Arbitrage power and the disappearing financialized firm

Ronen Palan
City, University of London, UK

Richard Phillips
Iconomist/City, University of London, UK

Abstract
Modern corporations have increasingly been adopting a decentred, layered, and multi-jurisdictional form as a strategy of boundary manipulation known amongst tax lawyers and accountants as ‘regulatory arbitrage’. The argument we put forward in this article is that the scholarly work that treats these strategies as mere tax avoidance practices has contributed to an underestimation and misrecognition of the way contemporary multinationals operate in markets. These strategies, which we explain in terms of arbitrage power, exploit the difference between exchanges in an imaginary ‘smooth’ market of the economic textbook and a global market that is divided among legal authorities, each imposing their own rules, regulations, and taxations. Arbitrage power exploits differences between the location of market exchange and the location of the registration of property title transfers, combining this with a manipulation of formal systems for recognizing business enterprises in order to escape some or all of the rules and regulations of society. The result is a marked difference between the ‘brochure multinational’, the way multinationals are seen and presented in their glossy brochures, and the way multinationals are legally and practically organized nowadays.

Keywords
Multinational enterprises, power, jurisdicitional arbitrage, evolutionary economics, globalization

Introduction
The world economy is dominated by a relatively small group of companies known by various names, including ‘multinational corporations’ (MNCs), ‘transnational corporations’ (TNCs), and ‘multinational enterprises’ (MNEs). Although numbering only in the thousands, collectively...
these organizations are responsible for about 33% of global output, 49% of export, and 23% of global employment (Garcia-Bernardo, 2021). Between a quarter and two-thirds of global trade in merchandise and commercial services takes the form of intra-firm trade, that is, exchange among affiliates of the same firm (Lanz and Miroudot, 2011; UNCTAD, 2016). Together with global banks, these companies form a core organizational and administrative vehicle driving the processes we call ‘globalization’.

This article centres on some puzzling but little-studied aspects of these organizations. The first puzzle is the chasm between economic, political, and sociological perceptions of these organizations, on the one hand, and their legal manifestations on the other. Despite the common reference to singular economic entities, such as ‘Apple’, ‘Amazon’ or ‘General Motors’, supposedly transacting in markets, buying off politicians, or failing to pay tax, and the like, an economic organization of this nature does not have a legal status, nor can it participate directly in markets or pay tax (Robé, 2011). In contrast, a corporation is a legal entity, licensed by a legal authority. As such, a corporation cannot be multinational. Most economic organizations operating across borders right now are set up as a network of legally independent persons, or subsidiaries, held together through equity ownership. Only those subsidiaries participate in market exchange.

The second and related puzzle is that instead of concentrating exclusively on their supposedly core task of coordinating factors of production or developing mean, lean, efficient organizational structures, modern multinationals are creating ever-more complex networks of subsidiaries and affiliates held together by layers of holding entities (Palan et al., 2021; Phillips et al., 2021). A New York Federal Reserve Bank study found the number of subsidiaries and affiliates owned by some of the largest US banking holding companies rose to an average of 3,400 in 2012, up from about 1,000 in 1990 (Avraham, Selvaggi and Vickery, 2012). One analysis of the 100 largest non-financial corporate groups in the world found this group of companies averaged above 700 subsidiaries per group in 2018 (Phillips et al., 2020). A follow-up study conducted 18 months later revealed a rise, on average, of 7% in the number of subsidiaries per group.

The argument we put forward in this article is that modern MNEs are adopting a decentred multi-subsidiary, multi-jurisdictional, layered form as a strategy of boundary manipulation known amongst tax lawyers and accountants as ‘regulatory arbitrage’. These strategies exploit “gaps between the economics of a deal and its regulatory treatment, restructuring the deal to reduce or avoid regulatory costs without unduly altering the underlying economics of the deal” (Fleischer, 2010: 227). These strategies evolved because of the dual role the law plays in modern, regulatory-rich societies. The law constitutes the formal dimension of what Douglas North calls ‘the rules of the game of society’ (North, 1990), by which he means the regulatory environment that enables and constrains action. On the other hand, the law also plays “an essential role in the social recognition, conceptual definition, and historical evolution of business enterprises” (Orts, 2013: 1). Corporate organizations employ formal systems of recognition for business enterprises to escape or avoid some or all of the rules of the game of society. They do so by camouflaging themselves, manipulating their boundaries, and reshaping their ‘presence’ in different regulatory realms, so that states, which both sanction and apply the constraining rules, have difficulty recognizing MNEs and hence may struggle to apply the rules to them. Such techniques can result in a significant reduction in costs, such as taxation, rules of financing, corporate governance, or imposition of liabilities, and can even add value to existing transactions (Karayan et al., 2002). In fact, these techniques are so valuable to these organizations that they are shaping the entire legal and corporate structure of modern MNEs.
The arbitraging firm is not only seeking cost reduction and regulatory avoidance; it is a firm that poses serious analytical problems to the theoretical universe imagined by economics and political science. For reasons already discussed at length in the literature, economic organizations tend to be seen as “simple aggregates of lower level entities” (e.g., products, consumers) (Kay, 2000: 9). Organizations are personified and take on human traits, such as rationality, utility, and the like (Fama, 1980; Kay, 2000). Accordingly, regulations tend to focus on multinationals as if they were singular entities endowed with individualistic traits. This theoretical edifice and the regulatory realm it begets reach their limits when those ‘simple aggregates’ deploy strategies of disaggregation as a competitive market strategy. In the next section, we return to John R. Commons’ institutional perspective to make sense of the technique of jurisdictional arbitrage and ask what it tells us about the way firms operate in a market economy.

