

University of Michigan Law School

University of Michigan Law School Scholarship Repository

Law & Economics Working Papers

8-22-2022

On Firms

Sanjukta Paul

University of Michigan Law School, sanjukta@umich.edu

Follow this and additional works at: https://repository.law.umich.edu/law_econ_current



Part of the [Antitrust and Trade Regulation Commons](#), [Labor and Employment Law Commons](#), and the [Law and Economics Commons](#)

Working Paper Citation

Paul, Sanjukta, "On Firms" (2022). *Law & Economics Working Papers*. 234.

https://repository.law.umich.edu/law_econ_current/234

This Article is brought to you for free and open access by University of Michigan Law School Scholarship Repository. It has been accepted for inclusion in Law & Economics Working Papers by an authorized administrator of University of Michigan Law School Scholarship Repository. For more information, please contact mlaw.repository@umich.edu.

ON FIRMS

Sanjukta Paul*

This paper is about firms as an instance of economic coordination, and about how we think about them in relation to other forms of coordination as well as in relation to competition and markets. Our thinking about these matters influences, of course, how we structure intrafirm coordination itself,¹ but also the *scope* of firm-based coordination within markets,² what types of *inter*-firm coordination we permit,³ and our regulation or prohibition of *other* forms of economic coordination (on pro-competition grounds).⁴ Changing how we think about firms has the potential to transform all of these regulatory directions because it is such a significant analytic unit for both law and theory.

The paper begins by explaining, at a broad level, the treatment of firms and other forms of coordination within competition law, which helps us think about how the law links up its ideas of competition with its ideas of various forms of coordination. Next, we consider the ‘base case’ of the firm, namely where the scope of coordination rights is

* I thank Steve Salop, Nathan Tankus, and Marshall Steinbaum for valuable discussions and comments, as well as participants in the University of Chicago Law Review Symposium on *Law and Labor Market Power*. In particular, Steve Salop has gone far above and beyond the call of duty in critically and sympathetically engaging the ideas contained here. I bear all responsibility for errors.

¹ For recent examples of the resurgent discussion within corporate law regarding the purposes of the corporation and workers’ participation in intrafirm decision-making, see, e.g., Matt Bodie and Grant Hayden, RECONSTRUCTING THE CORPORATION (2021), (with its own discussion of the “theory of the firm”); Anel Kovvali, Leo E. Strine & Oluwatomi O. Williams, “Lifting Labor’s Voice: A Principled Path Toward Greater Worker Voice and Power Within American Corporate Governance,” 106 *Minnesota Law Review* 1325 (2022); David Ciepley, “Beyond Public and Private: Toward a Political Theory of the Corporation,” 107 *American Political Science Review* (2013).

² Contemporary competition law’s relatively lax merger policy can not only be described as a regulatory expansion of firm-based coordination, but also was expressly understood as appropriately reflecting instances in which the presumed efficiencies of firm-based coordination outweighed market competition. See, e.g., Robert Bork, THE ANTITRUST PARADOX, 107–09 (citing Oliver E. Williamson, Economies as an Antitrust Defense: The Welfare Tradeoffs, 58 *AM. ECON. REV.* 18, 21 (1968)); see also Oliver E. Williamson, Allocative Efficiency and the Limits of Antitrust, 59 *AM. ECON. REV.* 105 (1969).

³ This is because the theory of transaction costs has been generalized beyond the context of firms to justify various forms of ‘integration through contract,’ or vertical restraints. For a recent summary, see, e.g., Callaci, Steinbaum, et al, “Vertical Restraints and Labor Markets in Franchised Industries” (working paper, July 19, 2022) (discussing Blair & Kaserman argument for liberalizing vertical restraints on efficiency grounds).

⁴ Many forms of coordination—“cartels,” or looser horizontal coordination beyond firm boundaries; labor unions; and public market coordination of various kinds—are, at least as a first cut, condemned as distorting allocative efficiency by reducing competition. Some are conditionally tolerated to fix narrowly defined “market failures” but their social benefits are not considered alongside the social benefits of firm-based coordination in a true apples-to-apples comparison (which is what the argument of this paper ultimately urges). Because allocative efficiency arguments, based in price theory, ultimately *assume* firms as the foundational unit of coordination, the argument of this paper also aims to induce a reconsideration of the reflexive presumption that these other forms of coordination represent, *ipso facto*, efficiency losses.

coextensive with the scope of operational integration (i.e., where firm boundaries correspond to the boundaries of the production unit). While generally associated with the rise of markets and competition, this section highlights how even the base case firm suppressed competition and is not obviously explained or justified on technical efficiency (as opposed to power perpetuation) grounds. Finally, the paper discusses the reigning framework for thinking about firms, which may be offered as an explanation or justification for the special treatment of firms (in both reality and theory) among forms of coordination, ultimately upon productive efficiency grounds. But power perpetuation by incumbent control groups is often a better explanation than productive efficiency for the centralization of decision-making rights and flows of incomes that seem to typify firms. And the ‘theory of the firm’ often assumes (or assumes away) contested legal rules that also affect productive efficiency outcomes, while using an impoverished conception of productive efficiency, particularly where labor effort is concerned. Moreover, the intrinsic character of competition and markets give us no good reasons to limit the normative bases for our legal choices about economic coordination to productive efficiency in the first place, much less to choose hierarchy as a unique solution to efficiency problems.

