The question of data has been at the center of recent debates around competition policy in the digital era. Concerns in this area are wide-ranging, and encompass privacy, collusion, barriers to entry, exploitative practices, and data-driven mergers.

Data can serve several purposes: for instance it can be used to improve algorithms, to target advertising, or to offer personalized discounts to consumers. While this heterogeneity of uses for data has sparked a large literature in economics, the multiplicity of models makes it difficult to draw general conclusions about the competitive effects of data.

In this paper we introduce data into a competition-in-utility framework à la Armstrong and Vickers (2001). The three key features of data are that (i) it allows to generate more revenue for a given level of utility, (ii) it is a byproduct of firms' economic activity, and (iii) it is a club good (non-rival and excludable).

We provide a sufficient condition for data to be pro-competitive, and apply it to several environments illustrating the variety of uses for data. This analysis sheds light on the tension between the static and the dynamic effects of data: this is precisely when data increases short-term equilibrium consumer surplus that it can be used as a barrier to entry or that it can result in market-tipping.

We then use this framework to study data-driven mergers. We consider two data-connected markets A and B: the data obtained as a byproduct of firm A’s activity can be used by firms in market B. While the concerns expressed so far by antitrust authorities revolve around the idea of input foreclosure, we show that a merger between firms in the A and B markets also affect the incentives to collect data. A critical condition for the merger to be desirable is for data-trade to be impossible absent the merger.