Politics and economics of individuated subjects

It is commonplace today to think of the social environment as if it were populated by individuals and their products (physical and institutional). ‘Events’ are enacted (and hence observed) as a concatenation of individual actions. Outcomes are conventionally assumed to be the result of the purposeful actions of rational individuals. As Kenneth Arrow points out, “It is a touchstone of accepted economics that all explanations must run in terms of the actions and reactions of individuals” (Arrow, 1994: 1). To this, Hayek adds: “there is no other way toward an understanding of social phenomena but through our understanding of individual actions directed towards other people and guided by their expected behaviour” (Hayek, 1948: 6). In a rebuke presumably of Marxist theory and other left-leaning theories, Paul Krugman asserts: “Economics is about what individuals do: not classes, not ‘correlations of forces’, but individual actors” (Krugman, 1996: 1). ‘Empiricism’ – a contested concept – is therefore assumed to be the study of human action and reaction.

These claims about the primacy of the individual are presented as common sense and intuitive. But the ‘observation’ of those individual actors in their social setting can only take place within an agreed-upon concept of an environment in which those subjects ‘act’. To identify something we need a prior consensus, whether implicit or explicit, about the nature of the environment enveloping that ‘something’. There must have been a prior history, an epistemological turn, that gave rise to the modern theory of action (Weber, 1978). The epistemological turn emerged as a set of ideas closely aligned with the concepts and methodologies of nineteenth century scientific thought, particularly nineteenth century physics (Mirowski, 1992). A conception of social time and space as emulated metaphor for what today is called ‘Newtonian space-time’ became the standard approach in the social sciences as well. In this analogous space-time social universe, events are located on a three-dimensional spatial grid, so that something taking place here cannot simultaneously take place there. Social time has an affine structure, with predictable time intervals and an orientation from past to future. As in the Newtonian universe, human action is viewed as a movement along this time-space grid involving an expenditure of energy.

‘Steady state’ is the presumed default position of this universe centred on ‘individuals’ and ‘action’ (Weber, 1978). Action, or the expenditure of energy, is done for a reason, and the reason is subjective; that is, the individual will only shift from a steady-state position to action if he or she perceives some ‘utility’ or gain. Thus, this specific notion of time and space has strong conceptual affinities with utility and value operating through a system of supply and demand (Marshall, 2009).
Paul Samuelson points out that economics has simplified its assumptions, removing traces of moral, utilitarian, and welfare connotations (Samuelson, 1938). What is left, he suggests, is the idea that economics deals in measurable quantities. To be science at all, Jevons argues, economics “must be a mathematical science. To me it seems that our science must be mathematical, simply because it deals with quantities. Wherever the things treated are capable of being greater or less, there the laws and relations must be mathematical in nature” (Jevons, 1879: 3; see also Walras, 2013). And as Waterman (2019) has pointed out, because economic analysis is a mathematical activity, as one’s mathematical understanding becomes richer, so too does one’s ability to provide explanations for economic phenomena.

The focus of economics on ‘measurable activities’ has had some profoundly paradoxical consequences. The concept of economics as mathematical science begat, in turn, the theory and practice of national statistics, and, over time, those statistics began to be used for ancillary purposes in the form of regulations. Regulated actors seek to limit their exposure to regulation. A common strategy adopted by regulated actors is mutating into less recognizable and measurable forms. Measurable economics activities became subject to what became known as Goodhart’s Law: Any observed statistical regularity will tend to collapse once pressure is placed upon it for control purposes. This way, economics created, paradoxically, the conditions undermining its theoretical predicates over time. Or put differently, economics inadvertently encouraged mutations of economic activities to avoid ease of measurement. Non-standardized (and hence difficult to measure) but crucial sets of ‘economic’ activities, such as those performed by lawyers and accountants, moved into the breach. The result was an expanding range of economic activities that cannot easily be measured. These activities are either ignored or treated as if they were standardized and measurable – even though they are not. Either way, errors creep in.

There is nothing new here. Acknowledging from the outset that certain types of social activities, including those with ‘economic’ consequences, are not measurable, economists’ preferred solution has been to split the social realm into the ‘spheres’ of economics and politics (there are others, but they are not relevant to the present discussion). For Adam Smith, Waterman remarks, economics was different from the ‘the art of government’, to which “it may contribute but may not determine” (Waterman, 2019: 329). In the economic sphere, economic action is oriented towards profit, from the Latin profectus, ‘growth, advance, increase, success, progress’. In politics, the motivating rationale for action is power and identity. Each sphere can be thought of as a marketplace with its own ‘endogenous’ system of allocating scarce resources, whether material or immaterial. But only one sphere, economics, deals in measurable quantities. In fact, politics was originally thought to be a different order altogether. And although we know this not to be the case, probably because of lingering Newtonian time-space notions, the concept of the sphere is closely associated with the idea of ‘political systems’ or ‘market’ spaces as homogenous, ‘smooth’, and differentiated only at the boundaries (wherever those boundaries might be).

Economists are perfectly aware, of course, that influencing the ‘political sphere’ can be profitable and, hence, in some way, may be considered an economic activity. The two fields of study, politics and economics, both accept, therefore, that ‘sphere-hopping’ takes place and, in fact, can be highly lucrative. Economic actors may sphere-hop to politics to work the levers of an authoritative system of resource allocation for parochial, pecuniary gains. Political actors, in turn, driven by material or pecuniary gains, may use their power within an authoritative system of resource allocation to remunerate themselves or their cronies. Sphere-hopping is not easily susceptible to economic analysis. “The interaction between institutions and organizations that shapes the institutional evolution of an economy” merely creates,
Keynes argues, “frightful muddles” (Keynes, 2016: 361). It may be best to assume sphere-hopping between economics and politics does not alter economic fundamentals, so that success in either sphere is determined by the sphere’s own set of dynamics and rules.

