# Costs of information for firms

*Keywords:information, costs, scaling, reading, firms*

## Extended Abstract

While information is crucial to firms, they are typically characterized in economic terms of capital and labor (1). A major line of economic measurement focuses on the variation amongst firms in the aggregate, or the demographics of firms (2), studies of which have highlighted remarkable regularities in firm inequality building on Gibrat's observation of a heavy tail in wealth (3, 4), growth and how its variation scales with firm size (5), and death which agrees surprisingly well with a constant rate of failure given any single industrial sector (6). In more recent years, such mathematical laws have been polished with as more detail has become available (7). Importantly, the diversity of firms in the aggregate hint at shared mechanisms driving firm behavior and thus serve as a way to test quantitative theories about firm wealth dynamics (8, 9), constraints (10), or management (5). As one counterpoint to the focus on economic variables, the field of collective knowledge focuses on the importance of social skills (11), coworker complementarities (12), teamwork (13), and cultural evolution (14). Yet, missing from the collective knowledge literature are precise measurements that span firms broadly of the information used by employees because of data limitations. Also going beyond economic parameters is the study of the mission statement, which points to the centrality of the ideas that instigate, instead of derive from, capital flows (15). Mission statements, however, reveal little about the information flow through a firm and their study largely qualitative. Finally, we might point to the study of patents as a reflection of information processes of a firm (16), but it is also the case that only a few firms file almost all patents. What is missing from the literature is the demographic account, or variability, of firms along the essential, information dimension to complement the overwhelming attention on firms as economic entities (17).

Here, we discover several aspects of the prodigious information consumption rate of millions of firms and connect them to their economic properties. To do so, we analyze an extensive data set, or "intent data," consisting of hundreds of millions of records of content accessed by firm employees within a large universe of publishers including The Wall Street Journal, Bloomberg, Forbes, Business Insider, and CBSi, along with more specialized groups of sites. Most are anonymous but span technology, marketing, legal, biotech, manufacturing, and a wide range of business services. In principle, the data would allow us to determine which news article a particular employee at a firm accessed and when. For each article, we have up to ten associated topics that have been identified with a proprietary topic modeling algorithm. These different properties permit us to analyze employee reading at different scales of resolution from the individual article, which then belong to sets of content distributors or "sources," and that may overlap in larger "topics." Importantly, the comprehensive nature of the data set allows us to obtain a multiscale portrait of how firms seek out information down to the individual acts of information acquisition.

We demonstrate that firms display power law scaling in the information dimension that is related to but different from known scaling along traditional economic measures,

revealing an economy of scale in *information*. Then, we consider how the volume of reading is related to its diversity, which reveals that there are pivot sizes at which larger firms are significantly more limited in the diversity of information they consume. Finally, we reconstruct the topic space in which firms live to consider how economic growth is accompanied by expansion in the space of topics. We show that firm extent in the space is a combination of topic space structure and bimodal information strategies indicating the role of exploration and exploitation. As a corollary, we predict that firms consume a prodigious amount of information over their lifetimes in a way that scales close to but not quite as fast as assets. These three results indicate the important role of information as a reflection of and perhaps as a driver of economics.

# References

1. Barro RJ, Sala-I-Martin XI (2003) Economic Growth. (MIT Press), Second edition.
2. Carroll G, Hannan MT (2000) The Demography of Corporations and Industries. (Princeton University Press, Princeton, N.J).
3. Gibrat R (1931) Les inégalit économiques. Sirey.
4. Axtell RL (2001) Zipf Distribution of U.S. Firm Sizes. Science 293(5536):1818–1820.
5. Stanley MHR, et al. (1996) Scaling behaviour in the growth of companies. Nature 379(6568):804–806.
6. Daepp MIG, Hamilton MJ, West GB, Bettencourt LMA (2015) The mortality of companies. J. Royal Soc. Interface 12(106):20150120–20150120.
7. Perline R, Axtell R, Teitelbaum D (2006) Volatility and Asymmetry of Small Firm Growth Rates Over Increasing Time Frames, (SBA Office of Advocacy), Technical Report 285.
8. Gabaix X (2009) Power Laws in Economics and Finance. Annu. Rev. Econ. 1(1):255–294.
9. Zhang J, Kempes CP, Hamilton MJ, West GB (2021) Scaling laws and a general theory for the growth of companies. arXiv:2109.10379 [physics].
10. Kempes CP, Koehl MAR, West GB (2019) The Scales That Limit: The Physical Boundaries of Evolution. Front. Ecol. Evol. 7:242.
11. Deming DJ (2017) The growing importance of social skills in the labor market. The Quarterly Journal of Economics 132(4):1593–1640.
12. Neffke FM (2019) The value of complementary co-workers. Science advances 5(12):eaax3370.
13. Wuchty S, Jones BF, Uzzi B (2007) The increasing dominance of teams in production of knowledge. Science 316(5827):1036–1039.
14. Henrich J, Boyd R (2002) On modeling cognition and culture. Journal of cognition and culture 2(2):87–112.
15. Rajasekar J (2013) A Comparative Analysis of Mission Statement Content and Readability. Journal of Management Policy and Practice 14(6):131–147.
16. Jaffe A, Trajtenberg M, Fogarty M (2000) The Meaning of Patent Citations: Report on the NBER/Case-Western Reserve Survey of Patentees, (National Bureau of Economic Research, Cambridge, MA), Technical Report w7631.
17. Larsen J (2001) Knowledge, Human Resources and Social Practice: The Knowledge-Intensive Business Service Firm as a Distributed Knowledge System. The Service Industries Journal 21(1):81–102.

