Bloomberg Tax

Tax Management
International Journal™

Data, User Participation and Value Creation: Drilling Down on Taxing Digitalization

By Shan Sun[*]

International tax is struggling to keep up with the digitalization of modern economics. Rules and principles developed over a century ago seem unsuitable to tax today's economy fairly and effectively. At the center of the discussion is the role of data and user participation in value creation. This article explores what it means for a renewed international taxation framework to adapt to the digital economy. It first explores the concept of value creation to ensure it captures the nature of the digital economy. It illustrates with specific examples when it may be considered suitable to take data and user participation into account for international tax. The article emphasises on the importance of understanding the underlying technology of digital economy for a fair and effective international taxation framework. It also explores two scenarios of tax base allocation by focusing on user participation and market data access respectively.

PRELUDE: FUELING AN ECONOMY

Data is fueling this century's economy as oil did the last one's, driving change and growth in everyday life. In the world of international taxation, the key question to ask becomes whether technology empowered by data fundamentally changes the formula of the foundation to reconsider the international taxation framework.

The complexity of tax challenges from the digital economy[1] almost resembles that of the three-body problem in physics — using underlying information on position, mass, and velocity at some specific point in time to determine the motion of three bodies under commonly recognized laws. As an analogy, it may be helpful to briefly revisit the origin of the underlying theory of taxation before diving into the deep debates on the intricacies of taxation on the digital economy.

Taxation at its very beginning evolved from the need for income redistribution[2] by central governments to fund public services, infrastructure, and the aftermath of war. Two elements remain pivotal to the considerations of taxation: labor and capital, both tangible and financial.[3] Overlaying these elements over time and location leads to economic and productivity growth, and thereby creates value. It therefore also provides a foundation to consider international taxation frameworks. The question is: Where does data fit in this equation? It is neither labor nor capital. It is extracted, refined, valued, and traded in many ways

[*]The author works as a global transfer pricing senior manager at Omnicom. Opinions and any errors are the author's alone and do not represent any organization's view irrespective of affiliation. Any comments or feedback, please direct to the author at shanxsun@gmail.com.

[1] Specifically, in writing this article, the main comments are focused on the OECD's *Tax Challenges Arising from Digitalisation — Interim Report 2018* (the "OECD paper") issued on March 16, 2018, with some overlapping issues from other papers issued by the European Union and the United Kingdom (see References, below).

[2] Ultimate problem of international jurisprudence, James Lorimer, 1884. *See* Wilfred C. Jenks, *The Significance Today of Lorimer's Ultimate Problem of International Jurisprudence* (1940); Myron J. Frankman, *International Taxation: The Trajectory of an Idea From Lorimer to Brandt* (1996, 2000).

[3] Aside from labor, the early literature of international taxation predominantly focused on the revenue capacity of each nation in considering taxation. *See* Wilfred C. Jenks, *The Significance Today of Lorimer's Ultimate Problem of International Jurisprudence* (1940); Wilfred C. Jenks, *Some Legal Aspects of the Financing of International Institutions* (1942). The assumption here is that there is a direct causal relationship between revenue and labor and capital from the productivity growth theories in economics. *See* Robert M. Solow, *A Contribution to the Theory of Economic Growth* (1956); Trevor W. Swan, *Economic Growth and Capital Accumulation* (1956). I further dissected capital into tangible capital (e.g., plant and machinery) and financial capital (e.g., funding) for ease of discussion to a modern setting in tax.

ISSN 0090-4600

never seen by the world economy before, creating new monopolies and business dynamics. It changes rules for markets and business, demanding new approaches from regulators. Getting it wrong would be disastrous for the world economy, so the decision should not be taken lightly.

Value Creation 2.0

To answer the aforementioned question, it is important to look deeper into what value creation really entails in the digital economy. It is important because the concept of value creation is widely accepted by many tax jurisdictions in assessing their respective tax base, yet no definitive consensus is discussed and agreed upon. Referring to Porter's view (Porter, 1985), the OECD paper stated the two factors that lead to the creation of superior business value: increasing the price of goods, and reducing costs (para. 66). Furthermore, it classified companies into three business operation structures or so-called ''value creation processes,'' depending on the level of digitalization: value chains, value networks, and value shops (para. 71). This article extends the concept further and proposes three possible ways to create value. To be more precise, irrespective of how business operates (with various degree of digitalization), value is created if there is:

(1) an increase in current business profit (essentially in alignment with Porter);

(2) an emergence of competitive advantage over rivals; and/or

(3) an increase in future business growth potential.