**Politics and economics in evolutionary thought**

We believe the above is a reasonable summation of what may be called ‘mainstream’ thought. Although subject to some critical inquiries, this notion of social space and time has achieved something approximating axiomatic status. But this time-space configuration has two problems. The actors, ‘simple aggregates’, tend to mutate, and ‘spheres’ of action – those analytical constructs that are used to put into sharp relief foundational dynamics of the economics system (markets) and political system (governance) – become unstable. These problems have been pointed out long ago by John R. Commons (1990: 1924). For Commons, neither Weber’s theory of action is a good starting point, nor the concept of the different spheres or diverging modes of resource allocation. The first does not work because scientific observations picks up not the ‘action’ of individuals as they shift from a metaphorical steady-state, but the way individuals trans-act with one another within the context of a rule-bound institutional environment (Palan, 2020). The second does not work for the simple reason that market transactions take place simultaneously in two realms, so that the exchange of goods or services is replicated as an exchange of property titles in the legal sphere. Without replication, a transaction will either be invalid or add enormously to costs, as parties to the transaction will need to pay privately for functions currently performed by states. Economic transactions take place within a regulatory environment, and this is a political sphere. Hence, politics, economics, and the law are inseparable as far as Commons is concerned. Commons argues that the simultaneous replication of exchanges in the two realms is a constitutive act of economic life to the point where economics should be centred, not on a feeling or the commodity, but on the concept of the ‘transaction’.

Unfortunately, Commons’ reasons for placing transaction at the heart of economics has been forgotten. The economic profession prefers Ronald Coase’s reformulation of Commons’ transaction theory to suit neoclassical theory of pricing and marginal cost (Coase, 2007). It is not unreasonable to assume that in the abstract world of neoclassical economics, wherein markets represent homogenous, ‘smooth’, undifferentiated spaces, the replication of exchange will be automatic, simultaneous, and instantaneous and, hence, non-discretionary. If that were the case, then the replication of transaction in legal and regulatory realms may add costs, ‘transaction costs’ in Coase’s language, but would not challenge the notion of spheres of action. MNE operate, however, in globalized markets parcellated among divergent national authorities, each enacting a somewhat different set of rules and regulations. Under such conditions, the exchange of goods or services ceases to be automatic and becomes discretionary. Discretionary replication can impact the value of the product or good exchanged, along with ‘factors of production’ that are supposed to be coordinated by businesses. Put differently, the supposed separate realms of law and politics are getting involved in something they should not, in calculations of pricing and marginal utility.

For the modern MNE, the regulatory environment in which transaction is replicated and the timing of the replication are important components of corporate planning and competitiveness. MNEs view space and time as malleable and important competitive tools. They have become specialists, not only in coordinating factors of production across different territories and spaces, but also in selecting the preferred regulatory environments for replicated transactions. They arbitrage spaces and optimize costs because they know very well
that the textbook theory of markets is an elegant abstraction, and any manager who follows
the textbook prescription is likely to fail. The empirical evidence points to the transition of MNE
management from a Newtonian space-time to a Commons universe.

What is a firm these days?

These theoretical musings have serious implications in the ‘real world’, beginning with the very
nature of the business corporation. Jean-Phillipe Robé rightly says, “What we all designate in a
loose sense as ‘Microsoft’, ‘IBM’ or ‘Toyota’ are not the legal subjects partaking in market
exchange” (Robé, 2011: 7). There is a marked difference between what may be described as
‘the brochure multinational’, the one presented by the firms themselves on their websites or
their K-10 submission to the US stock markets, and the spatial reality of the firm. We will
illustrate this point using three cases, Amazon, Sahara, and the China National Petroleum
Corporation (CNCP). These firms are some of the most important global players. Amazon is the
largest retailer in the world. CNPC was the fourth largest corporation in the world in revenue in
2020. And Sahara is an important African energy-trading company.

The brochure version of Amazon is presented on its K-10 submission as follows:

We seek to be Earth’s most customer-centric company. We are guided by four principles: customer
obsession rather than competitor focus, passion for invention, commitment to operational excellence, and
long-term thinking. In each of our segments, we serve our primary customer sets, consisting of consumers,
sellers, developers, enterprises, and content creators. In addition, we provide services, such as advertising
to sellers, vendors, publishers, authors, and others, through programs such as sponsored ads, display, and
video advertising. (Amazon, 2020)

Amazon continues:

We have organized our operations into three segments: North America, International, and Amazon Web
Services (AWS). These segments reflect the way the Company evaluates its business performance and
manages its operations. (Amazon, 2020)

This description resonates, we suspect, with the typical view of Amazon. The brochure
version of Sahara contains a series of now-familiar descriptive sections, including one called
‘about us’ with Sahara’s ‘vision, mission, core values, history, strategy’, another on its turnover
and employment, and so on. Sahara describes itself in the following words: “Sahara Group is a
leading international energy (power, oil and gas) and infrastructure conglomerate with
business operations in over 42 countries across Africa, Europe, the Middle East and Asia”.?

We could not find a brochure descriptive of CNPC, so we will have to make do with a
Wikipedia entry:

The China National Petroleum Corporation (CNPC) ... is a major national oil and gas corporation of China
and one of the largest integrated energy groups in the world. Its headquarters are in Dongcheng District,
Beijing. CNPC was ranked fourth in 2020 Fortune Global 500, a global ranking of the largest corporations

These brochure entities do not necessarily have a parallel legal presence in the legal sphere.
Figure 1 shows the legal structure of Amazon depicted by a technique we call equity mapping,
or EM in short (Phillips et al., 2020; Phillips et al., 2021). It is apparent from this that
Amazon’s actual legal structure does not correspond to its own description of itself.
Figure 1. EM of Amazon subsidiary structure, circa 2020. Source: Authors’ own.

Legend:

(1) EM: Equity mapping of subsidiary holding;
(2) Red Dot: Important subsidiary named in Amazon’s 10-K;
(3) Black Dot: Private subsidiary not mentioned in the latest report;
(4) Cluster: Large groupings of subsidiaries holding important intermediate holding companies and generating a spatial clustering effect;
(5) Conduits: Subsidiaries that own only one or even a few other subsidiaries, yet these holdings are themselves major shareholders in a large set of subsidiaries – i.e., they are visually distinguished as ‘in-between’ the main parent company and the disparate sub-clusters of the group.

