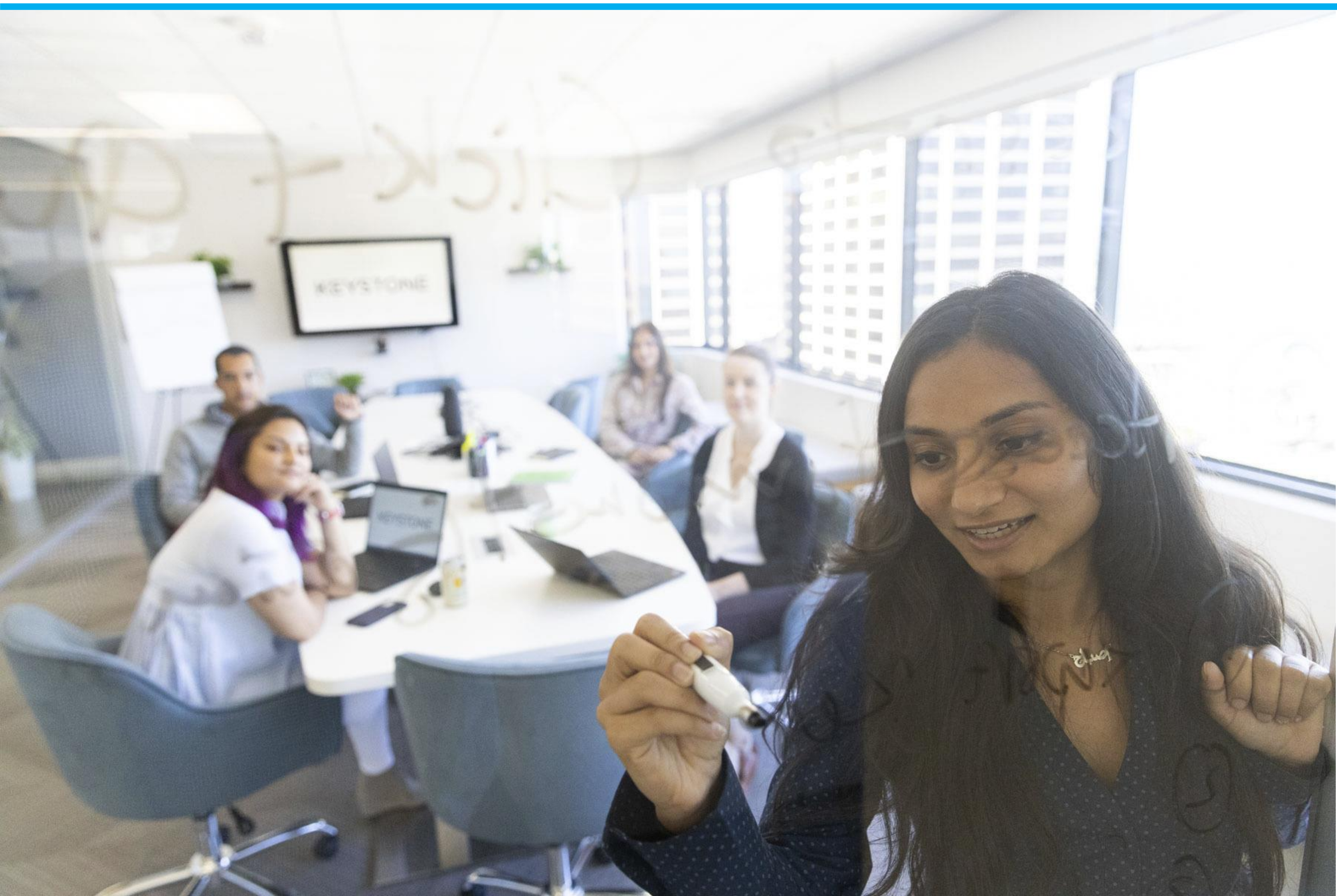
The Next Frontier in Merger Review:

How to Use AI and Other Big Tech Techniques to Assess Horizontal Mergers – Even in Traditional Sectors