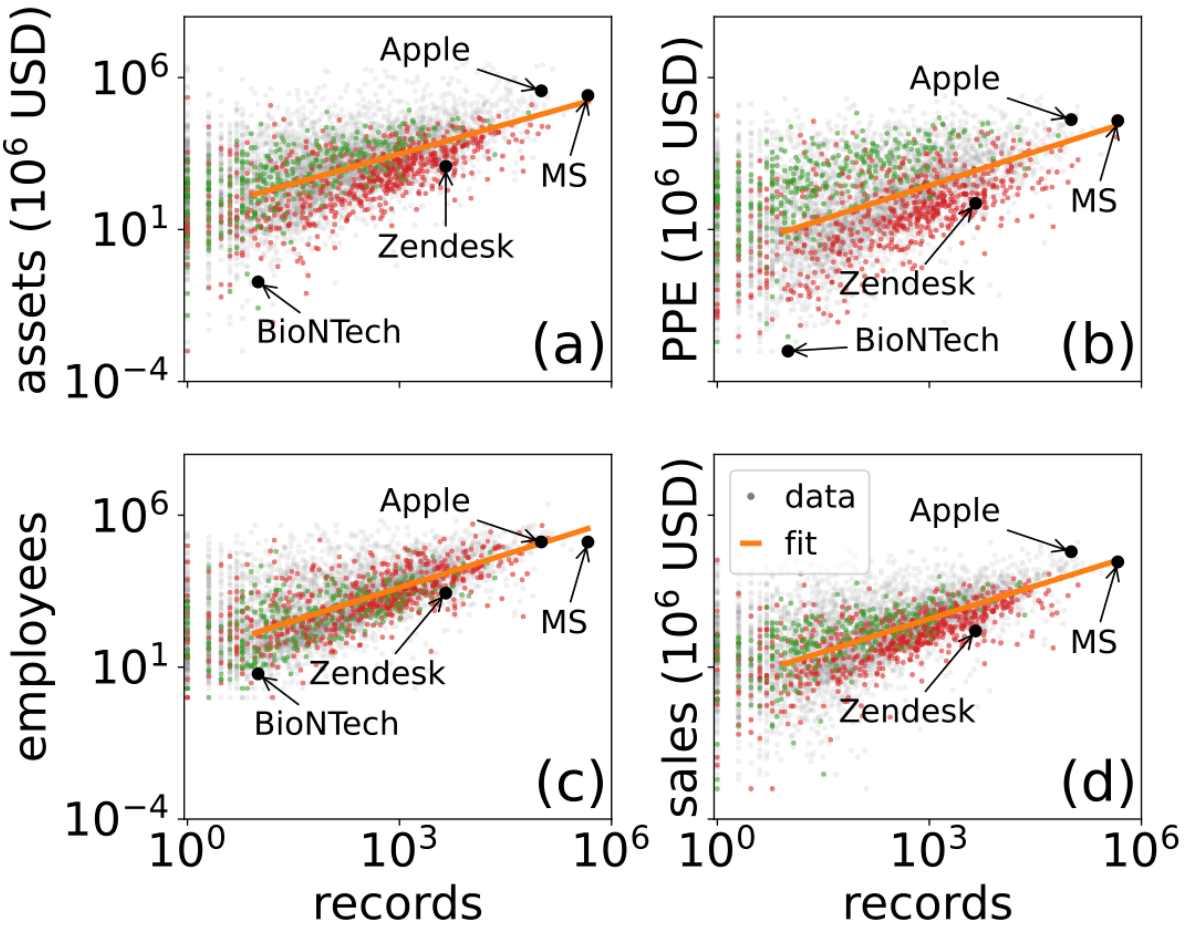


Figure 1. Scaling of record count with measures of capital (a) assets, (b) plants, property, equipment (PPE), (c) employees, and (d) sales. Each blue point is a firm in COMPUSTAT and our data set. Mining (green), service (red), and all other (gray) firms. Orange line shows a power law fit to $Y \propto X^{\beta}$ with exponents (a) $\beta = 0.65 \pm 0.02$, (b) $\beta = 0.74 \pm 0.02$, (c) $\beta = 0.73 \pm 0.01$, and (d) $\beta = 0.72 \pm 0.01$ using one standard deviation from bootstrapped fits as error bars. Fitting range $R \geq 9$ given by power law fits to the distributions. We highlight well-known firms Microsoft, Apple, Zendesk, and BioNTech. BioNTech had no reported sales in 2018.