For example, Amazon Overseas Holding Inc. (Delaware) is a standalone subsidiary with no further sub-holdings, whereas the majority of Amazon’s non-American subsidiaries are held via a branch headed by Amazon.com Sales Inc. (Delaware). Amazon Web Services (AWS) may be described as its own ‘segment’, yet various Amazon ‘web service’ entities are distributed throughout different parts of the structure. For instance, although there is a separate branch controlled by Amazon Web Services Inc. (Delaware), Amazon Data Services Inc. (Delaware) resides in a separate branch not linked through equity control to AWS. Other European ‘Web Service’ companies, in turn, appear to reside in the same cluster of ownership as entities involved in the retail segment.

Amazon’s legal structure bears little resemblance to the impression given in the brochure. Yet Amazon is a typical MNE adopting what is known as the parent-subsidiary model. The evolution of this model is crucial to the transition from an emulated Newtonian time-space world to a Commons world. It just so happened that the transition took place a decade or two after the advent of ‘neoclassical’ economics and was observed closely by keen-eyed economists, such as Thorstein Veblen and John R. Commons.

The parent-subsidiary model evolved from the late nineteenth century as a set of pragmatic solutions to intractable problems posed by ‘real’ global markets. As businesses sought to expand their manufacturing base across borders, they had to find ways of ensuring
their legal rights in other jurisdictions. The fledgling multinationals experimented with several administrative solutions to their problems. During the first phase of internationalization, several organizations contracted individuals who served as their foreign agents in host countries whilst others set up foreign branches. New options emerged following an amendment to the US laws of incorporation introduced in 1899 by the ‘mother of trusts’, the State of New Jersey. Soon emulated elsewhere in the US and beyond (Arsht, 1976; Cheffins, 2015; Yablon, 2006), this legal innovation allowed a corporation to own stocks in other corporations, creating a facility by which one corporate entity could control another through equity holding and by so doing ensure the two operated as one unified business enterprise (Grandy, 1989). The model became popular, not least because the device of the public corporation offered several advantages over agents or branches. More specifically, it allowed subsidiaries to raise funds in local markets whilst simultaneously restricting potential liabilities to the parent company (Blumberg, 1993; Ferran, 1999).

Due to these and other advantages, MNEs began to replace their system of foreign branches with the parent-subsidiary model. The results, described by some as ‘fiction’ (Greenfield, 2008), were such that for all intents and purposes, at least in the eyes of the law, MNEs now consist of separate legal entities, and all are subject to the rules and regulations of their respective countries. That abridged history of the evolution of the artificial legal person concept helps explain the broad contours of an MNE like Amazon.

We conducted a similar EM analysis of Sahara and CNPC. The results were astounding. Although Sahara describes itself as a ‘group’ in its brochure, the various group divisions listed in its brochure do not have any legal formal relationship. ‘Sahara’ is, in fact, a set of independent companies with no known legal relationship between them! Sahara truly does not exist as anything approximating a corporate group. It seems to act as a unified economic entity in the business world but, contrary to perceptions, there is no formal governance structure. This may not be as unusual as it sounds. US tax laws, for instance, recognized long ago that unincorporated legal entities can act in the economic sphere. US law employs a pragmatic measure for such unincorporated bodies based on their resemblance to conventional state-law corporations (Speck, 2015: 2584). Nigerian law does as well, thus allowing entities such as Sahara to trade (Rogers, 1981).

CNPC, the fourth largest MNE in the world by revenue, combines both versions of corporate groupings. CNPC has established a publicly owned and listed group in PetroChina Co Limited (80% owned by CNPC). PetroChina appeared to be structured as a Western group subject to the trappings of internal governance rules and regulations demanded by Western investors. CNPC has created another set of subsidiaries, many operating internationally, that are not linked in a formal way to PetroChina. CNPC contains, therefore, different organizational formats, each presumably used for a different purpose. It is a Western+ corporation, something barely known or understood, even by employees of such organizations. We try to capture this odd combination of a group that is not a group in Figure 2 on the next page.

Zone 2, outlined in blue, includes the subsidiaries as well as other forms of equity ‘partnerships’ less than or equal to 50% of PetroChina. Zone 2 represents a Western type of MNE. It is subject to the reporting and consolidated accounting rules typical for an MNE. Another zone contains about 110 entities controlled by CNPC, represented in Zone 1 in the tan colour. These are privately held companies. This zone is legally distinct from PetroChina and thus not subject to the same corporate governance requirements and auditing as PetroChina itself. Only six of the 110 Zone 1 entities offer consolidated financial accounts. All other financial reports are unconsolidated.
All three groups discussed above are ‘going concerns’ and may be considered MNEs. Amazon is perhaps more typical of the predominantly Western model of an MNE. Sahara is a Western-version and CNPC a Western+ version. Both alert us to the fact that groups can arise and function successfully as ‘going concerns’ in a far more legally diverse manner than assumed. These groups may lack a formal governance structure, like the one preferred by stock markets; indeed, their different components may lack a legal relationship whatsoever. Their modes of internal control, governance or distribution of liabilities are unclear, and they challenge our very notion of what a firm is. And yet they are known as MNEs. Like shadow banks, these organizations seem to swim like an MNE and quack like an MNE, so they must be an MNE.

