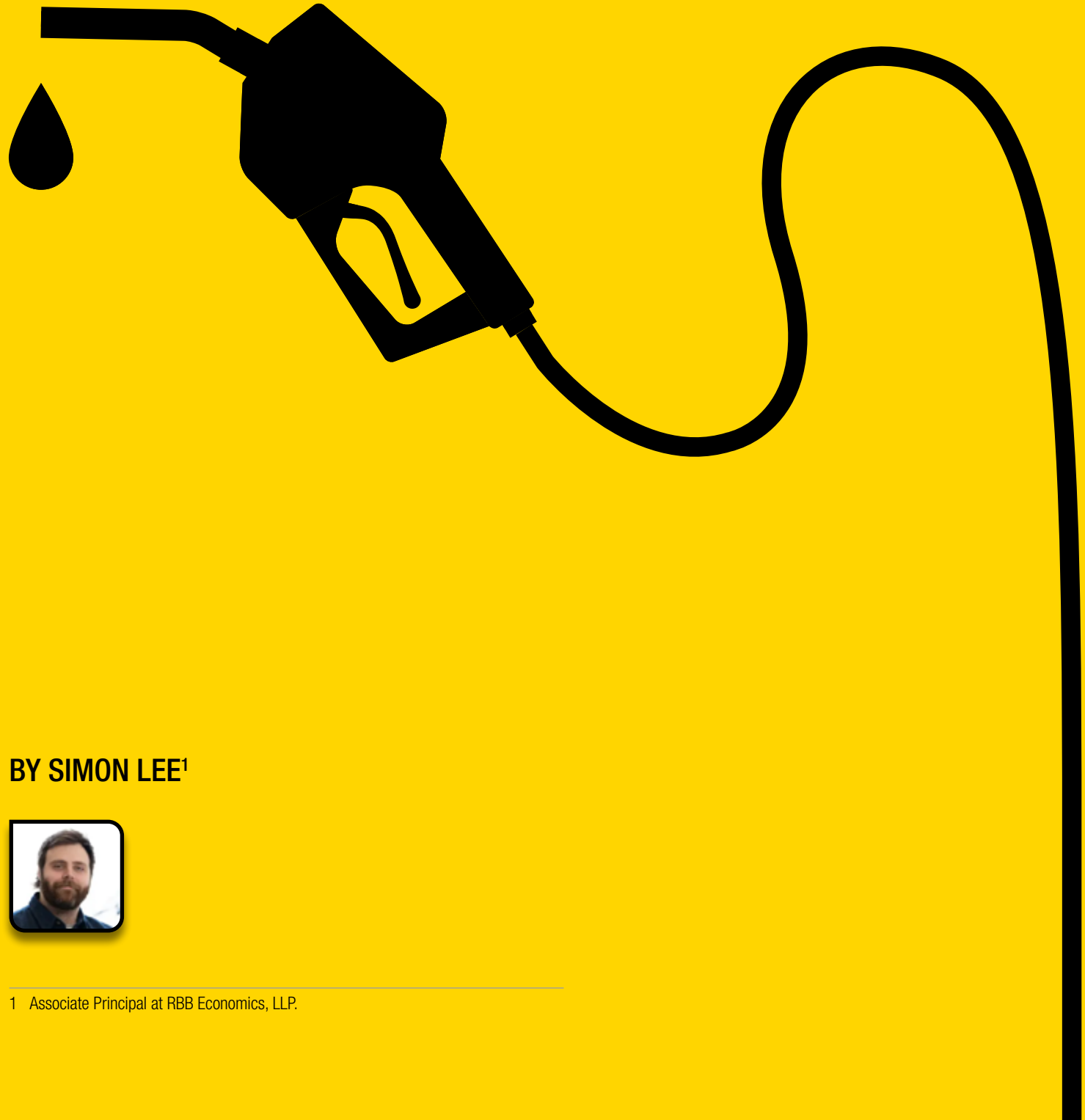
# WHEN DOES MORE DATA MEAN MORE POWER?



BY SIMON LEE[1]



1 Associate Principal at RBB Economics, LLP.

# CPI ANTITRUST CHRONICLE March 2026



Visit www.competitionpolicyinternational.com for access to these articles and more!



## WHEN DOES MORE DATA MEAN MORE POWER?

*By Simon Lee*


This article examines the growing role of data in competition analysis and cautions against the simplistic assumption that more data necessarily translates into greater market power. The author explains that data can generate competitive advantages by improving product performance, enabling more precise targeting and monetization, strengthening customer retention through personalization, and reinforcing network effects. At the same time, the author argues that these advantages do not automatically translate into anticompetitive outcomes. The article concludes that competition authorities should resist presumptions of "data dominance" and instead assess whether data advantages are sufficiently unique and difficult to replicate that they meaningfully prevent rivals from competing effectively.