April 2024
KEYSTONE
keystone.ai

By Patrick Bajari, Gianmarco Calanchi and Tega Akati-Udi



Introduction

Since joining the Department of Justice (DOJ) in July 2022, Susan Athey, Chief Economist of the Antitrust Division, has strived to expand the Expert Analysis Group within the Antitrust Division by incorporating people with "particular expertise in algorithms and modern tools for analyzing larger datasets, as well as machine learning."[1] On the other side of the Atlantic, in 2019, the UK's Competition and Markets Authority launched its Data, Technology and Analytics (DaTA) unit to help it "stay ahead, using the latest in data engineering, machine learning and artificial intelligence techniques."[2] And the European Commission has recently created the Data Analysis and Technology Unit to create a consolidated team driving the use of data science and market monitoring.[3]

So, one trend we can certainly see in merger review is towards more sophisticated data analysis: we expect to see "Big Tech-like" techniques deployed in merger reviews. Quite a change!

But does this mean that data analysis will have more space in merger reviews? Another parallel trend we observe among the leading agencies is towards greater reliance on qualitative evidence and internal documents. At face value, this would mean *less* space for quantitative analysis, not more. How do these seemingly opposing trends – one towards more quantitative analysis and one seemingly towards less of it – reconcile?

In reality these apparently opposing trends have a common root: on the one hand, increasing skepticism towards some of the "more traditional" econometric techniques employed in merger review, which have often been found wanting when assessing *ex post* their predictive effect.[4] And, on the other, increasing interest in **embedding business evidence and, altogether, a "business perspective" in the assessment of the likely effects of mergers.**

Business internal documents clearly meet the brief – albeit with the usual precautions on how to read them jointly with other sources of evidence. But sophisticated analyses based on Big Data techniques do too. Many companies are growing increasingly sophisticated in their ordinary course of business analyses. They increasingly rely on vast amounts of granular data (either collected or purchased) and use sophisticated tools to analyze it. **This trend extends well beyond Big Tech or digital markets, involving firms in many traditional sectors such as retail.** Using quantitative techniques that mirror those employed by these companies is an effective way to embed business evidence and a "business perspective" into the quantitative analyses that are relevant for merger control reviews. As Susan Athey put it:

> "We're studying businesses…so of course, a core discipline in a business school is economics. But there's also behavioral science. There's operations, statistics, now a lot of machine learning, sociology. All of these disciplines are present, and they

[1] Susan Athey, Keynote on Computational Antitrust at the DOJ, 9 January 2024. Keynote: "Computational Antitrust at the DOJ" (Susan Athey) (youtube.com)
[2] The CMA DaTA unit – we're growing! – Competition and Markets Authority (blog.gov.uk)
[3] MLex | Digital enforcement triggers restructure in EU antitrust department
[4] See, for example, Kwoka, John E., Does Merger Control Work? A Retrospective on U.S. Enforcement Actions and Merger Outcomes (April 4, 2012). Antitrust Law Journal , Vol. 78, 2013, Available at SSRN: https://ssrn.com/abstract=1954849 or http://dx.doi.org/10.2139/ssrn.1954849

inform the business disciplines…that type of interdisciplinary approach. It makes sense in academics. It makes sense in business. And it makes sense for those studying businesses."[5]

So, whether we will see more data analysis or less in merger review going forward will ultimately depend on the ability of economic advisors to produce more "realistic" analysis, based on state-of-the-art quantitative techniques, and move away from some of the "more traditional" econometrics that have been found wanting in terms of their predictive power.

Keystone categorizes state-of-the-art quantitative techniques into two broad groups: (i) techniques that improve upon more traditional approaches to analyses, and (ii) the application of novel AI/ML tools pioneered by Big Tech in a host of sectors.

But what are these more "realistic" analyses? In this article, we examine a few of these techniques, grouping them into two broad categories: (i) techniques that improve on the "more traditional" econometric analyses for merger review and (ii) novel Machine Learning ("ML") and Artificial Intelligence ("AI") tools currently used by Big Tech companies (but not only by them) in the ordinary course of business to learn causal effects. We have deployed some of these techniques in recent merger cases in traditional industries and found them very effective at addressing the typical limitations of the "more traditional" techniques.