The boundary of the firm

Brian Cox, the famed pop musician turned physicist turned documentary maker, once observed how similar elements of chemistry and physics can interact to create diverse worlds. By the same token, Amazon, Sahara, and CNCP alert us to the way similar sets of ingredients can produce diverse organizations. As these three firms indicate, the concept of the firm is dynamic, fluid, and changeable. Firms can take a variety of forms; their formal structures are constantly evolving as new subsidiaries and affiliates are set up and new paths of control emerge or disappear. As their formal structures change and evolve, firms’ operational tasks and functions can metamorphose to such an extent that the formal ‘brand’ name is now doing something else entirely. The boundary of the firm is fluid as well. The location of the legal boundary is discretionary, as is the format taken by the group. Indeed, a brochure hides a great deal of variety and fluidity under a title or brand name.
Property law and temporal arbitrage

As far as we can tell, the first set of replication issues to have an impact on corporate strategy and organization did not appear in taxation. Rather, it arose from changes in the sphere of property laws. Profound changes in property law took place in the US, argues John Commons, stimulated by the rise of large businesses centred on new technologies of locomotion in the late nineteenth century (Commons, 1924). The changes were catalysed by the decision of the Federal government to expand the railway system through franchises, a system that brought about the rise of ‘robber barons’ (Josephson, 1962). As a trans-state railway line controlled by a franchisee traversing state boundaries was being built, railway companies came into conflict with the states along the route. Those states had an interest in controlling the cost of transport through their territories, something naturally opposed by the railway companies. The conflicts often ended in the courts. The courts tended to accept states’ rights to control prices, but they also accepted that price control denied the railway companies both current income and future income streams factored into their business model and investment decisions. In doing so, argues Commons, the courts accepted that the railway companies possessed something with no clear status in law until then, a right over future income streams. This was an intangible right, yet it had a value, calculated by factoring in anticipated future earnings discounted against current interests (a method of calculation accepted by the courts). This right was capitalized into the firm’s stock market valuations and entered by accountants under an archaic concept derived from English common law as ‘goodwill’ value on the firm’s balance sheet (Atkinson, 2009; Palan, 2015).

Of course, the future plays an important role in marginalist thought as well. For Mengers, the essence of capitalism is the extension of manufacturing processes to the future (Moss, 2010). In Marshall’s view, “material goods consist of useful material things, and of all goods. rights to hold, or use, or derive benefits from material things, or to receive them at a future time” (Marshall, 2009: 45). The future also played an important role in the debate following the introduction of the Modigliani-Miller invariance theorem (Modigliani and Miller, 1958), with some explicitly arguing that investors capitalize on future earning capacity.

Commons reads these trends differently. He argues that the courts’ decisions regarding the railway issue amounted to a subtle but important alteration in the nature and concept of property. Intangible rights demarcate lines of control over future income streams and can only be valued based on the notion of exclusivity, not of physical objects, but of access to a certain resource in the future. As he puts it, “The change in the concept of property, from physical things to the exchange value of things is a change from a concept of holding things for one’s own to withholding things from others’ use” (Commons, 1924: 52). A right over trademark, patent or intellectual property is a right to exclude the enjoyment of that good or service by others. The ‘value’ of such intangible rights is determined not only by demand and supply, but also by the market estimation of the withholding capacity of the title owner. Suppliers of goods and services may not necessarily rush to meet demand as assumed by marginalists but will often seek ways to control supply to raise prices.

Be that as it may, these changes in the concept of property had profound implications on business strategy and planning, and ultimately on business’ emerging relation with time. In simple affine time, business is supposed to maximize profit by producing at a cost lower than the market price. Profit is generated through the ‘accumulation’ of discrete sales, and all profitable transactions are added up to yield an annual profit. But if the ‘value’ of the going concern factors into future income streams, investors may be equally, if not more, interested in the capitalized value of companies. Companies could aim for a strategy of demonstrable
capacity of controlling the future by showing a ‘withholding’ capacity (Palan, 2015). Far from concentrating exclusively on coordinating factors of production, a business could opt to extract maximum value out of a transaction in ‘real time’ or to build up ‘withholding’ capacity, thus leading to higher market capitalization.

The two strategies are not necessarily in conflict with one another, but they can lead to the development of very different organizations. For example, consider Amazon. Amazon has made it public knowledge that it has forsaken profit maximization in favour of a strategy of expansion and market capitalization. To ensure its website is the first port of call for would-be consumers and, in this way, build up its ‘withholding capacity’, Amazon is often happy to sell at cost or even below cost. Because of its size and market reach, it may be distorting the clearing operations of the market. Amazon shows there is no longer such a thing as a ‘representative firm’.

The strategy of capitalization is the result of the replication issue. The changes in law and accounting took place first in the US, and it took time for these concepts of property to be copied and implemented elsewhere. Whilst American companies were capitalizing futures already in the late nineteenth century, benefitting from their ability to draw money from the future, others could not. Admittedly, Western European countries followed suit very quickly. But Communist enterprises, including Soviet and Chinese state enterprises, could not borrow capital from the future and fell way behind their capitalist competitors.

Many developing countries also lacked the necessary legal and a stable political system to sustain this type of property rights. Lacking advanced stock markets and the legal facilities to convert the rights over future income streams into a fungible capital-raising mechanism, they increasingly relied on an archipelago of secondary financial markets. This option spread like wildfire around the world in the 1960s. The majority of jurisdictions were either remnants of the British Empire or previous members of the Empire (e.g., Singapore, Hong Kong, and Cyprus), holding to English common law and having close institutional connections with the City of London (Palan, 2010). MNEs began to use those regional centres, setting up subsidiaries, which, as independent legal persons, could raise capital and take advantage of their brand names as weapons of futurity. The phenomenon of intermediated investment holding or regional treasury operations began to evolve and played an increasingly important role in the construction of MNEs (Polak et al., 2011).

As a result of these subtle changes in property rights, and the staggered introduction of changes around the world, MNEs developed a whole new set of skills that did not involve the traditional focus on the coordination of factors of production. Many businesses, particularly large ones, began to ‘metamorphose’ (in the eyes of economists) into something else and appeared increasingly geared towards financial rewards. Yet from a business perspective, the strategy of building up withholding capacity is simply aimed at taking advantage of the array of opportunities, penalties, and rewards made available by existing rules and regulations. The puzzle is why doing ‘business as usual’ is considered an aberration or a departure from what business should be doing.