The three sources of value creation are by no means exhaustive nor are they mutually exclusive. Contrarily, they are often intertwined; or sometimes even inseparable. It is especially true given the important role research and development activities plays in creating intangible properties, which underpin much of the technological advancement today. The three facets of value creation may appear redundant at the first glance from an economist's perspective. They may be considered the same thing in essence with only temporal difference. However, in the context of the digital economy, it remains critical to make such distinctions. Finance and economics literature has long assumed that maximizing shareholder return — i.e., profit performance — is the goal of corporations. However, this theoretical assumption is no longer fitting for many start-up companies, especially in the technology industry, where often it is the growth potential in near short to medium term, instead of profitability, is emphasized as the key factor to achieve funding and ongoing operation. In fact, in today's economy, there are established businesses with low or even no prior profit. Often the nature of such companies (e.g., network effect) determines that the most important business strategy is to achieve scale and stay ahead of competition. To say such companies have not yet created value, because they are not yet profitable or accumulated substantial profit in the traditional economic sense, would be off the mark by a long way — rendering an inaccurate, even incorrect corresponding tax base. In other words, the seemingly trivial temporal differences between the three concepts of value creation listed above are vital to the survival and operations of many businesses today. The three concepts — summarized as the increased profit, competitive advantage, and growth potential — interpret what value creation entails in the digital economy. They hence should be considered separately and thoroughly for each business to arrive at a fair and correct tax base.

The extension of the value creation concept was made economically possible by many attributes, from hardware (e.g., availability of fast and affordable processing chips) to software and human knowledge (e.g., algorithms). Central to this change is the arrival of the behemoth amount of data collection. Data has become so readily available and abundant that formulating complex algorithms and robotic automation processes are viable, to the point that frontier researchers have begun calling data the ''new electricity.'' An up-to-date international taxation framework under this context begs the question: Does data play such an important role in value creation that it has fundamentally altered the previous core structure of taxation architected around labor and capital over time and location? It is also particularly relevant when attempting to understand the role of data and user participation in value creation, as illustrated in further detail below.

The Role of Data and User Participation

Answering this question is no simple task. In the face of differing country views and the often contentious state of international relations, governments around the world have not reached any consensus. One of the questions at the heart of the disagreement relates to the role of data and user participation in value creation, and thereafter the appropriate allocation of the tax base. In consideration of the extended interpretation of value creation proposed above, this article provides some concrete examples to illustrate why there is no straightforward one-size-fits-all answer. Rather, the solution is very much dependent on the details of how each business within the digital economy works, with a thorough understanding of the underlying technology upon which the business relies and the channels in which it creates value.

Starting with a quick thought-provoking example, let us look at the case of the automatic vehicle (AV) business. AVs may be built on simulations — i.e., fabricated data extracted from separately constructed traffic infrastructure in physical locations away from real life, on 3D computer modules. Alternatively, AVs may also be built and tested on habitual data of real user drivers in a particular country, city, or town (think of a cab driver on winding London roads in the United Kingdom versus a newly qualified driver on

ISSN 0090-4600

multi-lane roundabout in Austin in the United States). A quick review of the literature could distinguish various present approaches in an attempt to achieve the reality of AV very differently; for example, a short summary includes:

- building complete simulations in 3D on computers and run virtual miles;
- building road and city infrastructure based on real-world places and run physical tests on these fabricated structures in real life;
- testing simulations based on a much smaller real-world driving data; or
- deploying assisted-driving mode directly into real cars to collect real-world data first-hand.

It should be clear from the scenarios described above, that taxing companies who built their automatic driving system based on real user input data the same way as those who built on data from artificially simulated environment, as if they are the same, is inherently incorrect. Blindly continuing along this line, by charging a universal percentage on revenue, for instance, will ultimately result in unfair treatment to different businesses, hurting competition, or even unintentionally leading to monopoly. So what could be done to avoid such generalization from an international taxation perspective?

Let us look at another example to understand a little further what it really means to take account of data and user participation by gauging the extended concept of value creation discussed above. We illustrate with a relatively straightforward business example: music streaming. This type of business is classified into the ''reseller'' group by the OECD paper as a business that essentially trades content (para. 58). However, from a technical perspective, these companies are likely to be considered as an oversimplified combination of a traditional music track ''warehouse'' and a recommendation system. On a closer observation, this type of business can be dissected in greater detail using the extended interpretation of value creation articulated above.

(1) Value Created Through Increase in Business Profits (or Revenue)

The business purchases and controls rights from vendors (e.g., artists) and resells them to consumers. Arguably the more music tracks the business manages to own, the more likely a consumer would choose it as their primary source for music due to its abundancy of inventory (i.e., song tracks). The more the inventory, the more easily the business attracts customers and takes advantage of scale. In this sense, the so-called ''user participation,'' in my view, is no different from any traditional business in the sense that the more variety in tangible goods a business stocks, the more likely it will attract more customers and establish itself as a viable going concern.

