INTRODUCTION

Conceptualizing the business corporation: insights from history

David Gindis

Hertfordshire Business School, University of Hertfordshire, Hatfield AL10 9AB, UK
Corresponding author. Email: d.gindis@herts.ac.uk

(Received 21 April 2020; revised 28 April 2020; accepted 28 April 2020; first published online 1 June 2020)

Abstract
The purpose of this symposium is to shed light on the genealogy of the idea of a business corporation, an economic institution which has long been regarded with a mixture of awe and apprehension. Each of the four original contributions addresses the history of some of its key features. In the process, each contributor reveals some of the insights that history has to teach us regarding the central concepts that inform contemporary debates about the nature of the corporation, the contours of the corporation’s purpose, the sources of corporate power, the functions of corporate law, the duties of directors, the status of shareholders, and the legitimacy of corporate rights.

Key words: Business corporation; corporate personality; economic and legal history; history of economic; legal and political thought; limited liability; membership and governance; ownership

JEL classification: B15; K20; N20; N80; O16

Is the business corporation one of the greatest human inventions, or is it the source of much of what is wrong in the world? This question is not merely rhetorical; serious efforts have been made to defend both positions. Today, perhaps more so than at any time since the robber barons, the ‘evil corporation’ has become a cultural trope and ‘corporate evil lists’ far outnumber ‘love letters to big business’. In the popular imagination, excessive executive pay and soaring dividends are difficult to disassociate from job insecurity, heartless foreclosures, private data misuses, environmental degradation, and corporate scandals. For some in America, the key to curbing corporate power involves abolishing corporate constitutional rights. For others, the ‘benefit corporation’ movement holds the most promise. Corporate leaders seem to be on the defensive: in a widely noted move last year, the Business Roundtable said it was abandoning the shareholder-centric model of corporate governance and committing to an inclusive stakeholder-oriented approach.

The intensity of the public debate seems to be matched by the level of academic interest in these and other related issues. While the literature on business corporations is hardly new, the bookshelf of anyone taking an interest can now be readily graced by exceptional studies of the concept of a corporation from the legal, philosophical, political, or sociological perspectives (e.g. Baars and Spicer, 2017;
Choudhury and Petrin, 2017; List and Pettit, 2011; Mansell, 2013; Muñiz-Fraticelli, 2014; Orts, 2013; Ripken, 2019; Singer, 2018). Impressive historical work, including some by economic historians, has also been produced in the last few years (e.g. Cheffins, 2019; Harris, 2020a; Lamoreaux and Novak, 2017; Lamoreaux and Wallis, 2017; Ogilvie, 2019; Pettigrew and Smith, 2017; Pettigrew and Veevers, 2018; Turner, 2016; Wells, 2018; Winkler, 2018). This symposium makes an original contribution to this literature.

Ron Harris (2009: 613, 2020a: 255–256) distinguishes several approaches in the historiography of the modern corporation: the first views it as a Roman invention; the second and third as a concept revived and applied by canonists and civilists, respectively, after the rediscovery of the Code of Justinian in the late 11th century; and the fourth sees it as originating in the medieval Germanic tribal traditions. From all four perspectives, it is clear that the idea of a corporate body endowed with a separate legal personality and its own pool of assets is distinctly European. In medieval Europe, institutions including religious orders, abbeys, monasteries, cathedral chapters, boroughs, towns, guilds, charities, and universities were recognized as such. Some were chartered, others were not. But while the corporate form was present in virtually all aspects of medieval social life, it was absent in the realm of long-distance trade.

Italian merchants trading along the Silk Road, for example, employed the commenda, a form of limited partnership that originated in the Arabian peninsula and that had functional equivalents in the Indian subcontinent and China. Family firms and ethnic networks were also involved. Much knowledge and technology made its way into Europe in this manner. There are some indications that ‘proto-corporations’ may have existed in ancient India, late Imperial China, or the Arab world, but it is not clear from the relatively small literature on the topic that these were functionally equivalent to their European counterparts or that any of their features migrated into Europe (Harris, 2020a: 331ff). Regardless, by the end of the 17th century, while its trading partners continued to rely on kin and commenda-type partnerships, European long-distance trade was mostly organized using the corporate form. This underpinned Europe’s rise to maritime dominance.