Agency law and spatial arbitrage

Diverging business strategies involve important decisions revolving around the legal format of the corporate entity, known as agency law. There are many varieties of legal formats, public corporations, private corporations, partnerships, and so on, and these differences are linked to location. Economists and economic geographers have long understood one side of the equation: location is a crucial dimension of corporate organization and governance (for a good
summary, see Beugelsdijk et al., 2010). But this notion of place and space is driven by a ‘productionist’ assumption built into late nineteenth-century types of theory, with location theories tending to ignore the complexities of legal rules of agency, property, and contract.

During the 1920s, long after the successful reformulation of economics by the marginalists, the broad principles of taxation for multinational enterprises were agreed upon (Avi-Yonah, 2005; Kudrle, 2019). The taxation of active business income was based on ‘source’ jurisdictions where the business was located, whilst passive income (or investment income) was based on the jurisdiction where investors resided.9 The concept of ‘source’ income refers to the location of activities in space, irrespective of the legal structure adopted by an organization. On the surface, redundancy seems built into the principle because taxation applies to the legal person, not the named ‘MNE’. As separate legal entities, each subsidiary and affiliate is subject to the rules and regulations – including the rules of taxation – of its country of registration. The result is that the amount of taxation paid by an MNE is a summation of the taxes paid by each subsidiary and affiliate. And since each subsidiary tends to operate in its country of registration, ‘source’ taxation is automatically assigned to those entities. In theory, the alignment between the principles of taxation and the legal structure of the firms ensures ‘source’ taxation will be paid.

Reality has proven very different, and here, the difference between the legal structure and behaviour in the economic realm is crucial. Because these organizations tend to operate under a singular economic umbrella but use the facility of legal incorporation of supposedly independent entities, a core task of the central ‘bureaucratic’ apparatus is to make decisions about the location of activities from a regulatory perspective. Not surprisingly, evidence has begun to mount that MNEs are shifting profitable activities to low-tax jurisdictions (Clausing, 2016; Cobham and Janský, 2018; Crivelli et al., 2015; Garcia-Bernardo, 2021; Kaye, 2014; Zucman, 2013). The reorganization of activities along rules of taxation has nothing to do with national characteristics. It is simply the result of diverging national rules and regulations. Such behaviour does not fit the traditional model of the firm, so the tendency has been to dub it ‘artificial’ (aberrant as far as theory is concerned) (Desai et al., 2006; Dharmapala, 2008; Slemrod and Wilson, 2009). Research has identified, or is supposed to have identified, certain portions of the organizations that are now ‘sham’ or ‘artificial’ (Damgaard et al., 2019) – a concept presupposing one reorganization of corporate entities across space is less ‘artificial’ than another. But as every corporate transaction is supposed to be scrutinised for its regulatory efficiency, the exact criteria for identifying artificiality have never been too clear. Indeed, in the view of managers, accountants, and lawyers, the only artificiality is the normative distinction made by academics and regulators between legitimate and artificial activities.

Be that as it may, the concept of artificiality has become entrenched, stimulating a whole set of rules and regulations aimed at combatting artificiality. The effect has been, argues Judith Freedman, the development of ever-more exotic devices of artificiality (Freedman, 2008). These involve far more sophisticated uses of spatial and even temporal manipulations, specifically the reorganization of control over subsidiaries and affiliates with the goal of exploiting gaps, loopholes or omissions in the laws of one country to take advantage of the rules of another. These techniques are known as ‘jurisdictional arbitrage’ (Fleischer, 2010; Panayi, 2006; Phillips et al., 2020). In so doing, multinationals are not simply arbitraging the laws and regulations of different countries; they are also arbitraging the common perception of what they are (Blumberg, 1993; Greenfield, 2008). They are happy to present themselves in their brochures as single, unified actors operating as a ‘multinational corporation’ whilst, in reality, taking advantage of the parent-subsidiary model to arbitrage national rules and
regulations, most of which are aimed at corporations instead of groups. The language of the
firm has a performative power that has led regulators to pursue the wrong side of the
equation. Regulators are caught in a game of a whack-a-mole, chasing an ever-disappearing
subject.

**The law of many prices: Contractual law arbitrage**

A common mistake is to assume arbitraging is only about cost reduction in the form of
taxation. Fragmentation of the law, including rules of governance and contract, encourages the
arbitraging of contractual laws as well. In a seminal article, Frank Partnoy notes that the
predominant assumption in economics is the 'law of one price':

> The 'law of one price' states that identical goods must have identical prices. If identical goods do not have
identical prices, there will be pressure from financial arbitrage-simultaneous buying at a low price and
selling at a high price-until the prices of the goods converge. Thus, financial arbitrage is the key mechanism
in the adjustment process that leads to identical goods having identical prices. (Partnoy, 2018: 1018)

Regulatory arbitrage, he argues, does exactly the opposite. It entrenches price diversity
for similar goods and, in fact, pits jurisdictions against one other. The difference in price can
be attributed to tax, subsidy or other rules affecting the pricing of the good or service in
question. To take one example, since August 2004, limited partnerships and limited liability
companies organized under Delaware law have been permitted to contractually modify, limit,
and even eliminate the fiduciary duties of their managers as specified under common law.
This allows managers a great deal of discretion and removes what is called 'agency cost'. For
instance, in Delaware, managers do not need to honour the mandatory distribution of profits
to shareholders and have considerable discretion about whether to draw upon market
financing or self-finance (Cheffins, 2015). As a result, managers have greater leeway deciding
whether to draw on external finance or reduce dividend distribution while increasing internal
contribution. A Delaware subsidiary can optimize financing costs far better than, say, a similar
subsidiary in Pennsylvania. The notion of a ‘global’ financial market notwithstanding, where a
funding subsidiary is located may significantly alter the cost of financing a venture for MNEs.