(2) Value Created Through Achieving Competitive Advantages

The situation becomes more interesting when the music streaming industry has several competing players. Which business will a customer choose to use? For this particular type of business and in their current status, I believe the resulting choice largely depends on the quality of the business's music recommendation ecosystems. Specifically, consumers choose to stay loyal to one particular streaming business because of its offerings in music recommendations, playlists, and other tailored services. It is a classic example of business transitioning from goods only to goods and services or beyond. It merits further exploration to understand what it means from the perspective of data and user participation.

For a streaming business to know what to recommend requires data from various sources to build algorithms or models. Some main models used in the recommendation ecosystems for music and other content streaming business include the following:

a. Collaborative filtering model. The streaming service needs to have user behavior data — such as user preferences, length of time engaged with a track, savings to library, etc. It would also require data on user interactions (via shares, collaborative playlist, follows, etc.). This data is then numerically vectorized for machine processing and comparison to discover what music a user might potentially enjoy based on users' past behaviors. Recommendations are made by minimizing the differences between the numbers representing the acquired user data. This type of data is very powerful, highly dependent on user-specific data which is proprietary to the streaming business.

b. Natural language processing (NLP) model. To stay relevant and lead the trend in music, the streaming business crawls everywhere on the Internet for relevant content, which might be a blog about a concert, a tweet from an artist, or some critic on a specific music genre. Afterwards, NLP algorithms are used to overlay and fine-tune the recommendation heuristically at a macro level. In essence no one can recommend good music without a deep understanding of the industry. The data used here is necessary for a reasonable recommendation system but not proprietary to the streaming business, nor is it user specific.

c. Acoustic analysis model. The third type of model used by a streaming service for recommendation is a result of pure engineering on audio files, borrowing ideas from machine learning in image recognition — convolutional neural network (CNN). The acoustic analysis model addresses the issue where new, unknown artists and their specific work (such as a single release) is not being picked up by the previous two types of models. CNN recognizes certain characteristics of a song being parsed through, which are then classified into groups based on different characteristics such as tempo, loudness, key, mode, and even danceability. This

ISSN 0090-4600

way one would not receive a recommendation of a heavy metal track as music for sleep. In this scenario, although the data — i.e., the music itself — is not proprietary or user specific, and the associated rights belong to the streaming business. The analysis of the audio files is extremely subjective and dictated by engineering alone (from the construction of CNN to the interpretation of the song characteristics). It is where software engineers and domain experts really shine and add value.

The description of the above algorithms used in the streaming business illustrates that one way the business creates value is to distinguish itself from competitors by creating a superior recommendation ecosystem. Note this is different from the value creation via the sheer volume of copy rights inventory described above.

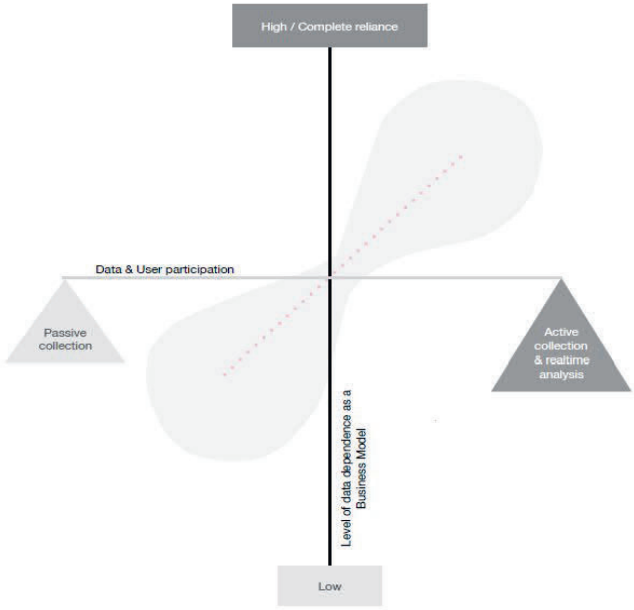
(3) Value Created Through Obtaining Future Growth Potential

The models described above evolved from just a few years of research. There is no doubt that new technology will emerge, and models and algorithms will be updated at an unprecedented speed. Data is not static. It inevitably leads to the conclusion that efforts in research and development, deployment of talent, and other attributes that drive future growth potential, are of vital importance to future economies and businesses, irrespective of the level of digitalization at that time. Understanding the core technology and the data a business uses, in my view, will lead to clarity on where the value is created and hence a reasonable tax base allocation.

