A Growth Model of the Data Economy

Farboodi and Veldkamp

Discussion by Ezra Oberfield

Princeton University

March 2021

A Growth Model of the Data Economy

- Basic model that can be generalized in many different directions
- Applications/Extensions of basic model
  - Data feedback loop / role of scale
  - Market for data / spillovers
  - When to see specialization in data collection
  - Comparative advantages in production or data collection/processing
- Lays out agenda to study many issues relating big data to firm dynamics and growth

A Growth Model of the Data Economy

- Easy to come away thinking question was: Is big data new conceptually?
  - Not sure it is possible to succeed in giving answer to whether "economics" are new
- Instead, should be viewed as a call to arms to go out and measure things
- Could focus a bit more on laying out roadmap for future empirical work

Is big data new conceptually?

- Many features of data have clear analogs in ideas/technology/productivity
  - Ideas are a type of data
  - Joint output (learning by doing)
  - Feedback look, returns to scale (learning by doing)
  - Partially non-rival (ideas)
  - Can be sold to other firms (licensing)
  - Firms willing to produce at a loss (learning by doing, organizational capital)
  - Firms focusing data production (R&D firms)
  - Eventual diminishing returns (Are ideas getting harder to find?)
  - Stock of knowledge (accumulated R&D)

- Problem: both ideas and data are vaguely defined and can take many forms
  - No "right way" to model ideas/technology
    - ★ Romer/Grossman-Helpman/Aghion-Howitt/Kortum-Klette
  - No "right way" to model big data
    - ★ Many types of data/algorithms/equipment

- My view: Not that productive to ponder how similar they are
  - Instead: thinking concretely about big data can help answer some important questions

Issues Specific to "Big Data"

- Privacy:
  - Consumer do not like it
  - Firms can price discriminate better
- Legal regimes
  - For ideas: patent regimes, licensing regimes
  - For data: Who has property rights? How long can it be kept?
- These are not the focus of the paper (studied elsewhere)

Big Question for Growth

- Many things can improve a firm's (or an economy's) productivity/quality/profitability

- Big questions:
  - What precisely are the mechanisms that generate increasing returns?
  - How strong are increasing returns?
  - What is internalized?
  - Is there scope for policy to improve outcomes? Which policies?

- Key issue: sources of productivity growth take so many forms
  - it can be hard to get know what to measure/estimate

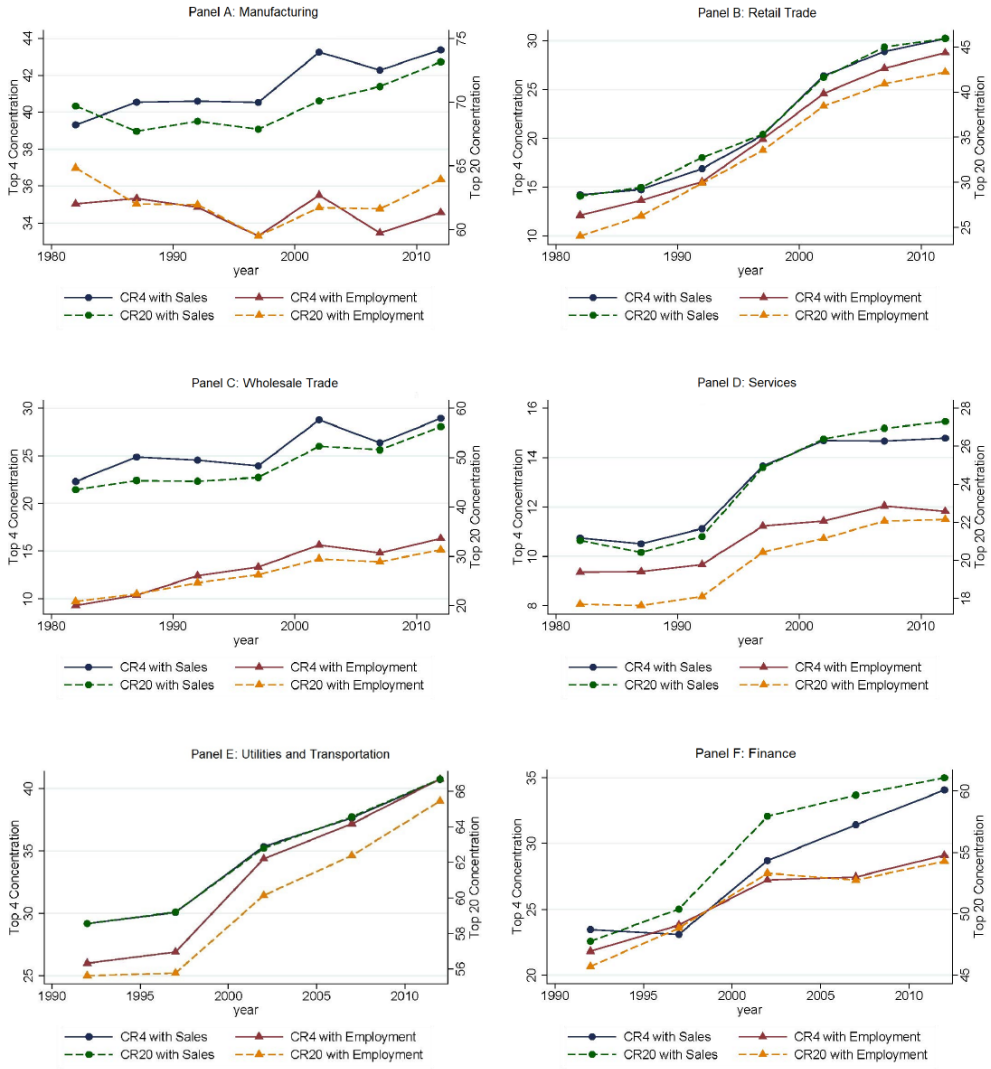
- Value in being concrete about particular source of productivity growth (in this case, big data)
  - How strong is data feedback loop?
  - How much has big data changed firms' returns to scale?
  - Does big data change antitrust prescriptions?
  - What are the measurements we need to answer these questions?
  - Being concrete makes it easier to map model to data