Another issue affecting the price of goods is that access to legal defence in contractual
disputes is determined by the host country’s courts and norms (Dupuy et al., 2009). Perceptions of a costly or unresponsive court system represent a risk factored into price
values. This might count as a transaction cost. But MNEs often respond to these types of
transaction costs by not ‘internalizing’ them (Coase, 2007). Instead, when they doubt the
quality or independence of the courts in the host country, multinationals often insist on the
transferral of contractual obligations to a third country. They typically set up an intermediary
subsidiary or set of subsidiaries in a preferred regulatory environment offered by this third
country (often as a joint-venture), ensuring that as many of the contractual obligations
associated with the investment as possible are transferred there.

Parnoy’s theory of the ‘law of two prices’ refers to a debate in economics about price
divergence. To speak to this literature, Parnoy assumes a hypothetical scenario of two
regulatory authorities. For MNEs, however, the potential for arbitraging is not limited to two
regulatory authorities but can extend to several. Therefore, MNEs face, not the law of two
prices, but the law of several prices, often ranging within the hundreds. ‘Factors’ such as
labour, finance, insurance, and the value of raw material and semi-manufactured goods can
vary significantly between countries. Equally, similar goods or services are likely to be priced
differently across jurisdictions. Sophisticated organizations understand very well that what economists call ‘coordination of factors of production’ is vastly improved by reallocating factors amongst jurisdictions. They plan their legal structure accordingly, adding to the number and ownership pathways of subsidiaries, since alternative pathways may change the economics of their investment. Simply put, coordinating factors of production under the law of one price takes one form: the traditional firm. The coordination of factors of production in economies of multiverse prices is a different proposition altogether; it takes the form of the arbitraging firm. This may answer the question about the proliferation of quasi-economic subjects; these subjects are arbitraging national rules and regulations, creating an optimized path of least resistance through the regulatory mire.

**Economics is the study of firms that are not firms**

Economics is about individual actions and reactions with firms as ‘simple aggregates’ (Kay, 2000). The ‘simple aggregates’ of economic theory are now busily self-dividing in a way not sanctioned by theory, with no individual action or reaction in sight. Lacking the theoretical tools to describe such behaviour, theorists have tended to treat arbitraging as ‘white noise’, either ignoring it or viewing it as an aberration on the margins of economic activities. But economists can ignore this white noise only up to a point.

Problems began to surface with an apparently simple question raised by Edith Penrose. Penrose was interested in finding out whether it was possible to establish a general theory of the growth of firms, why and how they grow to certain size and not another. The problem, she realized, is that the question of the growth of the firm is closely linked to the question of the boundary of the firm. After all, we can only measure the growth of ‘something’ if we can agree on its ‘size’, and size is discerned by boundaries. These assumptions go with the territory of what we call the ‘theory of the firm’. This rather innocuous set of propositions forced Penrose to confront the sort of widespread boundary manipulation activities that we describe in this article. But rather than describe those issues as arbitrage, Penrose concludes that the most elementary assumption of the economic study of the firm, namely, that it has boundaries and there are plausible ways of determining where one firm ends and another begins, is in fact extraordinarily difficult to establish. In her words:

> That a firm has boundaries follows from the nature of the categories we think in, however, not because we can clearly ‘observe’ them in reality. The boundary of the firm is what distinguishes it from the market and therefore must ‘exist’, whether or not it is ‘real’ since the firm/market dichotomy has been perhaps the major building block of an economist’s analytical thinking. (Penrose, 2009: xvi)

As Penrose explored this notion of the boundary, and the relationship between theory and observed behaviour, she came up with a second remarkable idea, namely that the problem with the boundary of the firm might have a temporal dimension: “At some point in the past a reasonably identifiable corporate form existed” (Penrose, 2009: 20). But by the early 1980s, determining the boundaries of the firm was proving highly problematic for a variety of reasons, chief of which, she felt, was size:

> As an industrial firm becomes larger and larger, and its operations become progressively more decentralized... [some firms] increasingly acquire the characteristics of a financial holding company, lose those of an industrial firm, and finally become virtually indistinguishable from an investment trust. (Penrose, 2009: 19)
As these organizations evolve, a similar organizational structure bearing the same brand name can change tasks and functions. It becomes something else entirely. It is the old cliché: beyond a certain point, quantitative change turns into qualitative change.

Penrose did not employ a fashionable term like financialization, but the implication of her ideas is that the boundary of the firm is closely associated with financialization; otherwise stated, finance is playing an increasingly greater role in the evolution of modern firms. “Can we safely assume”, she asks, “that the principles that govern the growth of an industrial firm are equally applicable when the organization is metamorphosed into an essentially financial firm” (Penrose, 2009: 19)? Her answer is no; in fact, “on the contrary, the techniques suitable for analysing the growth of firms engaged in the actual organization of production and distribution are probably very different from those required for the analysis of the growth of a purely financial organization” (Penrose, 2009: 19).

At this point in her argument, Penrose acknowledges that the firm/market dichotomy, the bedrock of microeconomic analysis, is empirically suspect, and she alludes to a theory of organizational metamorphosis or evolution. Nonetheless, she prefers to focus on aspects of the firm that existed (presumably) prior to the supposed metamorphosis. Penrose is explicit: the ‘firm’ of her studies is a truncated version of actual firms, and represents only those aspects susceptible to microeconomic analytical tools.

Penrose’s work proved highly influential. In his wide-ranging assessment of the economic theories of the firm, Neil Kay notes that contemporary economists acknowledge that “there was no irony or contradiction in the idea that in the theory of the firm, the firm was not a firm” (Kay, 2000: 14). In other words, those who wish to continue working with the standard microeconomic analytical tools – analytical tools designed specifically with the understanding that markets and politics are separate realms of interaction – can only do so by focusing on some behavioural aspects of modern firms. Put differently, the firm inhabiting the microeconomics universe is a “theoretical construct”, as Douglas Allen calls it (Allen, 2005: 899), whose primary purpose is to ‘fit in’ with those microeconomic tools. Whatever else those real living organizations are doing tends to be ignored.

