Data as a Networked Asset

A Brief Summary

Link to the full paper: https://yeli-macrofinance.com/data_value.pdf

Data is non-rival: one firm's use of data does not diminish its availability to others, allowing the same data to be used by many firms simultaneously (Jones and Tonetti, 2020). Data also exhibits externalities: information collected by one firm on its customers can benefit other firms by revealing consumer preferences and broader economic forces and by serving as training data for prediction models. Non-rivalry, combined with externalities, expands the potential uses of a firm's data, often transcending industry boundaries. We ask: What is the marginal q of data as a productive asset? How is data collected by one firm shared with others, and what are the scope and network structure of data sharing? How does data sharing affect firms' decision-making and valuation and propagate shocks across firms? Which firms are systemically important in the data economy? And what are the impact and unintended consequences of data regulations?

We trace inter-firm data flows originating from mobile applications (apps), which have become the primary channels for data collection in the economy. Firms' apps collect data and transmit it to software development kits (SDKs) that specialize in data aggregation and analytics. A crucial function of these "data-sharing SDKs" is merging information from different firms' mobile apps associated with the same consumer into a comprehensive customer profile; by sharing data with the SDKs, firms gain access to these valuable signals in return. Using granular information on app-level SDK installations, we construct a measure of data-sharing conduits between firms—two firms are connected when a common set of data-sharing SDKs is installed on their apps. Data sharing turns out to be a powerful economic linkage. It induces comovement in firms' operational performances, with an economic magnitude more than double that of the product similarity measure of Hoberg and Phillips (2016). Moreover, data-connected firms exhibit stock-return correlations that cannot be explained by standard asset-pricing factors or by other common exposures, such as product similarity, supply-chain linkages, and common analyst coverage.

Our paper is the first to uncover this new form of economic linkage, and it speaks directly to policy debates on data privacy. With rising consumer awareness of privacy, regulations such as GDPR in Europe and CCPA in California have imposed strict rules on data collection and sharing. In a difference-in-differences framework, we examine the introduction of Apple's App Tracking Transparency (ATT), which allows app users to block identity tracking and thus makes it difficult for SDKs to merge data from different firms. We find that ATT significantly weakens the performance comovement between data-connected firms—an unintended consequence of privacy initiatives. This experiment also validates that firms' overlap in data-sharing SDKs is fundamentally about linkages in data: ATT does not affect the contribution of other inter-firm overlaps, such as product similarity, to performance comovement.

Because comovement ultimately emerges from shocks propagating from one firm to another, we further provide direct evidence using cyberattacks, which destroy firms' data stocks and impair data collection by diminishing the customers' willingness to engage with the mobile apps. When a firm suffers a cyberattack, its data-connected peers experience significantly greater deterioration in operational and stock-market performances than firms that are not data-connected.

Which firms, then, are systemically important in the data economy? Answering this question requires accounting not only for direct spillovers but also for higher-order ones—firm A's data may be shared with firm B, which in turn shares data with firm C—and for their persistent effect over time, as data drives customer engagement and engagement generates further data (Farboodi and Veldkamp, 2026). We therefore develop a dynamic network model in which data plays the role of productive capital. Each firm forecasts customer preferences with a precision that depends on both its own data and data from connected peers; more accurate forecasts enhance customer engagement, which boosts customer capital, product demand, and cash flows, and creates new data for the firm and its peers. The model takes the empirically measured network of data-sharing conduits as input, and simulations of the calibrated model replicate the data-induced comovement and the ATT effects. This framework allows us to derive a network-augmented Gordon growth formula for data-driven firms' valuation that captures direct and indirect data spillovers and the self-perpetuating growth of data over all horizons; absent data sharing, it collapses to the standard Gordon growth model. Decomposing the aggregate valuation of all firms into individual firms' contributions, we find that a firm's contribution can differ substantially from its size: before the introduction of ATT, the ratio of Meta's aggregate-valuation contribution to its own valuation was 4.2—removing Meta would reduce the present value of aggregate cash flows by an amount equivalent to 4.2 times its own valuation—and after ATT, this ratio declined to 1.2.

References

Farboodi, M. and L. Veldkamp (2026). A model of the data economy. The *Review of Economic Studies*, forthcoming.

Hoberg, G. and G. Phillips (2016). Text-based network industries and endogenous product differentiation. The *Journal of Political Economy* 124(5), 1423-1465.

Jones, C. I. and C. Tonetti (2020). Nonrivalry and the economics of data. The *American Economic Review* 110(9), 2819-2858.