This is an opportunity but also a challenge: deploying these techniques imposes a significant change on how economic advisors[6] conduct quantitative analyses for merger review. They need to develop three new core sets of internal capabilities:

(i) **Big-Tech-level data engineering and logistics.** When many terabytes or petabytes of data are involved, computational constraints bind. We are no longer in a world in which running a regression takes a few lines of code and a few seconds to run. Instead, it requires a deep understanding of basic software and data engineering to distribute the computations across multiple machines in the cloud. It requires expertise with appropriate open source or proprietary cloud-based technologies (e.g. tools and services on Amazon, Microsoft or Google's Cloud). Staffing a merger project with software engineers with experience developed at Big Tech companies is essential.

(ii) **Experience with "Big-Tech econometrics."** Linear regressions – the econometric technique typically employed with "normal-sized" datasets – are not as effective when dealing with thousands or millions of control variables. ML has been shown to be far superior.[7] We will discuss some of the ML techniques employed in merger analysis below. Staffing a merger project with econometricians with experience developed at Big Tech companies is essential.

(iii) **Knowledge of (and experience with) company data and third-party datasets**. Companies often use similar types of systems to collect data and employ similar processes, although they differ in the complexity

[5] A Conversation with Susan Athey - ProMarket

[6] Academic economists are also adopting these Big Tech econometric techniques. See Susan Athey, The Impact of Machine Learning on Economics, May 2019 The Economics of Artificial Intelligence: An Agenda (nber.org)

[7] For example, Google's Chief Economist Hal Varian explains that ML presents two key advantages over standard econometrics: (i) a large number of (control) variables means there is a high risk of "multicollinearity" (i.e., the control variables are correlated with each other and it is not possible to tell apart the effect of the variable of interest from other controls), which ML can deal with; and (ii) we can estimate more flexible relationships than with simple linear models. Hal Varian, June 2013, "Big Data: New Tricks for Econometrics."

and amount of data collected, with Big Tech being the upper bound. Staffing a merger project with people with experience developed at Big Tech companies is essential. They really understand (i) what data companies collect in the normal course of their business and how they look at that data – i.e., what tools and statistics they use in their course-of-business analyses and how they *interpret* the data for their business strategy – and (ii) what third-party datasets are available for purchase, and their advantages and limitations (these datasets can be very expensive, so it is important to know their limitations in advance). This goes well beyond the scope of what can be learnt within the timeframe of a single case.[8]

Group 1: Big Data techniques that improve on "more traditional" merger analyses

Big Data techniques can be used to improve on four types of "more traditional" econometric analyses: entry (or intrusion) analysis, demand estimation, merger simulation, and merger efficiencies.

We can group the "more traditional" econometric analyses for merger review into four categories.

Category 1: Entry (or intrusion) analysis.[9]

For example, in a retail merger, take a store of company A and estimate how its revenues (as well as quantities and prices in more sophisticated versions) changed in percentage terms after the opening of a company B's store. The stronger the entry effect, the closer A and B are as rivals. This technique is commonly used to test closeness of competition between the merging parties and with other rivals. Diversion ratios[10] are also commonly derived from entry effect estimates.

The "more traditional" analyses typically rely on monthly store-level sales data collected from the merging parties. These data may not be granular enough to identify a statistically significant effect of entry on the incumbent's sales, particularly when the market is very competitive, because the effect can be very subtle and therefore harder to measure.

This can put the merging parties in a difficult position: these models are more likely to deliver statistically significant estimates in concentrated markets, where the entry of a new rival has stronger effects on incumbents' sales, than in competitive markets, where the entry effect is subtler. A statistically insignificant estimate carries nowhere near the probative value of a small and statistically significant estimate, because an estimate may be insignificant even when the model is not well specified, or when there is insufficient data to narrow