Antitrust exemptions

One way to query our thinking about the relationship between firms and markets is to consider the place of firms within the area of law that is commonly understood today to aim at perfecting market competition. Antitrust law — which in general limits, regulates, and channels both economic coordination and competition — contains a number of exemptions, or categorical authorizations of economic coordination. Foundational among them is the authorization of firms as the basic unit of economic coordination.

The firm suspends or suppresses competition in various ways. It engages in price coordination and market allocation activities that are standard focuses for antitrust law when they take place beyond firm boundaries. This coordination, and the implicit legal authorization it receives, is perhaps easiest to see when one firm displaces two or three previously existing firms in a given market, or when comparing a firm and a functionally similar cartel,⁵ but it exists more generally. To be sure, intrafirm coordination remains visible for some purposes — notably, monopolization claims, an area of enforcement that has precisely also retreated under the force of arguments

⁵ For a discussion, see Sanjukta Paul, *Antitrust as Allocator of Coordination Rights*, 67 *UCLA L. REV.* 378 (2020) (focusing on antitrust’s single entity doctrine and discussing the firm exemption in functional terms).

relating to the putative efficiencies of intrafirm coordination (even or especially within firms holding significant market power). Conversely, the fact that some coordination beyond firm boundaries is authorized does not undermine the primacy of firm-based coordination. The permissive contemporary attitude toward vertical restraints—contracts or other arrangements between actors in adjacent markets that preempt a material business decision by one or the other party (e.g. with whom to deal, or what prices to set), pertaining to a transaction other than the one between the contracting parties themselves—was indeed underwritten by the transaction cost stream of thinking that was originally generated as part of the “theory of the firm.”⁶ Nor does the existence of joint venture and related doctrines undermine the primacy of firm-based coordination. While the law has some malleable joints, it is presently not controversial that a joint venture or other inter-firm association must affirmatively justify its coordination on terms (e.g., that the association is *required* to create a particular product or benefit) not demanded of traditionally organized firms.⁷

Moreover, in its recognition of the firm exemption the law not only permits coordination but reflects and encourages a particular *form* of it in which both participation in decision-making (over prices, production decisions, the organization of production) and flows of pecuniary benefits are generally centralized.⁸ I have referred to this participation in decision-making within firms in terms of “coordination rights” in order to emphasize the continuity between the substance of what takes place within firms and what takes place—or could take place—outside them in other types of association.

Within positive law, antitrust exemptions are categorical authorizations of economic coordination. Aside from exemptions, antitrust law also frequently authorizes coordination in non-categorical ways in the application of its first-order rules (rules

⁶ Blair, Roger D., and David L. Kaserman, *LAW AND ECONOMICS OF VERTICAL INTEGRATION AND CONTROL* (1983); Brian Callaci et al, *Vertical Restraints and Labor Markets in Franchised Industries*, https://papers.ssrn.com/sol3/papers.cfm?abstract_id=4155571 (working paper 2022) (discussing Blair and Kaserman); Id. (discussing vertical restraints); see also Paul, *Fissuring and the Firm Exemption*, *L. & Contemp. Probs.* (2019) (same).

⁷ See, e.g., *Broadcast Music, Inc. v. CBS, Inc.*, 441 U.S. 1 (1979). As Steve Salop has distilled it, under current law even “production integration among association members is not enough to justify marketing integration (including joint pricing). This point is made clear in *NCAA v. Board of Regents of the University of Oklahoma*, 468 U.S. 85 (1984), where the joint marketing (i.e., price and quantity) was not permitted despite joint production (scheduling, football rules, etc.)” And as a practical matter, recognized joint ventures in the cases have frequently been between very powerful firms. See e.g., *Texaco v. Dagher*. This aspect of the law is therefore probably better understood as akin to vertical restraints and the putative “price leadership exemption” (see Tankus and Herrine, *supra*) in extending the scope of firm-based coordination rather than as a challenge to the primacy of firm-based coordination. Could the doctrine be pushed and expanded to *become* a challenge? Certainly. See Laura Alexander and Steve Salop, “Antitrust Worker Protections: Rejecting Multi-Market Balancing as a Justification for Anticompetitive Harms to Workers,” *U. Chi. L. Rev.* (2023) (this symposium).

⁸ *Id.* Antitrust law is not singularly or primarily responsible for this particular form of coordination, of course, but it is not neutral about it either: it not only reflects but reinforces it.

prohibiting some price-fixing beyond firm boundaries and authorizing others, prohibiting some sorts of domineering conduct by dominant firms and authorizing other sorts, prohibiting some corporate mergers and authorizing others, for example). It makes sense to call a set of such authorizations of economic coordination an “exemption” when it is categorical in some way, or durably appears so. The key feature of exemptions then is not that they are express but that they are categorical. So the labor exemption, the agricultural cooperative exemption, and the firm exemption are all exemptions, even though only two of them are currently expressly stated in the positive law. There may be border cases—when a first-order rule becomes exceedingly permissive (the authorization of ‘price leadership,’ for example) it may become sensible to start thinking of it as an exemption.⁹ But the distinction is useful, in part because of the ready correspondence between existing exemptions and certain relatively stable, socially recognizable forms of coordination.

While we take some of these forms of coordination for granted, an exemption from competition also implies a set of substantive criteria. Put another way, antitrust exemptions are not merely formal categories. This may sound trivial when put so—partly precisely because the law eventually tends to take for granted the social forms that it has itself also shaped—but the existence of substantive criteria implies that the exemption typically exerts some causal influence in the world, whether past or present, in shaping the type of economic coordination it authorizes. The criteria the law uses to decide whether a given instance of economic coordination qualifies for that exemption give that exemption its particular shape and character.