There is a strong consensus among sociologists, institutional economists, and economic and legal historians that the constitution of perpetually-existing, impersonal organizations in both public and private spheres is an indispensable stepping stone toward modernity (Coleman, 1990; Greif, 2006; Hodgson, 2015; Kuran, 2011; North et al., 2009; Weber, 1947). An essential institutional feature of such organizations is their separate legal personality, or corporate personality, which is inherently impersonal because it is defined without reference to specific individuals and can survive changes in organizational membership. Contracts become more secure when they are made with an organization whose identity is distinct from the identities of any of the individuals involved, and when they are bonded by the organization’s own asset pool, against which creditors can enforce their claims in court (Armour et al., 2017; Arrufada, 2012; Hansmann et al., 2006).

These features were recognized, but not explained, by legal commentators long before the birth of modern social science. In one of the first treatises devoted entirely to the subject, Stewart Kyd (1793: 5–8) wrote:

6 Economists have by and large been absent from the recent debate about the nature of the corporation, with a few exceptions (e.g. Deakin et al., 2017; Gindis, 2016; Langlois, 2019; Lawson, 2015; Pagano, 2010). It is also worth noting that some of what has appeared in economics journals was authored by non-economists (e.g. Deakin, 2017; Veldman and Willmott, 2017). Outside economics, numerous interdisciplinary special issues have been published (e.g. Mansell et al., 2019).

7 These would include socially or legally recognized bodies featuring one or more, but not all, of the characteristics commonly associated with the corporate form.

8 This difference was not limited to trade. For example, China relied on clans and Europe on corporations for the provision of public goods and administrative services (Greif and Tabellini, 2017).

9 It was mostly through colonial transplantation in the 19th century that the business corporation came to be used throughout much of the rest of the world.

10 As North et al. (2009: 148ff) put it, this is a ‘doorstep condition’ for the move from limited to open access orders.

11 Kuran (2011: 97ff) has argued that the absence of corporate personality in Islamic law and the resulting reliance on personal and family networks were important factors, among others, holding back the economic development of the otherwise enlightened Arab world.
12–13) defined the corporation as ‘a body politic, or body incorporate’, namely ‘a collection of many individuals, united into one body, under a special denomination, having perpetual succession under an artificial form’, vested ‘by policy of the law’ with the capacity ‘of taking and granting property, of contracting obligations, and of suing and being sued’. To the corporation’s separate capacity for property, contract, and litigation recognized by Kyd, most commentators today would add limited liability, transferable shares, delegated management under an independent board structure, and investor ownership (Armour et al., 2017). The contributions to the symposium examine the history of these features, some of which are the objects of considerable dispute.

One of the most enduring debates among business ethics, corporate governance, corporate law and management scholars is whether the primary purpose of the corporation is to maximize shareholder value. Looking at this question through the lens of agency theory, proponents of the Anglo-American shareholder-oriented model of corporate governance see directors as duty-bound to serve the interests of their principals, the shareholders, whose delegated authority they exercise. This assumes that shareholders are the corporation’s ‘owners’ and is often taken to mean that a shareholder is in fact a part-owner of the corporation’s assets. It is further assumed that shareholders are the corporation’s sole ‘members’, with the significant implication that employees and other parties are ‘outside’ the corporation. While directors may take the interests of these outside parties into account, their ultimate obligation is to promote the members’ interests.

A prominent critique of shareholder primacy focuses on the fact that, as a matter of law, directors owe fiduciary duties not to shareholders but to the corporation itself, understood as a legal person or legal entity entirely separate from the shareholders. On this view, it is this entity, not the shareholders, that owns the assets under the directors’ control, meaning that shareholders cannot be the principals on behalf of whom directors act. Moreover, shareholders cannot be said to own corporations, because as legal persons they are not things that can be owned. Shareholders do own shares, by virtue of which they have some limited residual income and residual control rights, but this does not afford them further special treatment vis-à-vis other stakeholders. On the contrary, it is possible to imagine a more participative stakeholder-oriented model of corporate governance in which directors balance the interests of a broader range of constituencies and pursue objectives other than the maximization of shareholder value.

Two contributions to the symposium engage with this debate. David Ciepley (2019) offers a strong defense of the premise underlying this critique of shareholder primacy. Ciepley traces the origins of the Anglo-American misrepresentation of shareholders as owners and members to the early days of the English East India Company (EIC), which began its life in 1600 not as a unified business corporation as we understand it today but as a hybrid between a merchant guild and a partnership. The EIC’s immediate predecessor, the Levant Company, was a merchant guild that monopolized trade with Turkey and the Eastern Mediterranean. Its members participated in the profits of trade and were united in a ‘body politic’: they elected and were represented by their own governors, regulated their own conduct, and voted on the admission of new members. But each merchant traded with his own capital, on his own account. The EIC was chartered on this basis. It had a separate legal personality, with the capacity to own assets, but did not possess a unified and permanent capital stock.