[8] To illustrate this point, building an attribution model using clickstream data in Google Search requires a deep knowledge, among others, of what constitutes a "session" or how the "price" is exactly measured. For example, it is not uncommon for a session to have a half-life of just a few days: if the sessions in the control group and the treatment group are different in measurable and unmeasurable ways, this can make a critical difference to the final estimates of effects. Similarly, it may be important to understand whether the "price" in our dataset is an end-of-the-week price, a median price, or some other value. These are just two of many aspects of the data which require deep prior knowledge.
[9] It is also possible to conduct a similar analysis based on the exit of a rival instead of entry. For simplicity of exposition, here we focus on entry.
[10] Diversion ratios measure the percentage of company A's lost sales when it raises prices (or closes shop) that are captured by company B – the diversion ratio from A to B.

down the confidence intervals. So, these models tend to be biased against the merging parties. **Big Data techniques can assist with this problem.**

- **First, using Big Data techniques we can make the best use of the merging parties' data.** Large retail chains often keep record of the prices and quantities of each product (or SKU) sold at their stores. So, rather than using monthly sales revenue data at store level, we can use weekly (or even daily) revenue (before and after discounts and promotions) and quantity data for each SKU sold in each store, over a few years' period. Using terabytes or petabytes of SKU-level data gives us much more leverage to reduce bias compared to a store's total monthly sales. With appropriate techniques we can control for a tuple of SKU, store and time effects and compare price/quantities for the same SKU across similar stores before and after entry, rather than looking at store sales in aggregate. The more we control, the more accurate our estimates.
- **Second, we can exploit customer-level datasets available from third-party providers.** Many data providers offer datasets that track the behaviors of individual consumers. For example, some providers offer payment card transaction-level data that contain information on all the transactions made by tens of millions of consumers' cards, some also identifying the precise outlets in which customers made their purchases. Other providers track consumers' movements through their mobiles, so it is possible to monitor individual consumers' movements across stores. When combined with the information on store openings or closures, these data permit us to track patterns of change in consumer spend and store visits when a new store is opened (or closed) in the area. We can generate detailed customer profiles based on information on a customer's complete history of past purchases (including text and pictures associated with those purchases). And we can marry the tools of machine learning/deep learning with the classic question of separating causation from correlation.

While traditional econometric analyses typically rely on store-level sales data, Big Data techniques can leverage more granular data for more accurate and localized estimates and a wider breadth of insights.

Using these datasets presents several advantages.

- **First, greater precision of the estimates:** using an incomparably more granular dataset reduces the variance of the entry estimates, so we can obtain a statistically significant entry effect estimate even when this effect is very subtle. For example, one approach is to estimate the entry effect for each incumbent store individually using a random sample of other incumbent's stores that have not experienced entry as control group, and then pool together all these entry estimates across all incumbent's stores. This approach produces a very precise estimate of the entry effect at network-level (which can also be localized for each overlapping area, see next point). This entails a significant computational challenge, which requires Big Tech-level data engineering skills, but delivers very accurate estimates of the entry effect.
- **Second, better localization of the estimates:** not only can we estimate the entry effect even when it is subtle, we can also estimate how it varies depending on the relevant circumstances of each local area

– e.g., how it declines with the distance of the entrant from the incumbent, how it varies with the number of stores belonging to the incumbent, to the entrant or to third parties, etc. This is very important to derive reliable diversion ratio estimates not only at network level but also for each local area of overlap between the merging parties. But it requires vast amounts of very granular data, which can only be processed using Big Data techniques.

- **Third, wider breadth of insights:** using SKU-level data, we can separate the impact of entry on prices and quantities, as well as on different groups of products. This permits us to test, for example, whether a particular rival may be competing for part of the product range. Using consumer-level data, we can learn the diverse impact of entry on different customer groups. This provides far richer insights into competition in the market.

Category 2: Demand estimation.

When company A increased (or decreased) prices, how much did its sales decline (increase) by? How about the sales of company B? This technique is commonly used to estimate diversion ratios.

Big Data techniques can help prevent endogeneity in demand estimation by controlling for significantly more variables and leverage enormous and granular datasets.

The "more traditional" analyses typically rely on monthly or weekly SKU-level price and volume data and estimate how a product's (SKU) sales quantities historically changed when its price or the price of other products changed, while controlling for all other relevant factors that changed in the interim.