For example, antitrust law does not simply bless all economic coordination that takes place within the boundaries of a corporation or other entity duly formed under the law of business associations. If it did, then *any* coordinating parties could incorporate (or form a partnership, etc.) and thereby insulate their internal coordination from antitrust scrutiny. Instead, antitrust law imposes its *own* criteria to define an association that will count as a firm, or a “single entity,” and whose internal coordination is therefore insulated for key antitrust purposes.¹⁰ Similarly, the law of the labor exemption imposes (often quite exacting) demands and strictures upon labor coordination, which in turn shapes it.¹¹ The law of the agricultural exemption also sets out the types of associations and associational activity that will be authorized, shaping

⁹ Nathan Tankus & Luke Herrine, *Competition Law as Collective Bargaining Law*, in CAMBRIDGE HANDBOOK OF LABOR IN COMPETITION LAW (2022).

¹⁰ Paul, *supra*.

¹¹ See, for example, Labor Management Reporting and Disclosure Act of 1959.

that activity and those associations.¹² Exemptions contain criteria for permissible coordination and thus help to shape the social forms of coordination.

Indeed, even beyond the *current* legal criteria that accompany and define an exemption, criteria contained in earlier phases of law also likely helped to shape the directions that economic associations took. Judge-made antitrust rules soon after the passage of the Sherman Act (together with changes in state corporate law) ushered in the great merger movement and reinforced a particular template for the business firm,¹³ which had come into existence only relatively recently and also with the aid of legal developments. At the same time, incorporated farmers' groups struggled to access the basic firm exemption in antitrust's formative era, and similarly, labor incorporation would have been unlikely to have solved labor's Gilded Age antitrust problems.¹⁴ Both these types of coordination eventually gained their own antitrust exemptions (and to different extents, areas of affirmative 'coordination law'). Like the more ad hoc judicial pronouncements of yore, these early and mid-twentieth-century statutory enactments not only authorized a preexisting form of economic coordination but further defined and shaped it.¹⁵ The influence of the law upon the directions taken by the labor movement and labor associations is a canonical strand of American labor historiography.¹⁶ The legal regulation of agricultural coordination also channeled and shaped organizing among farmers in enduring ways.¹⁷

To be sure, antitrust exemptions also often correspond to social categories of coordination that have at least some independent social existence and recognition, beyond their legal construction (and because some of these forms are taken as given by theory). Labor unions, farmers' associations, and business firms have had durable social existences to various extents. That social existence can sometimes implicitly lend legitimacy to the legal exemption—and can also sometimes detract from it (think of many portrayals of the social dimensions of labor unions in popular media over the

¹² See, e.g., *Antitrust Status of Farmer Cooperatives: The Story of the Capper Volstead Act*, United States Department of Agriculture: Rural Business Cooperative Service, Cooperative Information Report (2002).

¹³ See, e.g., Naomi R. Lamoreaux, *The Great Merger Movement in American Business, 1895–1904* (1985).

¹⁴ *Reeves v. Decorah Farmers' Co-Operative Soc'y*, 160 Iowa 194 (1913); *Ford v. Chicago Milk Shippers' Ass'n*, 46 Ill. App. 576 (1892). These cases and the broader issue of the types of associations that were permitted to access the firm exemption in this formative period are discussed at much greater length in separate forthcoming work: *SOLIDARITY IN THE SHADOW OF ANTITRUST: LABOR & THE LEGAL IDEA OF COMPETITION*, chapter 2.

¹⁵ Wagner Act; Capper-Volstead Act.

¹⁶ Labor scholars have thoroughly demonstrated the influence of various phases of law upon the social evolution of labor organizations. For just two paradigmatic examples, see William Forbath, *LAW AND THE SHAPING OF THE AMERICAN LABOR MOVEMENT* and Karl E. Klare, *Judicial Deradicalization of the Wagner Act*.

¹⁷ See, e.g., Victoria Saker Woeste, *THE FARMER'S BENEVOLENT TRUST: LAW AND AGRICULTURAL COOPERATION IN INDUSTRIAL AMERICA, 1865-1945* (1998).

decades).¹⁸ At the same time, the fact that exemptions are not merely formal implies that the law also shapes the social reality that it sometimes appeals to for legitimacy, and that theory assumes.

A given authorization of economic coordination may either serve a given policy goal or subvert it. For example, in the criteria it uses to implement the firm exemption, law relies on factors that have little to do with promoting competition. Instead, the criteria antitrust law uses to permit or prohibit coordination under its firm exemption largely emphasize operational efficiency, which centralized decision-making is presumed to serve. To be sure, these criteria and their effects often get *described* in terms of promoting competition. But if that is true at all, it is in an indirect, *causal* sense, i.e., that the criteria have the *effect* of constituting the units that will then engage in competition. But a firm exemption constituted by principles of internal competition, or democratic coordination—or one that balanced operational efficiency with democratic coordination, or any number of other permutations—would *also* constitute units that engage in competition with each other. One might make policy arguments evaluating the various options among one other, but there is no logical sense in which one criterion (except perhaps the one that insists upon maximizing internal competition)¹⁹ is actually more “pro-competitive” than another.