Its merchant members shouldered the higher costs of trading with the East Indies by setting up their own, separate joint-stock companies. When the companies’ ships returned, their assets were liquidated and distributed, along with the proceeds, to the merchants and their investors. This institutional structure prevented the EIC from competing effectively with the Dutch East India Company (VOC). Chartered in 1602 to coordinate the operations of six pre-existing city-based merchant partnerships, the VOC had a central board, comprising directors representing the six cities, with significant discretionary powers bestowed by the federal assembly of the Dutch Republic. It raised capital by issuing shares to outside investors, who had no say in the VOC’s governance. The funds thus raised were locked in, but shares became freely tradable in what then was the first stock market. In 1623, the VOC began issuing its own debt instruments. These features helped it accumulate capital and achieve the scale that made it the most profitable and indeed the dominant player in the spice trade.
The EIC sought to catch up by mimicking the VOC. In 1657, it was re-chartered as a unified business corporation with its own permanently locked-in capital stock. This meant that the merchants and those who had previously invested in the separate joint-stock companies as partners and co-owners became outside investors, owning solely their shares of stock, exactly like the investors in the VOC, who had never viewed themselves as members. Contrary to the VOC, however, the EIC’s new charter conceptualized the stockholders as a body politic that elected its governors, just as guild members had done. The additional provision that stockholders were periodically entitled to receive the proportional equivalent of the book value of the company’s net assets – a provision few exercised, that disappeared from subsequent charters – sustained the notion that although legal title to the assets was vested in the corporation, the assets were actually held in trust for its members.

Ciepley suggests that the English tendency to conceptualize any corporation involving more than one individual as a body politic, and the implication that the relationship between the corporation and these individuals was one of membership, has deep roots. It seems that the Roman law distinction between the universitas personarum, or ‘member corporation’, and the universitas rerum, or ‘property corporation’, which was adopted in Continental Europe following the rediscovery of the Code of Justinian, never crossed the English Channel, or was lost early on. This prevented English lawyers from seeing the post-1657 EIC as something entirely different from a body politic: like the VOC, the EIC had become a property corporation without members. As such, the authority of its governing body came from its charter, that is, from the state. This, Ciepley concludes, is the only way to conceptualize the modern business corporation as a property-holding legal entity that is entirely separate from its shareholders.

In the second contribution to the debate, Samuel Mansell and Alejo Sison (2019) point out that a problem with the focus on shareholder ownership, or lack thereof, in the argument against shareholder primacy is that it leaves unanswered the question of whose interests directors ought to serve. Unless one assumes that corporations do not have members and that the source of directors’ authority is entirely exogenous, in which case it seems that this exogenous source gets to decide whose interests directors ought to serve, it stands to reason that directors receive at least some of their authority from a body of individuals constituting the corporation’s membership. Instead of justifying the exclusion of shareholders, critics of shareholder primacy can also attempt to include non-shareholder constituencies. Yet clear-cut criteria for attributing membership are lacking. What makes an employee as much (or more) on the ‘inside’ of a corporation as a shareholder? What about other stakeholders, such as suppliers, creditors, or customers?

Mansell and Sison argue that medieval canonists and civilists grappled, mutantis mutandis, with similar issues and connected the concepts of ownership, authority, and participation in a corporation in a manner that can shed a useful light on the contemporary debate. Medieval lawyers viewed the corporation as a separate property-holding legal entity, whose assets were collectively owned by its members, conceived as a ‘plurality in succession’. Ownership did not reside with any individual member, or with the aggregate of members at any point in time, but rather with an entity abstracted from concrete members and surviving membership changes. However, as Mansell and Sison observe, neither the corporation’s status as a legal entity nor this particular idea of collective ownership unambiguously clarified which actions the governing body was authorized to pursue. Not surprisingly, much litigation across Europe between roughly the 12th and the 15th centuries concerned the authority of various kinds of corporations to act in certain ways.

The main object of contention was whether the source of the governing body’s authority emanated from the members below or was bestowed by some superior authority. By the mid-13th century, the standard answer was that the governing body was authorized in part by the members that elected it and in part by the approval of a higher authority which confirmed the members’ right to act as a corporation. In the ecclesiastical context, while a cathedral chapter could, for example, elect a bishop, the

---

12 Consider the classic (and typically English) distinction between the corporation sole and the corporation aggregate.
13 For an influential defense of the view that corporations are ‘franchise governments’, see Ciepley (2013).
powers of the office were derived from the Church, and ultimately from God. Matters were more ambiguous in the civil context. While the governing bodies of guilds, universities, and the like derived some authority from their members, their authority could be confirmed by a variety of sources, including not just town councils or kings but also Roman law, common law, or custom. In any case, members did not have unqualified authority over the corporation’s actions.