# I. INTRODUCTION

As data-driven business models proliferate in digital and other markets, competition authorities (in Africa and beyond) have grown increasingly attentive to the role that data may play in shaping market power. While it is broadly accepted that large-scale datasets can offer competitive advantages, the critical question for competition enforcement is more nuanced: under what conditions might the possession of data, or the accumulation of data (e.g. through merger activity), actually pose a significant risk to competition?

This article provides an economic framework for distinguishing between cases where the possession or accumulation of data might meaningfully pose competition concerns, *versus* those where potential fears of "data dominance" are likely to be overstated. It is structured as follows.

· Section II outlines the key ways in which data can confer competitive advantages, and summarises a handful of recent South African and regional matters in which data-related concerns have featured.

- Section III provides a set of economic considerations that may assist in determining whether data are likely to raise competition concerns.
- Section IV concludes by reflecting on the implications for competition enforcement in digital and other data-rich markets.

# II. WHEN CAN DATA BE A SOURCE OF COMPETITIVE ADVANTAGE?

In many modern markets (particularly digital ones), data can be a key driver of competitive advantages. For instance, firms that can collect, analyse, and act on large volumes of data may be able to improve the performance of their products, personalise their services, target customers more effectively, and/or optimise their pricing and operations.

While these types of effects might normally be expected to translate into better outcomes for consumers (at least in the short term), there may be a concern that the possession or accumulation of data might result in such significant competitive advantages that rivals are unable to keep pace – potentially leading to their foreclosure and, over time, a weakening of competition overall.

This section sets out some of the ways in which data can strengthen a firm's market position, and outlines examples of where such concerns have featured in recent competition assessments and policy decisions.

### *A. Product Improvement*

Access to large, diverse datasets can improve the performance of models that make use of data, especially in highly automated markets where algorithmic performance can be a key differentiator between rivals. For example, data may allow a firm to refine its responses, enhance its recognition and understanding of input requests, better match input requests to suitable product and content outputs, improve fraud detection, and/or automate customer service with greater accuracy or lower unit costs. Accordingly, the possession or accumulation of substantial data assets may allow firms to train more accurate or adaptive models than their competitors would be able to closely replicate (without incurring significant additional costs).

### *B. Monetization and Targeting*

Detailed user data may enable firms to segment and target customers more precisely, which might then translate into higher click-through rates, better conversion outcomes, and increased sales or ad revenues. For example, a firm that can track users across multiple services or platforms, and which can combine behavioural, demographic, and transactional data, may be able to achieve superior targeting capabilities (or to offer these capabilities to third parties, such as advertisers). Therefore, the possession or accumulation of such data assets may enhance a firm's position when it comes to digital targeting, particularly where alternatives lack comparable reach or granularity.

### *C. Customer Retention*

Firms with access to detailed historical data on their users may be able to personalise products and interfaces in a way that increases engagement and loyalty (e.g. personalised playlists, content feeds, search results, and product recommendations typically depend on prior usage data). As a result, the more of such data that a firm holds, the better it may be at anticipating user preferences and creating a degree of "learning-based" lock-in. The possession or accumulation of multiple sets of behavioural data may therefore enable firms to deliver more tailored experiences than their rivals, thereby raising switching costs for users and making it more difficult for competitors to contest effectively for customers.

### *D. Network Effects*

In some markets, serving more users leads to more data. More data might then enable firms to provide more compelling offerings (e.g. through the mechanisms discussed above), which might then attract more users, thereby resulting in the accumulation of even more data. The possession or combination of significant data may thus enhance these self-reinforcing feedback loops, which may in turn exacerbate or entrench the mechanisms outlined above. If rivals find it difficult to replicate these effects, or to respond to them in other ways, such data-driven advantages may entrench dominance and raise barriers to entry / expansion.

### *E. Examples in Context*

Several recent South African and regional matters and policy decisions illustrate the types of concerns that may arise in the context of markets where data (may) play a prominent role.