And if the argument is that an operational efficiency criterion for the exemption *helps* a given entity compete with other entities, then it is not obvious that other criteria (such as democratic governance) could not in many circumstances aid in this as well. Moreover, it is not at all clear that the legal constitution of economic entities in a way that (putatively) maximizes each of their abilities to compete with the others is required or even recommended at all within a larger framework that prizes competition *between* the entities as a process that conduces to social goals. One might well think that entities should be internally democratic but also ought to compete with each other, for example. The usual causal pathways by which business rivalry is thought to deliver social and economic benefits do not particularly require legal criteria for constituting business and economic associations that specifically select for operational efficiency. In fact, one would suppose that the existence of sufficiently robust business rivalry would itself

¹⁸ See, e.g., *On the Waterfront* (1954).

¹⁹ It is entirely possible to imagine the law doing so. Indeed, the prospect of internal competition is sometimes volunteered by corporate leaders to defend mergers, as for instance in the current dispute over Penguin’s acquisition of Simon & Schuster. See, e.g., https://www.wsj.com/articles/penguin-random-house-ceo-defends-publishing-merger-at-antitrust-trial-11659645436?st=b0mf9bli7hfk2n&reflink=desktopwebshare_permalink

push democratically constituted entities in the direction of operational efficiency, without the law micro-managing it.

In effect, what the law has done is to internalize, into its very concept of “competition,” a particular vision of intrafirm relations, as it has grown up over time in response to social, economic, and legal factors. It is notable that law has *not* performed this same internalization activity with respect to the acknowledge purposes of other economic associations, even ones it conditionally tolerates (e.g., labor unions or farmers’ cooperatives).

Thinking about firm-based coordination in developmental terms

While the internal organization of firms is not generally described in neoclassical theory itself,²⁰ there is a kind of “base case” for the firm that functions as a kind of ideal type from which deviation must be either corrected or corrected-for; that base case can be described as a proprietary firm that makes a single tangible good in a single plant and exists in a market populated by other such firms. It is presumed that plant-level operational integration creates technical efficiencies that explain this form of coordination, and (tacitly) that justifies and legal privileges that authorize that coordination. The base case is the best example to support the firm exemption on its own terms, in that it does not involve all the later deviations from technological integration (and from single products and markets, not to mention from single plants) that many real firms involve. While discourse broadly presumes that the firm exemption is justified far beyond this base case, it is useful to consider the best case for it. Not only does the base case of the firm involve the suspension of competition just as much as the more horizontal forms of coordination it broadly replaced (which are conventionally considered anti-competitive), but it also *uniquely* suppressed competition in certain ways.

While competition is conjoined with hierarchical coordination through *technical efficiency* in many whig histories of business development and in the defense of the current status quo, in many critiques of capitalism the same conjunction is also affirmed, but understood to be mediated by *causality*. Thus, on the critical view, market

²⁰ See, e.g., Oliver Williamson, *The Organization of Work: A Comparative Institutional Assessment*, 1 *J. Econ. Behavior & Org.* 5 (1980) (“questions regarding alternative modes of internal organization do not arise naturally within, and in some respects are even alien to, the neoclassical tradition”); Aaron Gordon, Address to the American Economic Ass’n (quoted in Williamson, *id.*) (1975) (“Nor...should we forget the extent to which conventional theory ignores how and why work is organized within the firm and establishment the way it is”); Stephen A. Marglin, *What Do Bosses Do?: The Origins and Functions of Hierarchy in Capitalist Production*, 6 *Rev. Rad. Pol. Econ.* 60, 83-84 (1974) (“In the competitive model, there is no scope for supervision and discipline except for that imposed by the market mechanism.”).

competition itself generates hierarchical forms of coordination, as capitalists compete with each other to extract more from workers and the Earth. (The two explanations can also overlap.) The appearance of the modern firm as a production unit, though, calls into question both the efficiency explanations for hierarchical coordination and its connection to competition (whether that connection is glossed as positive or negative). Both reigning theory and critique of capitalism too often conjoin “markets” or “competition” with “firms.”

Development of the base case of the firm

The base case for the firm as an analytical unit of economic coordination is a proprietary firm making a single product at a single factory. This picture of the firm is implicit in the ideal theory of not only the Smithian market but also of neoclassical perfect competition. Only later did sprawling firms characterized by significant vertical integration become a common reference point—and theorists of various political persuasions theorized (or explained) them as deviations from the base case, whether they thought they required regulatory correction or not.²¹ To be clear, I am not saying that such ideal markets ever existed (even though the small proprietary firms associated with them were once much more common).²² What I want to establish here is that even in this imagined base case for the firm, a de facto “exemption” from market competition is already operating, authorizing a particular form of economic coordination instead of others. Many views—both celebratory and critical—of the system of production and commerce built upon firm-based coordination equate the rise of firms as units of production with the rise of “markets” and “competition.” But the firm-based form of economic coordination suppressed competition in ways that were parallel to earlier forms of coordination and also in ways that were distinctive. Moreover, we have very good reasons to consider or accept a power-perpetuation explanation for the rise of hierarchical firm-based coordination, rather than the technical-efficiency explanations that still predominate. And while critical views of this history are often more likely to accept the power-perpetuation explanation in place of or in addition to the technical-

²¹ See generally Alfred D. Chandler, *THE VISIBLE HAND: THE MANAGERIAL REVOLUTION IN AMERICAN BUSINESS* (1977); Adolf Berle and Gardiner C. Means, *THE MODERN CORPORATION AND PRIVATE PROPERTY* (1932). As noted, neoclassical perfect competition implies nothing about the internal organization of the firm; if anything, it assumes owner-operators and does not deal directly with multi-product firms.

