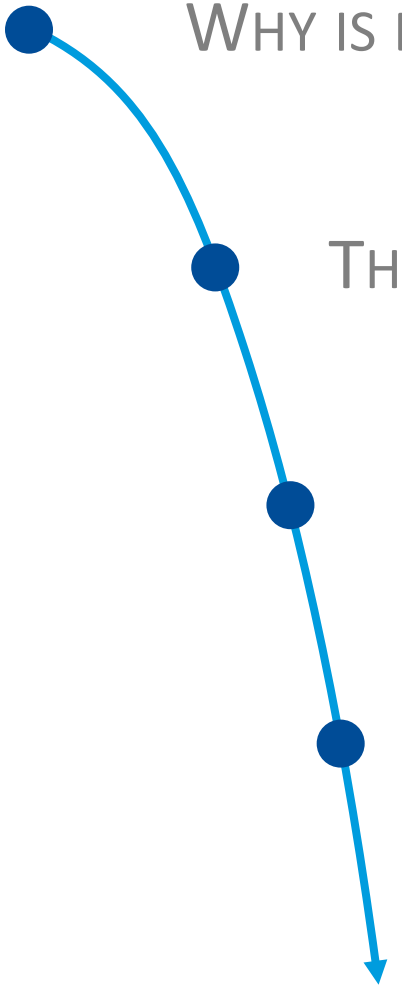
- Better data on borrowers (especially SMEs) may mitigate adverse selection and expand lending opportunities.

But also has the potential to create losers as well as winners:

- Data can weaponize existing market power:
  - Facilitates more granular price discrimination strategies; higher mark-ups.
- Exclusion and bias could also arise:
  - Availability of granular data may lead to the exclusion of high-risk individuals.
  - Granular data may undermine the basis for risk-sharing through insurance.
  - Existing biases and discrimination can be reflected in data-based decision making.

Implications: Stability

- Data will influence **financial stability trade-offs** arising from concentration and competition.
  - Private data hoarding will tend towards less competition in financial services raising concentration risks;
  - Wider data sharing will tend towards more de-centralization and competition but the latter could pose trade-offs between efficiency and excess risk-taking.
- Greater use and spread of **data increases risk of cyber-breaches**. Private incentives to invest in securing data are unlikely to internalize that theft or misuse of sensitive data can undermine trust in systems.
- Non-rivalry of data **influences operational risks**:
  - Use of cloud computing create nodes of risk, but their more sophisticated systems may represent a big improvement for institutions with low capacity.
  - The use of interconnected systems allows for the geographic diversification of operational risks.



WHY IS DATA IMPORTANT NOW?

THE ECONOMICS OF DATA

IMPLICATIONS FOR GROWTH, EQUITY & STABILITY

DATA POLICY FRAMEWORKS

Data policy frameworks: General principles

- Data policy frameworks must balance multiple competing objectives:

Growth

- *Promote competition and innovation;*
- *Exploit benefits of data by facilitating scale through data access;*
- *Ensure incentives for data collection and storage.*

Equity

- *Internalize privacy costs by assigning control rights over individual data;*
- *Establish data rules to protect vulnerable groups from exclusion.*

Stability

- *Address concentration-stability trade-offs;*
- *Create incentives to invest in adequate cyber security;*
- *Avoid operational risk concentration*

- Need for an integrated approach: sector-specific or narrowly defined policy interventions can have side-effects

Data policy frameworks: Some concerns

Status quo data policies raise a number of macro-relevant concerns:

Market opacity

Since control/access rights are unclear, data is likely being overused, with privacy insufficiently respected in the absence of meaningful consent.

Instability

There may be insufficient private incentives to protect user data, subjecting a variety of markets to instability in the event of a major breach.

Growing concentration/market power

Incumbents appear to be earning large rents from hoarding data, forming large barriers to entry that may be stifling dynamic competition.

Trade fragmentation

As countries use very different approaches to data frameworks, there is a risk of international fragmentation in data and goods trade.

THE ECONOMICS OF DATA: AN INTEGRATED PERSPECTIVE

YAN CARRIÈRE-SWALLOW & VIKRAM HAKSAR

The views expressed herein are those of the authors and should not be attributed to the IMF, its Executive Board, or its management.

LHOFT WEBINAR

MAY 15, 2020

STRATEGY, POLICY AND REVIEW

Background Slides

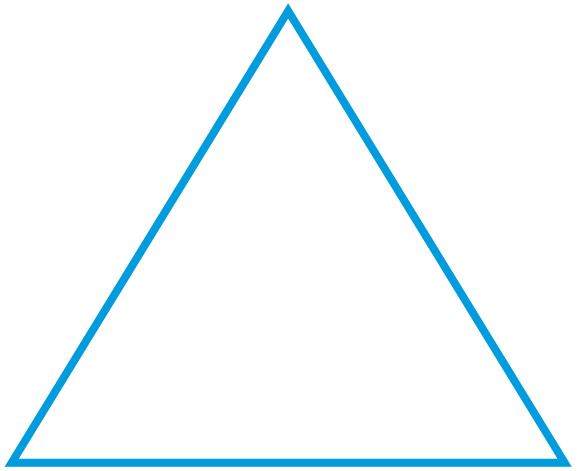
DA·TA (/ˈdatə,ˈdādə/), *noun*:

- A factual representation of a characteristic, action or natural occurrence
- Can be quantitative or qualitative
- Stored on mechanical or digital media
- Jones & Tonetti (2019) propose the following distinction:

What is new about data?

Falling costs

of technologies that collect, store and transfer data



Digitalization

data generated as a byproduct of digital economic activity

Better tools

AI/ML analytics and more computing power allow more value to be derived from data across a wide range of applications and sectors