The challenge – which can prove fatal – is ensuring that the estimates are not affected by "endogeneity," which means that some important controls have been omitted.[11] However, controlling for endogeneity can be very complex and is often unfeasible. Again, **Big Data techniques can assist.**

- **First, controlling for hundreds, thousands or even millions of variables, reduces the risk of omitting important factors.** For example, Big Tech companies would estimate demand elasticity by comparing the purchase behavior of groups of "otherwise identical" individuals, one of which experienced a price increase. And would use a large number of controls to ensure that individuals are truly "identical" other than for experiencing a price increase.

- **Second, when data is abundant, it is possible to trade quantity for quality.** For example, Big Tech companies collect price data indicating the exact second in which a price change occurred. These data can be used to compare purchases made just a few seconds before and a few seconds after a product price change. As customers are unaware of the exact time in which prices change, they are effectively assigned at random into the group that did not experience a price change (purchased before the price

[11] For example, one might observe that prices increased by 10% between two periods and that quantities instead of declining also increased by 10%: does this mean that increasing prices increases demand? The trick is in the reason why prices changed in the first place: was it in response to a demand increase (e.g., the products are sunglasses and summer is approaching), or a supply change (e.g., costs went up). When estimating demand, it is important to include controls that identify price changes caused by changes in supply (e.g., price increases), not demand shocks. Omitting these important controls could bias estimates and even turn them to the wrong sign, as in this example.

change) and the group that did (purchased immediately after). Clearly, looking at purchases immediately before and after a price change means dropping the vast majority of the observations in the dataset, which is unthinkable with a "normally" sized dataset, but is possible with Big Data. And what this gives back is a "quasi-experimental" design (called a "regression discontinuity" design), where customers are randomly assigned to the control and treatment groups, and therefore (in aggregate) are "otherwise identical."

- **Third, more granular data allow for an estimation of very "localized" effects.** Sometimes demand elasticity is estimated using scanner data aggregated across many stores as well as aggregating groups of similar products together under the same SKU. Using store- and SKU-level price, volume, and revenue data opens the door for more "localized" estimates of demand elasticities for each product of interest in each individual store separately. As mentioned above, it is possible to regress these very localized estimates against a series of variables of interest, e.g., to test how demand elasticity varies with distance from the other merging party/third party rivals, presence of particular customer types etc.

Category 3: Merger simulation.

Demand elasticity estimates can be plugged into theoretical models to estimate the likely price effect of the merger. The simple models (e.g., GUPPI) combine diversion ratios with estimates of gross margins. More sophisticated models can better account for demand elasticity and subtler price effects (e.g., competitors' response to the merging parties' price increase).

Big Data techniques can better account for demand elasticity and subtler price effects of a merger by embedding the sophisticated pricing algorithms companies use into merger simulation models.

The challenge for these models is to provide a realistic representation of the process through which companies set prices and account for all the relevant factors that affect prices. Often, these models provide a very simplistic, static, and short-term view of how firms set prices. **Big Data techniques can assist.**

Large corporates (well beyond Big Tech) increasingly rely on sophisticated pricing algorithms. For example, a company like Walmart may use an algorithm that crawls information on other chains' and digital platforms' prices for similar products and then update its own prices when it observes an impact on its own demand. And rivals often deploy similar algorithms too. So, there may be a dozen algorithms crawling each other's price data and responding to each other. These algorithms provide precious insights into how firms set prices an react to each other depending on the underlying parameters that the firms themselves consider in their ordinary course of business. To reflect the "business perspective," a merger simulation should therefore embed these algorithms into the model.

In practice, this could mean estimating how Walmart's prices respond to its' rivals' actions (price changes, launches of new products, etc.) and how other rivals respond to Walmart's – i.e., their best-response functions. Big Tech companies use "hedonic pricing" models for this purpose.[12] Hedonic pricing is a near century old technique which consists in regressing product prices against product attributes, to estimate how much value consumers attach to each attribute. To estimate Walmart's best-response function, we could regress Walmart's own prices against variables that capture rivals' actions (e.g., Amazon changing prices or launching new products), as well as product attributes. And would capture product attributes through ML models: e.g., neural networks, computer visions on all the pictures of a product on a web page, natural language modelling on all the words describing the product in reviews and its characteristics on a webpage etc.. ML is very effective because it captures "soft" product attributes, such as a slick product design, which are otherwise difficult to encode.[13]

Category 4: Merger efficiencies.