²² Instead, it seems that small firms in the nineteenth century were embedded in a complex as well as rapidly changing matrix of business conventions and customs and local regulation, in short in express inter-firm (and public) coordination of markets—rather than being regulated primarily by competition, as the Smithian ideal would hold. See, e.g., WILLIAM G. ROY, *SOCIALIZING CAPITAL: THE RISE OF THE LARGE INDUSTRIAL CORPORATION IN AMERICA 189* (1997) (discussing and showing that inter-enterprise price coordination was still conventional in the nineteenth century United States); William J. Novak, *THE PEOPLE’S WELFARE: LAW AND REGULATION IN NINETEENTH-CENTURY AMERICA* (2000); Paul, *Moral Economy*, Part I.

efficiency view (though many critics adopt a version of the technical efficiency story as well), the rise of hierarchical economic coordination is frequently assimilated to the rise of “markets” or competition in critical accounts as well.

Let’s consider an example of a proprietary firm—a single-factory hat-making firm, for instance.²³ Since this type of firm had fleeting historical existence (or predominance), generally sandwiched temporally between smaller and more informal workshops on one end and vertical integration on the other, its owner likely was often once a journeyman hatter himself.²⁴ We will also assume that there is some degree of technological complexity and operational integration in this illustrative one-factory firm, relative to the guild, workshop, or putting-out systems that it succeeded. In the case of hat-making, this might consist in pouncing machines, brim-rounding machines, or fur-cutting machines, among others.²⁵ One of these industrial innovations present in many hat-making factories, a special chemical treatment processes for the fur pelts bound for hats, resulted in serious mercury poisoning of workers, a condition sometimes known as “the Danbury Shakes.”²⁶

To see the firm exemption at work in this type of base case firm, consider what it displaced, and how that system of coordination is viewed through contemporary and later eyes. At one point, most material production in the U.S. was done in the context of workshops and (in predecessor forms in the British Isles) the more formal guild, rather than in what we now understand as proprietary firms.²⁷ This was true of hat-making

²³ The example is inspired by the most famous labor-antitrust case in American history, the Danbury Hatters Case, or *Loewe v. Lawlor*, in which the Supreme Court gave its approval to the use of federal antitrust law against a number of forms of labor organizing and protest. Many then and now have argued that this approval was contrary to Congressional intent. See, e.g., EDWARD BERMAN, *LABOR AND THE SHERMAN ACT* (1930); Paul, *Recovering the Moral Economy Foundations of the Sherman Act*, YLJ (2021).

²⁴ This was true of Dietrich Loewe, a German immigrant and proprietor of the shop at issue in *Loewe v. Lawlor*. These transitions were likely messy and overlapping. For instance, even among proprietary firms in the context of millinery (the making of women’s hats) and dress-making in the United States, where women once predominated both as entrepreneurs and workers, functional production and commercial processes in many ways “exceeded the rather narrow boundaries of the commercial enterprise.” Wendy Gamber, *THE FEMALE ECONOMY: THE MILLINERY AND DRESSMAKING TRADES, 1860-1930* 3 (1997).

²⁵ Stephen A. Collins, *Two Centuries of Hat-Making: Danbury’s Famous Trade* (booklet) (1985).

²⁶ Alice Hamilton, *EXPLORING THE DANGEROUS TRADES: AN AUTOBIOGRAPHY OF ALICE HAMILTON, MD* (1942); Shirley T. Wajda, “Ending the Danbury Shakes: A Story of Workers’ Rights and Corporate Responsibility” (National Endowment for the Humanities, December 12, 2020) (accessed at connecticutshistory.org).

²⁷ See, e.g., CHRISTOPHER L. TOMLINS, *LAW, LABOR AND IDEOLOGY IN THE EARLY AMERICAN REPUBLIC* (1993); S.R. Epstein, *Craft Guilds in the Pre-Modern Economy: A Discussion*, 61 *THE ECONOMIC HISTORY REVIEW* 155 (2008). While “workshop” is loose term that extends to the putting-out system, it also includes the much more autonomous forms of production that predominated in early America. Tomlins, *id.*; Wythe Holt, *Labour Conspiracy Cases in the United States, 1805-1842: Bias and Legitimation in Common Law Adjudication*, 22 *OSGOODE HALL L.J.* 591 (1984).

among many other related trades. Conventional theory since Adam Smith has tended to view guilds and workshops as hubs of anti-competitive price-fixing.²⁸

While the guilds and their analogue workshops in early America certainly represented a form of coordination, the implicit “exemption” from competition they enjoyed was very different from the base case firm’s. Generally speaking the organization of material production in this context was looser, more horizontal, and less formalized: there were far fewer distinctions between workers and owners, and crucially, many or most workers (journeymen and apprentices) could usually look forward to becoming owners, such that much (though of course not all) hierarchy was defined by the life-cycle.²⁹ This means production units were not characterized by the centralization of decision-making and pecuniary benefits to the same degree associated with what would come to be called the “classical firm.” Moreover, firms’ displacement of workshops was not one-to-one; generally speaking, a firm absorbed several workshops.³⁰