For example:

- In the *Vodacom–Maziv* merger, the South African Competition Tribunal initially prohibited Vodacom's proposed acquisition of a controlling stake in fibre operator Maziv. Among the concerns raised was the possibility that the merger would have enabled Vodacom to integrate mobile and fibre usage data, thereby giving it a competitive advantage in terms of customer targeting, service bundling, and network optimisation. According to the stated concerns, access to a consolidated pool of infrastructure-level and customer-specific data could have enabled Vodacom to foreclose smaller internet service providers (in particular by utilising data on customers' use of pre-existing mobile services to target similar customers with new fibre services), or to offer bespoke services that would have been difficult for rivals to replicate.[2]
- In *WeBuyCars*, the South African Competition Commission ("the Commission") expressed a concern that the data collected by the AutoTrader platform (e.g. data relating to buyer and seller behaviour, vehicle searches, pricing trends, and regional demand patterns) could have been leveraged to give WeBuyCars an advantage. More specifically, the Commission posited that such data could have been used to optimise WeBuyCars' pricing and operations, thereby risking the foreclosure of competing buyers and sellers of used vehicles (which ostensibly would not have had access to the same or substantively similar data).
- In the Online Intermediation Platforms Market Inquiry, the Commission raised concerns that online intermediation platforms (including e-commerce marketplaces, classifieds websites, travel aggregators, and app stores) utilise their access to user and transaction data to reinforce their market power. Specifically, it noted that these platforms may leverage data advantages to prioritise their own listings, optimise pricing or ad placement, and limit the visibility or performance of rivals and third-party sellers. The Commission also highlighted how the asymmetry in access to data between platforms and their users or business partners can entrench incumbency and reduce competition (both between users of platforms, and between platforms themselves).
- At the regional level, the COMESA Competition Commission has increasingly recognised the competitive relevance of data in digital mergers, particularly where platforms operate across multiple member states. For instance, COMESA's evolving regulatory framework explicitly identifies control over data, data-driven network effects, and user scale as key analytical considerations.

These examples illustrate how competition authorities are applying increased scrutiny to the potential for data to act as a source of competitive strength – not merely through scale, but *via* their use in pricing, targeting, and service optimisation. While potential outcomes will differ across cases and contexts, the broad underlying concern is that the possession or accumulation of data may give rise to advantages that are difficult for rivals to overcome, and that such effects may ultimately marginalise competitors to the detriment of the competitive process.

However, as noted above, data can enable firms to improve their competitive offerings to the benefit of competition and consumers (at least in the short term). As such, the relevant question relates to when, and under what conditions, the possession or accumulation of data should be seen as a credible and realistic cause for concern.

## III. WHEN ARE DATA UNLIKELY TO RAISE COMPETITION CONCERNS?

While data can be a source of competitive advantage, it does not follow that the possession or accumulation of data will necessarily be detrimental to competition. Indeed, in many cases the possession or accumulation of data is unlikely to be problematic, either because the data in question are

2 See, for example, the South African Competition Tribunal's reference to "information exchange concerns" in its press release, available at: https://www.comptrib.co.za/info-library/case-press-releases/tribunal-releases-its-comprehensive-reasons-for-prohibiting-the-vodacom-maziv-transaction [accessed 19 June 2025].

not especially (incrementally) valuable / useful, or because rivals remain able to compete effectively despite not (currently) having access to the same or even substantively similar data (in which case, as noted above, competition and consumer outcomes are likely to be enhanced by data).

Accordingly, the key challenge for competition authorities is to distinguish between those cases that are likely to have a material (adverse) effect on competition, and those that are not. This requires an effects-based approach that goes beyond simply looking at whether a firm possesses significant data, or would accumulate significant data (e.g. *via* a merger). Instead, such an assessment ought to consider whether the data in question would confer such significant competitive advantages that rivals would be unable to effectively compete, in turn resulting in a situation where the firm in question would be able to worsen its own competitive offering to the detriment of consumers.

The considerations below outline the types of scenarios where the possession or accumulation of data assets would be unlikely to give rise to sufficiently large competitive advantages for competition concerns to be warranted.

### *A. Substitutability and Access*

The competitive relevance of data depends on whether other firms can obtain substantively similar data through reasonable means. For instance, rivals might already possess such data, or may be able to access datasets of similar utility through their own operations, through partnerships, through commercial procurement, or through open-source alternatives. In turn, if the data in question are not exclusive or are not particularly differentiated in ways that matter for competition, then the possession or accumulation of such data would be unlikely to raise concerns about the creation of an insurmountable competitive advantage for the firm in question.

### *B. The Role of Data Analytics Capabilities*

Critically, data are only likely to carry competitive advantages if a firm has the capacity to extract and act on their insights. This typically requires skilled data scientists, strong engineering infrastructure, and appropriate machine learning tools. As such, even if a firm is able to accumulate a more extensive pool of data than other firms, rivals with stronger analytics capabilities may be able to derive more value from their smaller datasets, or even catch up over time by investing in the right tools and expertise. Therefore, competition assessments should consider not just the relative size and richness of datasets, but also relative abilities to operationalise and extract competitive value from data.