**Conclusion: Voice, exit... arbitrage**

In *The History of Sexuality*, Volume 1, Michel Foucault notes: “Power is tolerable only on condition that it masks a substantial part of itself. Its success is proportional to an ability to hide its own mechanisms” (Foucault, 1998: 86). Boundary control, or the control of the relationship between the firm and regulatory authorities, is where the action is these days. The ‘rational’ MNE is employing its group structure and capitalizing on the fact that an estimated one third to two thirds of all international trade takes the form of group intra-trade (UNCTAD, 2015) to arbitrage national rules and regulations. By so doing, the MNE can select (up to a point) a regulatory environment that suits it.

A common view amongst political scientists is that when faced with a global market for goods, services, and capital, dissected as it were, amongst national regulatory authorities, corporate entities and their controllers (managers, leading shareholders) have two options: they can either voice their objections to certain regulations (taxation, financial rules, rules of liability and conduct, labour and environmental laws, and like) or they can exit, giving up on certain markets (Hirschman, 1990).

But there is a third strategy, familiar to any ecologist: just like any other organism facing environmental challenges, an institution may opt to camouflage itself, arbitrage its existence, and reduce its regulatory burden.
The capacity to shape the institutional and regulatory environment is power. Whether an MNE opts for traditional forms of lobbying for political pressure or achieves similar aims through arbitrage is of secondary importance. Both are techniques of power but only one – possibly the one that has less effect – attracts attention. Not everyone has access to arbitrage; only the wealthiest groups and institutions have the resources needed to employ the lawyers and accountants who can create sophisticated arbitrage schemes that are strictly legal whilst ensuring these organizations operate within their preferred regulatory environment.

Arbitrage is a form of power that tends, as Foucault notes, to hide its own mechanics. MNEs are not hiding what they do entirely – they cannot. But they do not advertise it either. MNEs are using all available tools to talk as little as possible about arbitrage. Yet the basic rules of reporting and accounting give them no choice but to file reports annually, and, hence, give an inkling of arbitrage. The development of MNEs and their use of such power tools are masked, largely for other reasons, including the continuing commitment to the standard theory of economics and political science, and their individualistic lenses through which action is perceived and evaluated, sustained by a certain concept of space/time that supposedly envelopes economic activities. As a result, routine but important tasks performed by multinationals are treated as minor sets of activities or as illegitimate. The issues of transaction replication and the development of group structure of multinational enterprises are ignored while normative assertions about the artificiality of these practices replace actual analysis. Jurisdictional arbitraging is ether ignored as ‘white noise’ (not because it is unknown but because it takes place in a space that is a blind spot produced by disciplinary fragmentation) or otherwise considered insignificant.

Acknowledgements

This article is based on a keynote address given by one of us, Ronen Palan, on the eve of the Finance and Society Network Annual Conference in London, UK, on December 11, 2019. Many of the ideas in this article evolved in discussions with our friends and colleagues, including Lorraine Eden, Javier Bernardo-Garcia, Sandy Hager, Petr Jansky, Photis Lysandrou, Yuval Millo, Anastasia Nesvetailova, Hannah Petersen, Cecilia Rikap, Jean-Phillipe Robé, Amin Samman, Jason Sharman, Herman Mark Schwartz, Kirstin Surak and Elizabeth Thompson. Research for this article was supported by the European Research Council Advanced Grant no. 694943, “Corporate Arbitrage and CPL Maps: Hidden Structures of Controls in the Global Economy” (CORPLINK).

Notes

1. “It is commonplace for business to be conducted through a network of connected companies rather than by a single corporate entity” (Ferran, 1999: 26). At times, an MNE may set up an overarching controlling structure for the group holding and this may be called an ‘MNC’. This format is found mostly amongst ‘Western’ groups. At times, the MNC does not possess even the most rudimentary governance or legal structure yet manages to operate in markets. A recent study found about half of US firms have adopted a complex multi-subsidiary and multi-jurisdictional organizational structure (Lewellen and Robinson, 2013).

2. Stanley Jevons (1879: 18) expressed this idea when he said the “science of Economics ... is in some degree peculiar, owing to the fact ... that its ultimate laws are known to us immediately by intuition”.


3. What exactly do we mean when we say all action is explained in terms of actions and reactions of individuals? Does it mean individuals are the sole agents of order and change? Or does it mean individuals ‘play out’ structural or institutional forces, and, hence, observed behaviour describes what those individuals believe is of utility to them? Both scenarios can be described in the same way. Both can result in consistent choices made by agents and both may yield equilibria. Our sense is that ‘methodological individualism’ tends to refer to the first scenario.

4. Walras (2013: 84) says: “The whole world may be looked upon as a vast general market made up of diverse special markets where social wealth is bought and sold”.

5. That has not deterred political scientists from trying, with a (debated) modicum of success, to identify areas of measurable quantities in politics as well.

6. Paul Davidson argues that Keynes ‘demolished’ at least the time dimension of this configuration by introducing the notion of non-ergodic time. Linear repeated time does not work in standard economics: “The economy is governed by an existing ergodic stochastic process. One merely has to calculate probability distributions regarding future prices and output to draw significant and reliable statistical inferences [information] about the future. Once decision makers have reliable information about the future, their actions on efficient free markets will optimally allocate resources into those activities that will have the highest ‘known’ future possible returns, thereby assuring the highest possible global prosperity in the future” (Davidson, 1991: 134). Keynes left space untouched, but the issue of time and the relationship between past and future are no longer considered simple Newtonian equivalents.


8. In light of these findings, we suggest Sahara may have changed its legal structure.

9. Most leading countries, save the US, have abandoned the system of passive taxation. Over time and for a variety of reasons, passive income taxation, known as resident taxation, has been abandoned by the majority of countries (for a discussion, see Matheson et al., 2013). Amongst the major OECD countries, only the US and the Netherlands hold on to the principle of resident taxation – although even that is in some doubt (Avi-Yonah, 2019; Desai and Dharmapala, 2018; Matheson et al., 2013; Shaviro, 2010).

References