Conversely, a firm making a single product at a single factory (the base case) represents an exemption from competition in a way that conventional theory does not seem to fully recognize. While price or other non-operational economic coordination across a set of technologically less-advanced products produced by individual artisans may be easier to *see*, it would be more accurate to say that more complex technological production *obscures* the price coordination taking place at the factory level rather than actually erasing it. Perhaps there were previously ten hatters in a given market town, making hats in a semi-communal, semi-autonomous fashion, and observing collective pricing norms. Later, let’s say we have nine hatters employed in a small hat factory by a single owner (previously a jour himself).³¹ Even if they are now making use of

²⁸ Adam Smith, *WEALTH OF NATIONS*, ___; Sheilagh Ogilvie, *THE EUROPEAN GUILDS: AN ECONOMIC ANALYSIS* (2019). Smith’s oft-quoted passage assumed to be about price-fixing in the modern sense is really about price-setting in the context of the guild system. See also Gary Richardson, *A tale of two theories: Monopolies and craft guilds in medieval england and modern imagination*, 23 *J. HIST. ECON. THOUGHT* 217 (2001), for a very helpful survey of this predominant view of the guild system, and its limitations. Other historians have pointed out the various public-facing social benefits of guilds, from quality control to innovation—not to mention maintaining the livelihoods of workers—as well as the evidence that their price-setting was socially constrained. See, e.g., Epstein, *supra*; Richardson, *supra*.

²⁹ Stephen A. Marglin, *What Do Bosses Do?: The Origins and Functions of Hierarchy in Capitalist Production*, 6 *R. Rad. Pol. Econ.* 60 (1974); Holt, *supra*.

³⁰ *Id.* One might reply that guilds (and informal pricing conventions in the U.S.) engaged in market-wide pricing. That is generally true (always within public constraints, though, which were pervasive) but it was *also* true for early proprietary firms. See, e.g., Roy, *Socializing Capital*; Paul, *Moral Economy* (discussing pervasiveness of market-wide price coordination in early U.S. and its toleration by courts, also discussing public, town-level constraints on guild pricing).

³¹ For historical accounts of this process, see e.g., Wythe Holt, *Labour Conspiracy Cases in the United States, 1805-1842: Bias and Legitimation in Common Law Adjudication*, 22 *OSGOODE HALL L.J.* 591 (1984); BRUCE LAURIE, *ARTISANS INTO WORKERS: LABOR IN NINETEENTH-CENTURY AMERICA* (1997); W.J. RORABAUGH, *THE CRAFT APPRENTICE: FROM FRANKLIN TO THE MACHINE AGE IN AMERICA* (1988).

somewhat more centralized technical apparatus (whose ownership is centralized in the hands of the owner of the firm), the suspension of competition in this production unit, in this market, is functionally equivalent to its predecessor form. Commentary, to the extent it approaches the issue at all, tends to blur the separate question of whether such technological and operational integration *justifies* the allocation of such coordination rights as a policy matter—or even whether it would be practically possible to withhold them—and whether such coordination is occurring in the first place. Indeed, regarding the question of practical possibility, factory-level price coordination is *not* strictly necessary. For example, ownership in the output of the new factory could be divided among the ten hatters, who could individually price his or her “share” however they liked. (Ownership could also be divided in other ways, but this method would preserve as many independent pricing decisions as possible.)

We may also be tempted to wave away the persistence of price coordination rights in the ‘base case’ of the firm precisely on the strength of the (also contemporaneously emerging) idea of the self-regulating market. Why? Because in theory, the price coordination rights granted to an individual production unit, an individual firm, have no impact on market prices anyway: every firm is a “price taker” under conditions of perfect competition. (Of course, this idea is not usually applied consistently to ‘cartels’ nor to analysis of predecessor production forms like guilds.) This approach is not only a refusal to take the common existence of market power seriously, but also a refusal to take the constant human construction of prices seriously. A growing literature contests it.³² On the latter view, it is not just that firms in many markets have pricing power: it is that pricing power is a constant—the question is how it is *distributed*, not whether it exists.³³

³² Tankus & Herrine, *supra*; William Boyd, Ways of Price Making and the Challenge of Market Governance in US Energy Law, 105 MINN. L. REV. 739, 820-27 (2020); NEIL FLIGSTEIN, THE ARCHITECTURE OF MARKETS: AN ECONOMIC SOCIOLOGY OF TWENTY-FIRST-CENTURY CAPITALIST SOCIETIES 27-35 (2018) (discussing the idea of markets as structured institutions with “social relations between competitors to govern competition”); FREDERIC S. LEE, MICROECONOMIC THEORY: A HETERODOX APPROACH 189 (Tae-Hee Jo ed., 2018); Robert C. Hockett & Saule T. Omarova, “Private” Means to “Public” Ends: Governments as Market Actors, 15 THEORETICAL INQUIRIES L. 53, 62-65 (2014) (discussing “market-moving on the part of [both] government instrumentalities” and private parties); Tae-Hee Jo, What If There Are No Conventional Price Mechanisms?, 50 J. ECON. ISSUES 327, 332-37 (2016)

³³ Some firms in some markets certainly look like price-takers—but in many cases this is due to concentrated power in an adjacent market (think of consolidated, well-capitalized agricultural processors buying from small farmers, for instance) rather than perfect competition in theirs. In other cases, such as decentralized markets with some other coordination mechanism (an active trade association, say), firms might also look like price-takers, but only in the sense that a voter in a large polity or a small shareholder in a large public corporation is a price-taker: That any one voice does not control the outcome does not imply that the outcome is not the product of coordination. But in either of these scenarios, the price coordination rights granted to firms of course do matter—just as governance rights, or votes, in political or corporate governance matter. It is true that market governance built upon the firm exemption is usually less