Efficiencies need to be verifiable and quantifiable.[14] Firms increasingly rely on sophisticated supply chain models and algorithms, for example to predict demand and manage the supply chain accordingly. Through a detailed review of these algorithms, it is possible to describe the mechanism through which mergers can reduce marginal costs in the supply chain and accurately measure the magnitude of such reduction.

For example, a merger between retailers may generate scale economies which translate in lower marginal costs, not only in fixed cost reductions. When a product's demand becomes too small, retailers have to hold inventory for longer and transportation costs increase (e.g., it is no longer possible to pool deliveries). This is an issue that e-commerce platforms such as Amazon or eBay face and have perfected tools to measure and manage them. For example, they estimate the effect on transportation and other costs when demand varies, calling through features of the supply chain etc. through double-machine-learning models so that it is possible to localize the effects on particular types or groups of products and/or customers. These techniques can be used to estimate merger efficiencies, as we can quantify how transportation (or other) costs decline when combining the merging parties' sales.

Group 2: Taking it to the next level: measuring substitution the way Big Tech companies do

The first three techniques described in the previous section are fundamentally based on the same idea: if products A and B are substitutes, then if A becomes available or lowers its price, then B's sales go down (and *vice versa*). So, to

[12] See, for example, Bajari, P., V. Chernozhukov, A. Hortaçsu and J. Suzuki (2018) *"The Impact of Big Data on Firm Performance: an Empirical Investigation"*, NBER Working Paper No. 24334 (available at: www.nber.org/system/files/working_papers/w24334/w24334.pdf).

[13] Clearly, this approach is subject to selection bias. Walmart may lower the price of its products because they are not selling much, regardless of Amazon's price changes, so one challenge is to distinguish a price reduction in response to a price change from Amazon, from a price reduction that would have happened anyway. As above, Big Tech companies would "trade quantity for quality", looking at Amazon's prices a few seconds before and a few seconds after Walmart's price changes. Algorithmic responses are very fast, so they are likely to be captured even in a narrow time window. Whereas other price changes are unlikely to occur in that precise window.

[14] Section 3 of the FTC/DOJ 2023 Merger Guidelines. Merger Guidelines [2023] - U.S. Department of Justice and the Federal Trade Commission (ftc.gov).

measure substitutability, we look at the impact on B's sales of A becoming available (entry) or increasing its price. But this is not the way Big Tech companies look at substitutability. Big Tech companies tend to measure substitutability by identifying regularities in the pattern of consumers' behaviors through ML algorithms. Here we summarize some of the key techniques that they commonly use.

- **Search rankings**. Big Tech companies identify product substitutes using product rankings. For example, imagine searching for "men shoes" on Walmart.com. One product will feature as #1 in search results. But then, after some time, Walmart would refresh the rankings and the same product may appear, e.g., as #3. Looking at customers who viewed/purchased products immediately before and immediately after the change in ranking, we can see whether demand changed significantly when the product moved from #1 to #3, as many customers switched to the new #1 (which is consistent with the new #1 being a close substitute) or customers followed through and continued buying the same product in #3 position (which indicates a certain degree of differentiation).

- **ML techniques to profile customer groups**. Agencies typically draw the lines for customer segmentations based on qualitative criteria, such as customer or product characteristics, merging parties' internal documents, interviews and industry reports. But customer segmentations can also be tested quantitatively through supervised and/or unsupervised ML techniques[15] applied to granular customer data. These techniques can detect and measure patterns of user characteristics, purchase behaviors that are often not easily observable and cannot otherwise be measured. For example, unsupervised ML can be used to categorize product reviews into certain customer types based on patterns in the data. ML algorithms are much more effective than standard econometric techniques (which require some human judgement in identifying product groups) at spotting relevant patterns and are not constrained to find linear patterns only.[16]

- **Merger effect on diverse customer groups**. When prices, quality, and innovation can be set differently for different customer groups, there is scope for a merger to impact different customer groups differently. It is possible to employ AI tools to reliably estimate the different effects of a deal on different customer groups. By applying AI tools to granular consumption data, for example, all the texts and pictures in reviews of all the products a customer has purchased online, it is possible to first identify diverse customer segments and then predict the differential impact of a merger (both in terms of harm and of benefits) on those customer segments.