What Does This All Mean for International Taxation?

The example above illustrates how we consider a more detailed analysis might be necessary in analyzing whether and how data and user participation contribute to value creation. It may be helpful to think of the role of data and user participation in the context of business value creation on a scale, illustrated in the diagram below.

The diagram depicts the trend of business moving from passive collection of data with low reliance on it in creating value, to solely relying on actively collected user-specific data with real-time analyses. Using the earlier example, at the lower end of the spectrum may be traditional car manufacturers that are looking to build in computer-assisted cruise control and parking features to enhance their existing products. At the other end would be completely automated self-driving car manufacturers whose products would not exist at all without data collection and analysis. Even at the extreme end of the spectrum, it is worth noting that the scale and methods of data and user participation involved could be vastly different. The line drawn between the roles played by user-specific data and generalized data in value creation quickly blurs as illustrated in the automatic driving system example



above. To what extent should we decide that local user data is so vital that a change in the tax structure is warranted to capture the contribution made locally? It requires a deep understanding of individual business, as well as the underlying technology and data it requires.

Tax Base Allocation (1): Focusing on User Participation

The above examples may be a convoluted way of summarizing the importance of both data and user participation. Indeed, this article aims to emphasize that the engineering behind the data-driven digital economy may not be straightforward when it comes to value creation and the assessment of a correct tax base. The actual delineation of the three ways of value creation proposed here may also present additional difficulties. For example, how does one distinguish the deciding factor in consumer choice of purchasing the premium service? In particular, is it due to the abundance of music inventory, or the physical features of a car, or is it primarily because of the recommendation services and the social aspect thereof or the self-driving capability? If the decision is based on a mix of factors, how does one assess and allocate the difference and attribute profit accordingly? Such difficulties would warrant very carefully crafted comparable analysis as a base case for A/B testing.[4]

To further understanding of the underlying technology and its position in creating value for the business, one needs to understand the nature of the data the underlying technology requires. In particular, a break-

[4] Details of how A/B testing may work in assessing tax base and profit allocation is beyond the scope of this article. In a nutshell, A/B testing is a statistical technique to compare two versions of a single variable and determine which of the two versions is more effective in achieving the desired goal. It is widely used in the software engineering field.

ISSN 0090-4600

through point for tax to make sense of the digital economy may be, first, to understand the nature of data required to enable the technology — whether it is user-location-specific, proprietary to the business, or generic and publicly available. This understanding would be a powerful tool for tax authorities to assess tax base, when coupled with information on the engineering time spent on developing a product, and certain trackable metrics, such as the API calls[5] embedded in digital businesses. The examples here are a simple illustration to show the first step of this process — i.e., that if a tax authority wants to tax on the basis of user participation, specific user data proprietary to a business should be separated from the publicly available data to determine the value creation process. Inevitably, this area will need to be explored further for fair and effective international taxation framework going forward.

Tax Base Allocation (2): Focusing on Data Access

To complete the argument, the opposite side of focusing on user participation for tax may be that user participation should not be considered at the center of the discussion at all. It could be excessive and unrealistic for tax to grasp the nuances of every type of technology and the algorithms used behind it. Instead, the real focus should just be on data itself, irrespective of the level of user participation involved. Just like labor and capital, data has borders. Its generation and origin are trackable. It is even testable, if necessary, through exercises such as A/B testing,[6] to ''switch off'' a source of data input from a particular geographic location and evaluate the impact of the absence of that particular data input on any designated metric such as revenue or profit margin. In this sense, the digital economy provides tax professionals much more powerful tools to assess and determine an appropriate tax base. Given the nature of data, and the fact that no modern business could operate without it, consideration for international tax should be around individual market access right for data.

For most companies in the digital economy, the right to access data of a particular market is vital to success. A company could argue that no user participation or very little is needed in its complex technology-driven business model. However, in many cases, it would be impossible for a modern business to declare its operation does not collect or use data for decision-making of any sort. For some companies, data is the raw ingredient fed into the technology on which they are built. In such cases, even a business with no profit has benefited from the data by forming a potentially valuable and scalable business for the future.

Taxing market access to data could entail a complex calculation process and require much more discussion and deliberation among all countries concerned. The value of market data access may need to refer to many mutable and immutable attributes such as market share, population and potential revenue. One way to structure it could be like a Value Added Tax, charging a business for entry to a particular market, and the expense is deductible if the business made profit and paid tax on it. To reduce its impact on competition and start-up companies, loss-making companies may be exempt from such cost or, if warranted, the tax asset arising from its loss may be restricted to account for value creation through growth potential generated from accessing the market data now. Further, to ensure fairness, the concept of ''envy-free cake cutting''[7] might be helpful, where tax authorities and third-party financial institutions first determine the value of market data access right before multinational companies demonstrate and allocate tax base by reference to data used and value created in each market accordingly. An international tax framework taking account of the right to access market data could have endless permutations and is dependent on the nature of each business model in each market.