Now, what about the particular *form* of coordination that the base case firm represents? The fact that it was constituted substantially by centralization, both in terms of decision-making and in terms of flow of incomes, is not in serious dispute, regardless of normative positionality. While Smith’s contemporaneous discussion was framed in the main in terms of “division of labor” rather than (for the most part) expressly in terms of hierarchy, division of labor itself to some degree implied hierarchy in the form of a durable management role; in other ways it led causally to both concentrated coordination rights and concentrated distribution of incomes. Interestingly, a key part of that causal chain runs through reducing competition: by producing an intermediate rather than a final product, the artisan/worker was subject to a much smaller market for his or her product, and thus was more at the mercy of the putter-outer (his buyer), sometimes entirely.³⁴ And as I argue at length elsewhere, in the American context the courts’ regulation and limitation of journeymen’s organizations—the common law antecedent and template for federal antitrust law’s later regulation of labor organizations, and ultimately for its labor exemption—was expressly founded not on principles of *competition* but on the recognition of new types of *property* rights, i.e., the right of certain members of the firm to control the firm as property, without “interference” by the collective action of workers within that firm.³⁵ (The denial of coordination rights to workers under fledgling judge-made antitrust law was therefore not only an allocation of coordination rights *to* other actors, but also relied on a particular substantive vision of intrafirm coordination rather than on neutral principles of competition.)

In any case, why did this unique form of coordination eventually displace others? The conventionally accepted story is that it succeeded because it offered technical efficiency benefits, which ultimately “grew the pie” for everyone, even as both coordination rights and pecuniary benefits associated with productive activity were concentrated in fewer hands. A famous version of this explanation was given by Adam Smith using his favored example of pin-making.³⁶ There are strong reasons to dispute

formalized and more tacit than most systems of either political or corporate governance. (And it is true that a given firm’s price coordination rights may matter less or more depending on the broader context, but that is true for other types of governance as well.) For illustrations of patterns of market governance building upon the firm exemption, particularly focusing on the “price leadership exemption” and upon commodities exchanges, see Nathan Tankus and Luke Herrine, *Competition Law as Collective Bargaining Law*. We are in sore need of more empirical research dealing directly and on its own terms with existing patterns of market governance, their variations, and their interaction with the specific competitive dynamics that obtain in particular sectors.

³⁴ Marglin, *supra*.

³⁵ Paul, *SOLIDARITY*, chapter 3.

³⁶ Interestingly, Oliver Williamson pointed out that Smith failed to consider other modes of economic organization (as alternatives to independent artisan production), thereby “rigging” the comparison. Coase, however, did a version of the very same thing! Williamson, to his credit, cannot be accused of this. In his own account, he begins: “so as to avoid

this explanation. First, it is uncontroversial that the hierarchical organization of production, in the “putting-out system,” actually somewhat preceded the technological advances associated with factories, which are broadly understood to justify the hierarchical organization of the firm on technical efficiency and integration grounds.³⁷ Organizational centralization, in other words, preceded centralized technology. This does not mean that the hierarchical organization of economic activity generally associated with the firm could not have technical efficiency justifications independent of technological integration, but it does change the terrain of the debate significantly. Moreover, the technological changes that led to the industrial revolution started and was well underway under the older, supposedly inefficient guild system.³⁸ The fact that further technological change took place after both coordination rights and flows of income at the production level had been consolidated obviously does not show that this legal organization was *required* for further technological change. We do not have access to the counterfactual in which factories were co-owned by workers and managers, but we certainly cannot assume that technological change would not have occurred, or would not have occurred as rapidly. Nor can we assume, to the extent it may have occurred somewhat more slowly in certain cases, that this would not have been a net good for humanity at the time and for the planet and humanity in the long run.³⁹

The economist Stephen Marglin’s classic paper on the development of the putting-out and the factory systems, *What Do Bosses Do?*, concluded that instead of technical efficiency gains, the best explanation for emergent hierarchy at the firm level was instead simply about interested parties seeking to entrench a distributional arrangement that benefitted them. In other words, some people sought to consolidate

imputing benefits to hierarchy that can be had, in some degree, by simple nonhierarchical associations of workers, it will be useful to begin with an examination of worker peer groups.” Williamson, *Markets and Hierarchies*, 321. In his discussion of Smith, he notes: “only a single alternative to factory organization of the kind described is considered. The alternative is for each man to work ‘separately and independently,’ each pin being crafted separately, start to finish, before work on the next is begun. Intentionally or not, the comparison is thereby rigged in favor of factory modes of organization.” Williamson, *The Organization of Work*, 9. Precisely the same point can be made about Coase’s solution to the problem of transaction costs putatively posed by market-based organization of economic activity.

³⁷ See, e.g., Marglin, others.

³⁸ See generally George Unwin, *Industrial organization in the sixteenth and seventeenth centuries* (1904); Marglin; Epstein. Maybe this isn’t an accident: Adam Smith himself noted that the variety of tasks and requisite skill required in production contexts where the “division of labor” has not yet become too minute fosters a climate of innovation and invention. Smith, *WEALTH OF NATIONS* (discussed in Marglin, *supra*).