### *C. Dependence on Third-party Infrastructure*

In some markets, the ability to extract value from data is mediated by infrastructure or platforms that are controlled by third parties. For instance, in digital advertising, firms often rely on platforms like Google Ads to deliver targeted advertising (i.e., these third-party platforms use their own systems and algorithms to decide how ads are served and priced). As a result, even if a firm has access to significant data, it may have limited control over how those data translate into commercial outcomes. In such cases, the possession or accumulation of data may not meaningfully create or reinforce market power if the main tools for monetising data lie outside the control of the firm in question.

### *D. Marginal Returns to Data*

The incremental benefit from adding data to an existing dataset may decline beyond a certain point. For instance, while initial data inputs can significantly improve the performance of many machine learning and other calibrated applications, further additions beyond a certain threshold may yield only modest gains to the overall precision and quality of the models or applications in question. Where such a threshold has already been met, or where rivals have already surpassed that threshold, the possession or accumulation of even more data (regardless of volume) would be unlikely to materially enhance the capabilities of the firm in question, or to result in a situation where rivals are significantly marginalised.

### *E. Competitive Significance*

Some datasets may be more competitively significant than others. For instance, rich, granular, real-time behavioural data may enable more precise targeting or predictions, whereas generic, historical, or aggregated data may have more limited strategic use. Accordingly, competition assessments should avoid assuming that all data are competitively meaningful, and should instead evaluate the specific utility of data, and their relevance to competition, on a case-by-case basis.

### *F. Other Sources of Competitive Advantage*

While data can be a strategic asset, data are unlikely to be the only lever that firms have at their disposal. In practice, competitors may offset a data-related disadvantage through strengths in other areas (such as product innovation, customer loyalty, pricing, logistics, or brand recognition). Accordingly, a firm's lead in respect of data may not be decisive if rivals are able to compete effectively on these other fronts. It is therefore

important to evaluate data in the broader context of market dynamics, and not to assume that the possession or accumulation of data alone implies an unassailable market position.

### *G. Key Takeaway*

With the above considerations in mind, ultimately the core question from a competition enforcement perspective is whether data advantages are so significant, and so difficult to replicate or offset, that they would prevent rivals from competing effectively. If rivals are able to respond (for instance by accessing alternative data assets, improving their own analytics capabilities, or competing on dimensions other than data), then the possession or accumulation of data would be less likely to distort the competitive process. To the contrary, in such circumstances, data-driven improvements would be likely to improve consumer outcomes through the generation of new and better products and/or services, which may then spark competitive reactions from rivals. In these situations, the resulting effects are likely to be procompetitive rather than harmful.

## IV. CONCLUDING REMARKS

Given the range of ways in which data can create competitive advantages, authorities should approach data-related investigations with analytical discipline, but also with care to avoid over-enforcement. A presumption that "more data equals more market power" risks misdiagnosing competitively neutral or beneficial market conduct and mergers. Indeed, even where a firm possesses or gains access to significant data, it remains important to assess whether this meaningfully affects competitive dynamics in a way that might lead to the foreclosure of rivals and the consequent softening of competition.

Put differently, while it is intuitively appealing to treat data as a source of market power, it does not necessarily follow that all instances of data possession or accumulation will give rise to competition issues. What matters is not simply how much data a firm has, but how useful, unique, and replicable those data are, and whether they materially enhance the firm's ability to compete in the market, or to prevent others from doing so. In many cases, rivals may be able to respond effectively through alternative datasets, differentiated strategies (including in respect of other competitive parameters), or by leveraging third-party platforms. In such scenarios, the possession or accumulation of data may in fact intensify rivalry by prompting innovation, improving product performance, and/or delivering efficiencies that ultimately benefit consumers.

As data become increasingly relevant to how firms compete with each other in many markets, the challenge for antitrust assessments and merger control will be to distinguish genuinely harmful cases from those that are competitively neutral (or potentially procompetitive). Meeting that challenge requires an economically rigorous, context-specific, and effects-based approach, grounded in an understanding of how data actually drive competitive outcomes.

CPI COMPETITION POLICY® INTERNATIONAL

# CPI Subscriptions

CPI reaches more than 35,000 readers in over 150 countries every day. Our online library houses over 23,000 papers, articles and interviews.

Visit competitionpolicyinternational.com today to see our available plans and join CPI's global community of antitrust experts.