[15] Unsupervised Machine Learning involves finding clusters of observations that are similar in terms of their covariates whereas supervised Machine Learning entails using a set of features or covariates to predict an outcome using some labelled observations where both covariates and outcomes are observed (the training data) to predict outcomes in a dataset with covariates but without outcomes (the test data).
[16] See also Susan Athey, The Impact of Machine Learning on Economics, May 2019 The Economics of Artificial Intelligence: An Agenda (nber.org).

Conclusion

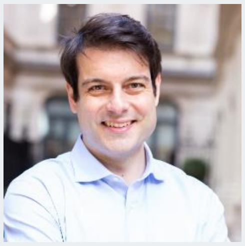
In conclusion, we see a trend towards embedding business evidence and, altogether, a business “perspective” in the assessment of mergers. To persuade competition agencies using data analysis, economic advisors will increasingly need to rely on sophisticated, Big-Tech-like quantitative analyses which provide a more realistic view of how businesses look at substitutability in their ordinary course of business.

About the Authors

Patick Bajari

Patrick Bajari is a Senior Partner at Keystone Strategy where he is the Chief Economist and Managing Director for Core AI. Prior to Keystone, Patrick worked as a technology VP for Amazon reporting directly to the CEO of the Consumer focused on helping Amazon build data driven solutions at scale in problems such as pricing, marketing, inventory, transportation, and forecasting. Patrick holds a doctorate in economics from the University of Minnesota and has held faculty positions in economics at Harvard, Stanford, Duke, and Minnesota. He is a visiting researcher at Harvard Business School and the economics department at the University of Minnesota. He has been recognized with a number of prestigious professional awards, including being named a fellow of the Econometric Society and a Fellow of the National Association of Business Economists.

Gianmarco Calanchi



Gianmarco Calanchi is a Senior Principal in the London office of Keystone Strategy. He has more than ten years' experience as a professional economist. He has advised clients in high profile competition and antitrust cases, including merger control cases, Article 101 and 102 cases and private litigations before the European Commission, national competition authorities and tribunals. His experience spans industries including financial markets, media, telecommunications, transport, energy, healthcare, pharmaceuticals, and payment systems.

Before joining Keystone, Gianmarco was Vice President at Compass Lexecon and led the economic analysis of over 20 horizontal and non-horizontal mergers at the UK Competition and Markets Authority (and at its predecessor, the Office of Fair Trading).

Gianmarco holds an M.Sc. and a B.Sc. in Economics and Social Sciences from Bocconi University, with BSC dissertation in statistics and MSC dissertation in game theory applied to competition policy.

Tega Akati-Udi

Tega is a Consultant II in Keystone Strategy's London office. Prior to joining Keystone, she worked at Frontier Economics, primarily in competition but also in aviation and strategy. She has worked on a range of cases including mergers, litigation and market studies within the UK and across Europe. She holds an MSc in Economics from Tilburg University and a BA in Economics from the University of Cambridge. Tega is passionate about promoting diversity within the field of economics and is a volunteer with The Black Economics Network.

Outside of work, you'll find Tega reading fiction, scouting out the latest musical in London, or training for a half marathon.

About Keystone Strategy

Keystone Strategy is a leading innovative strategy and economics consulting firm dedicated to delivering transformative ideas and cutting-edge solutions to Fortune Global 500 companies, top law firms, and government agencies. Our unique expertise in technology-led strategy, AI-driven digital transformation, product development, healthcare, finance, antitrust, and litigation enables us to create bold strategies that have far-reaching implications on business, consumers, and public policy. We leverage a unique combination of strategic insights from the world's leading industry and academic experts with the practical expertise of our accomplished professionals to deliver extraordinary impact for our clients. Learn more about Keystone Strategy at: www.keystone.ai