Data and the Advantages of Scale

Why does big data give advantages to large firms?

- Large firms have more transactions to observe
  - Google has more searches
  - Amazon has purchases
- Better opportunities to experiment (is this a corollary?)
  - Suppose Amazon wants to tweak placement of "Buy Now" button
  - Can try it on subset of customers. Scale $\Longrightarrow$ precise answers quickly
  - Small store can experiment, but low customer flow makes it difficult to learn quickly
- Large firms find it worthwhile to make large complementary investments
  - Netflix created ABlaze, interface to make implementing experiments easy
- Google has incentive to pay fixed cost to experiment and increase profit by 0.01%

Rise of Superstar Firms



Big Data and the Rise of Superstar Firms

- Are the forces new?
  - Larger firms always had more incentive to experiment, easier time experimenting
  - Could hire experts/consultants to collect/analyze data old fashioned way

- "Big data" tools allow firms to do this more/better
  - Easier to store data? Easier to sell data? Less costly to analyze data?
  - McDonald's introduced new menu items in some stores.
    - ★ Could always see how item sells
    - ★ Now, can estimate more refined responses e.g., heterogeneity

- Does big data explain the rise of superstar firms?
  - Farboodi, Mihet, Philippon, Veldkamp (2019)
  - Tambe, Hitt, Rock, Brynjolfsson (2020): digital capital much more concentrated in large firms than other types of assets

Scale and Policy Questions

- Are the anti-trust issues different?

- Does data feedback loop create barriers to entry?
  - Classic question with learning by doing (Dasgupta Stiglitz, 1984))

- What do we need to measure to answer these questions?
  - strength of internal/external increasing returns
  - How strong is data feedback loop?
    - ★ How strong is this loop for different types of data?
    - ★ Do the complement each other?
  - How large are barriers to entry?
  - If advantages of scale eventually diminish, when do they diminish?

- Are there the auxiliary predictions that would give confidence in estimates?
  - What are the complementary investments?

Should we subsidize data collection?

- If big data can be distinguished in a legal sense from other intangible capital, can tailor policy specific to data
- Should we subsidize/tax data collection?
- Should we subsidize data sharing?
  - Duplication of efforts?
- Should we do it differently for different types of data?
- Patents give temporary monopoly right to originator in return for disclosing the insight
  - Is there an analog with Big Data?

Zero Long Run Growth

- Data modeled as a posterior precision, Ω: forecast error $\sim N\left(0, \Omega^{-1}\right)$
- In paper, productivity is $A = \bar{A} - b(\text{forecast error})^2$

  $\Longrightarrow \quad E[A] = \bar{A} - b/\Omega$

  $\Longrightarrow \quad \lim_{\Omega \to \infty} E[A] = \bar{A}$
- What if $\log A = \log \bar{A} - b \log(\text{forecast error})^2$, with $0 < b < 1/2$?

  $\Longrightarrow \quad E[A] \propto \bar{A}\Omega^b$

  $\Longrightarrow \quad \lim_{\Omega \to \infty} E[A] = \infty$
- Is one more plausible than the other? What does perfect foresight do? An empirical question?

Zero Long Run Growth

- Most people accept models with no bound on productivity
- Alternative interpretation: data on productivity of production techniques?
- Consider infinite number of possible techniques
  - Each with different productivity (e.g., as in Kortum, 1997)
  - Only finite number have been discovered
  - Alternative interpretation: only have data on finite number of them (many armed bandit)
  - With improved data, can increase effective productivity, possibly without bound

Conclusion

- Stimulating paper
- Lays out agenda for studying many issues related to big data and firm performance
- Would help to give more guidance about empirical counterpart of key elasticities