CONCLUSION: KEEPING UP WITH THE CHANGES

This article provides an extended conceptual base for value creation to account for the unique nature of the digital economy. On the long and difficult journey of updating the current taxation system worldwide to catch up with changes in the digital economy, we need to understand in detail how value is created with carefully outlined parameters. We next need to understand the underlying technology within a business, its exact use of different types of data and the extent of reliance on user participation to create different values. This

[5] API stands for application program interface, which, broadly speaking, is a set of protocols and tools used to build software applications. APIs are used as building blocks for programmers and developers to build integrated digital services and are commonly used in almost all digital businesses. An API call is a request from a user to a computer server to use a specific application. User logins, saves, bookings and search queries are all examples of API requests and are completely trackable for user activities and product accesses.

[6] A?B testing is widely used in the digital economy. It simultaneously compares two versions of a test subject by varying only one variable to determine which version is more effective based on the test result and a pre-defined metric. For example, A/B testing can be used to determine if a play button in the middle of an application is better than placing it on the side to attract higher click rates.

[7] The envy-free cake-cutting theory originated from the Bible where two people are invited to eat a cake, the first person choosing how to cut the cake in half and the second choosing which piece to eat. Researchers from Carnegie Mellon University provided proof that such protocol is inspiring to solve gerrymandering in politics where two political parties are invited to divide and ''freeze'' a state into various political districts in turn. The protocol does not give the first player an advantage over the second. *See* Wesley Pegden, Ariel D. Procaccia, and Dingli Yu, *A Partisan Districting Protocol With Provably Nonpartisan Outcomes* (Oct. 24, 2017).

ISSN 0090-4600

article provides two scenarios where the focus could be on either user participation or data access itself to establish an appropriate allocation of tax base among different markets.

International taxation is an extremely complex area not just because of the subtleties in legislative interpretations of each country but even under guidelines following the same principles. Changes resulting from technology and data brought even greater challenges to the industry. I hope this article initiates wider discussions on what is the most appropriate way of taxing business in the modern economy and illuminates the importance of comprehending the drivers of business value. Ideally, future tax professionals will possess a deep and broad understanding of how an underlying technology works in creating value for each type of business, the types of data business requires, and to what extent different types of data play a role in driving that value forward. The advancement of technology drives much of the changes in the current society. It is also becoming apparent to many researchers directly involved in dealing with data that nothing is going to be static, including data itself, and neither should taxation be.

REFERENCES

- Sophia Ciocca, *Spotify's Discovery Weekly: How Machine Learning Finds Your New Music*
- European Commission, Press Release, *Digital Taxation: Commission Proposes New Measures to Ensure That All Companies Pay Fair Tax in the EU* (Mar. 31, 2018)
- Myron J. Frankman, *International Taxation: The Trajectory of an Idea From Lorimer to Brandt* (1996, 2000)
- HM Treasury, *Corporate Tax and the Digital Economy: Position Paper*
- Wilfred C. Jenks, *The Significance Today of Lorimer's Ultimate Problem of International Jurisprudence* (1940)
- Wilfred C. Jenks, *Some Legal Aspects of the Financing of International Institutions* (1942)
- Chris Johnson, *From Idea to Execution*
- Jinyan Li, *Protecting the Tax Base in a Digital Economy* (2018)
- Alexis C. Madrigal, *Inside Waymo's Secret World for Training Self-Driving Cars* (Aug. 23, 2017)
- Mary Meeker, *Internet Trends Report 2018*
- Netflix, *Data Science and the Art of Producing Entertainment at Netflix*
- OECD, *Tax Challenges Arising from Digitalisation — Interim Report 2018*
- Wesley Pegden, Ariel D. Procaccia, and Dingli Yu, *A Partisan Districting Protocol With Provably Nonpartisan Outcomes* (Oct. 24, 2017)
- Jacob Silverman, *The Menace and the Promise of Autonomous Vehicles* (June 2018)
- Robert M. Solow, *A Contribution to the Theory of Economic Growth* (1956)
- Trevor W. Swan, *Economic Growth and Capital Accumulation* (1956)
- Piergiorgio Valente, *Digital Revolution. Tax Revolution?* (Mar. 26, 2018)
- Brian Whitman, *How Music Recommendation Works—and Doesn't Work*

ISSN 0090-4600