But even such partial authority that members did have implied that decisions regarding who counted as a member were crucial. Medieval corporations often had full or active members alongside passive ones. Active membership involved the participation in, and the sharing of responsibility for, the corporation’s distinct collaborative activity over time, which was equated with its ‘common good’ and understood as an abstract purpose that persisted through membership changes. While the notion of common good finds its natural expression in the Aristotelian teleological tradition, Mansell and Sison show that functionally equivalent concepts can be found in the institutional economics literature that considers the firm as a real entity underpinned by multilateral relational contracts. In sum, they conclude that non-shareholding stakeholders can be members of the business corporation. At the same time, it implies that shareholders, as members, have a duty to exercise their moral judgment and voting power with respect to decisions taken in their name.

The notion of shareholder responsibility seems to conflict with the fact that shareholders are absolved from liability beyond the extent of their equity investment. Few would dispute that this is economically beneficial. A commonplace view among financial economists and scholars belonging to the law and economics tradition is that business corporations would be unable to raise capital without offering shareholders limited liability. By reducing the risk associated with any one equity stake, limited liability allows investors to diversify their portfolios, which helps liquid stock markets develop and favors the kind of capital accumulation needed for large-scale investments. Hence the limited liability statutes enacted in Britain and elsewhere in the mid-19th century were essential for the industrial revolution. This narrative has recently been supplemented by studies suggesting that limited liability was a feature of the VOC and other early business corporations, with the implication that it also played a role in the commercial, financial, and transport revolutions of the 17th and 18th centuries.

In the third contribution to the symposium, Harris (2020b) disputes both versions of the history of limited liability. There was no limited liability at the time of the early business corporations, Harris argues, because for the liability of stockholders to really become an issue, a corporation would have to be financed by both debt and equity, corporate insolvency would need to be a real threat, and a procedure for liquidating corporations, including a rule establishing the priority of bondholders and other creditors over stockholders, would have to be available. The EIC was financed mostly by equity. After 1623, the VOC resorted primarily to long-term debt finance, but its backing by the Dutch state meant that it never really risked insolvency. In both cases, there was no significant conflict between stockholders and creditors, which might have been addressed by a creditor priority rule, and no corporate liquidation procedure had yet been evolved.

Matters did not change when the corporate economy began to expand in the 18th century. The charters of the new insurance or canal companies, for example, made no mention of limited liability. Nor did the writers of the key treatises on corporations of this period. Debt was seldom used, and insolvency was not viewed as grounds for corporate dissolution, which could only be achieved by an Act of Parliament. Paradoxically, it was the emergence of the unincorporated company in Britain that brought the issue of shareholder liability to the fore. The Bubble Act 1720 had forbidden the formation of joint-stock companies enjoying the status of separate legal persons without the assent of Crown or Parliament, but shrewd lawyers used various common law instruments to replicate the financial attributes of joint-stock capital and transferable shares. They also attempted, mostly unsuccessfully, to shield shareholders from tort and other non-consensual liabilities with respect to third parties. This underscored the value of limited liability.
Entrepreneurs petitioning Parliament for the right to incorporate in the years leading up to the repeal of the Bubble Act in 1825 began citing limited liability as a motive for incorporation. But no single idea of what limited liability might look like in practice prevailed. When incorporation by simple registration became the norm with the Joint Stock Companies Act 1844, a Winding-Up Act 1844 was enacted to enable the dissolution of insolvent companies by creditors, who were allowed to collect from shareholders. Shareholder liability was then limited to the unpaid balance on shares in the Limited Liability Act 1855 and the Companies Act 1862. In parallel developments, various liability regimes were available in the United States. To illustrate, shareholders were liable up to the double of their equity investment in New York, whereas the liability of shareholders in Massachusetts was limited to the unpaid balance on shares. On both sides of the pond, the variety of liability regimes available throughout the 19th century was even more pronounced in the banking sector.

What Harris calls ‘full limited liability in the modern sense’ became a uniform attribute of business corporations in the mid-20th century. In Britain, for example, the practice of issuing shares that were not fully paid up was still common at the turn of the 20th century. California only shifted from a pro-rata unlimited liability rule to a default limited liability rule in 1931. The shift in banking occurred later still. By this time, the view that limited liability was a game-changing institutional innovation, without which corporations and much else could not have developed, had become commonplace. Yet the evidence suggests that the absence of limited liability did not hinder the development of the most successful corporations and the most prosperous economies in the West. A richer comparative institutional analysis of the variety of limited liability rules is needed. And a better understanding of the history of limited liability, Harris concludes, can also help us imagine alternative liability regimes for different sectors and different types of corporations in the future.

Limited liability is generally cited as a key difference between the corporation and the (general) partnership.14 As any introductory law textbook will clarify, partners are personally liable for the partnership’s debts. In some jurisdictions, partnerships are also said to lack separate legal personality. However, recent legal scholarship on corporations and other organizational forms suggests that partnerships are legal entities, functionally speaking, even in jurisdictions (such as the England) where they are said to lack legal personality. This is because partnerships, like corporations, exhibit a degree of ‘entity shielding’: partnership law provides that in the event of the firm’s insolvency, the claims of the partners’ personal creditors are subordinated to those of business creditors. The corporation’s comparative advantage, from this perspective, is that it exhibits a significantly stronger degree of entity shielding. Another advantage highlighted in recent work on corporations is that control over the locked-in assets lies with an independent board.

In the final contribution to the symposium, David Gindis (2017) shows that these ideas were anticipated at the turn of the 20th century by Ernst Freund, whose recognition as an important yet neglected corporate theorist is long overdue. Freund’s The Legal Nature of Corporations, published in 1897, was written at a time when the rise of large business corporations compelled many American observers to admit that the nature of the corporation had yet to be understood. But Freund’s book was not merely an attempt to come to terms with a novel legal and economic reality. Gindis argues that it is more appropriately understood as perhaps the earliest attempt to engage in what might be called the ‘rational study of corporate law’, namely the analysis of the ends sought by corporate law and the reasons for desiring them. In the process, Freund implicitly engaged in what today would be called comparative institutional analysis.

Corporate theory in late 19th century America was at a crossroads. Prior to the liberalization of state incorporation laws, when special charters were granted primarily to ventures deemed to be in the public interest, incorporation was viewed as a state-granted privilege. On this ‘concession theory’ of the corporation, corporations were state-created artificial persons whose powers were strictly limited by their charters, with the implication that acts not explicitly authorized by the state and expressed as such in their charters were legally null and void. But as charters became available by simple

14The cases of limited partnerships and limited liability partnerships are exceptions.
registration, the existence and actions of corporations came to be viewed as products of private rights and freedom of contract. The new ‘aggregate theory’ that emerged in the last decades of the 19th century denied the significance of the corporation’s separate legal personality, and claimed that all corporate rights and duties, particularly as related to asset ownership, were in reality the rights and duties of the corporation’s members, the stockholders.

For Freund, as Gindis explains, neither perspective was satisfactory. The aggregate theory’s portrayal of the corporation as little more than a sophisticated partnership was particularly problematic. Not only did it fail to explain why individuals forming a business association would choose the corporate form over the partnership, it also downplayed the indisputable fact that law treated corporations as holders of rights and duties that are entirely distinct from any of the individuals involved. Freund was not prepared to accept the idea prevailing in Continental Europe at the time that a corporation held rights because it had a will of its own. While a corporation might well be a real entity, it was not a ‘real person’.15 In any case, metaphysical speculations of this kind were unnecessary. The solution was to approach the problem at hand from the perspective of property law.

In modern terms, Freund reasoned that members of business associations faced collective action and commitment problems that were more or less mitigated by their choice of legal form, which differed in their capacity to secure property both against insider defection and against outsiders. This led him to argue that the chief characteristics of the corporate form, namely a distinct legal personality and an independent board, offered the strong protections that contractual arrangements and partnership law could not achieve. While the analysis was rudimentary by today’s standards, it provided the first modern explanation of and justification for the separation of ownership and control. In addition to showing why the treatment of a corporation as a legal person was necessary to secure corporate property, it also explained that the treatment of a corporation as a legal person could help secure public rights against corporations, presaging the jurisdiction of courts over corporate defendants and other developments.

Overall, each contribution to this symposium sheds light on the genealogy of some of the central concepts that inform the contemporary understanding of the business corporation. The symposium’s exploration of the issues identifies several new research agendas and is presented as an invitation for further research.

Acknowledgements. This symposium has been a long time in the making. The papers ultimately collected were published online over an unusually long period of time. I am very grateful to the editors, and especially to Geoff Hodgson, for their patience.

References

---

15See Gindis (2009) for a more detailed discussion of this issue.