³⁹ See *supra* regarding the “Danbury Shakes,” to take just one of the tragically numerous examples that litter our history. See also, generally NATE HOLDREN, *INJURY IMPOVERISHED: WORKPLACE ACCIDENTS, CAPITALISM, AND LAW IN THE PROGRESSIVE ERA* (2020). What would have happened differently if worker participation in such decisions had been the norm at the time when business leaders chose to continue to use the mercury process in fur pelt preparation, even after its effects became evident—and when they further increased its use after World War I? Hamilton, *supra*. We might ask the same counterfactual questions about broader decision-making mechanisms as they relate to the effects of rapidly changing industrial processes on the water, the air, and the land.

and entrench (for themselves and those they viewed as their social successors) their relative gains (in both control and income terms) in durable ways, and they succeeded. Accepting this explanation does not mean waving away technical efficiency questions as unimportant, and it does not even mean that the latter were not important in shaping behavior at the time. It just means agreeing that hierarchy did not so distinctively solve technical efficiency problems –across a variety of very different sectors at roughly the same time— that neutral solutions to operational problems, rather than the human urge to consolidate power in interaction with favorable existing legal and social tools, explains its entrenchment.

Importantly, the power-perpetuation story challenges the technical efficiency story, but it also contains little to support the competition-critique or market-critique position associated with many critiques of capitalism, which would attribute this power-perpetuation to the growth of “markets” or “competition” generally (thus distinguishing it from previous forms of hierarchy). Markets of course had existed for quite some time, and power-perpetuating phases of human organization have probably vied, alternated, or coexisted with more democratic and egalitarian forms of organization for a very long time as well.⁴⁰ The explanation for the rise of firm-based coordination given by Marglin and supported by other accounts⁴¹ does not make it fundamentally different from other instances of the power-perpetuation mode: it is about people and groups who already enjoy legally and socially sanctioned power over others using legal and social tools to entrench and expand that power. This account does not say that “markets” or even competition caused that result, much less that they did so uniquely. Indeed, we already saw that one crucial foothold gained by early ‘capitalists’ was constructed by *reducing* competition through the division of labor, effectively slicing what had previously been relatively larger markets for finished goods into many smaller (and less competitive) markets for intermediate goods. Initially at least, therefore, it may not have been that expanding markets drove wages down, but that *narrowing* markets did so (while aiding in the consolidation of legal and social forms that would codify this hierarchy in decades to come).⁴²

⁴⁰ David Graeber and David Wengrow, *THE DAWN OF EVERYTHING* (2021) (extensively surveying the anthropological and archaeological evidence).

⁴¹ Roy; *supra*; Charles M. Yablon, *The Historical Race: Competition for Corporate Charters and the Rise and Decline of New Jersey: 1880-1910*, 32 *J. CORP. L.* 323 (2007) (both discussing the late nineteenth-century corporate law revolution in similar terms).

⁴² On that consolidation of legal forms, through the common law of employment, see Chris Tomlins’ masterful discussion in *LAW LABOR AND IDEOLOGY IN THE EARLY AMERICAN REPUBLIC*. (See also Karen Orren, though Orren’s account emphasizes greater continuity with hierarchical medieval forms, while Tomlins—without contradicting that

Moreover, the archetypal firm suppressed competition in other distinctive ways, further casting doubt on the idea that competition rather than hierarchy is what is distinctive about the type of economy that firm-based coordination helped to create. Some new proprietary capitalists overtly hoarded certain specific pieces of market or industry information from their workers and even from their managers, to help ensure they did not strike out on their own and become rivals.⁴³ And as Marglin points out, even without overt action the *division of labor itself helped to ensure this*: a worker trained in just one piece of a highly segmented process is much less likely to be positioned to threaten the owner or coordinator of the entire process as a potential rival.

Importantly, these suppressions of competition did not just stamp out potential rival capitalists, but potential rival *forms of coordination* as well. The “Rochdale experiments” in democratic industrial organization, which were also a template for the American agrarian antimonopoly movement, were seen by some as a threat in real time.⁴⁴ At the broadest level, we might say that the social and economic inequality that the production-level consolidations brought about itself suppresses competition. Such inequality allows a small elite to monopolize all of the economic and social resources in which potential rivals would need to get at least a foothold in order to challenge incumbents—whether as rival capitalists or through alternative forms of coordination involving dispersed decision-making and more equitable division of incomes.

In short: the firm exemption from competition exists, even if we assume only firms whose boundaries correspond to tightly integrated production units like single plants; its appearance is likely explained better by power perpetuation than technical efficiency alone (aided by contemporaneous legal developments); and firm-based coordination suppressed competition at least as much if not more than more horizontal predecessor forms did.

Theorizing the theory of the firm

It might now be objected that whatever the genesis of firms was about, contemporary theory now explains them in efficiency terms—which is sufficient to justify their central role in theory (which in turn provides the normative benchmark to guide law) and in actual regulation and construction of the economy.

element of continuity—brings out the somewhat democratic and egalitarian forms that preceded the consolidation of common law employment in the nineteenth century.)

⁴³ Marglin, *supra*.

⁴⁴ Marglin; on Rochdale’s influence in the U.S., see Woeste, *supra*, and Paul (Moral Economy). See also Lawrence Goodwyn, *DEMOCRATIC PROMISE: THE POPULIST MOMENT IN AMERICA* (1976) (on radical agrarian cooperation in the post Civil War nineteenth-century U.S., often along cross-racial lines).

Unlike some things we call “theories,” the theory of the firm generally isn't (just) an attempt to inductively generalize from the existing things in the world that we call firms, in order to give a normatively neutral account of a category that comprises them. Rather, at the broadest level it aspires to answer the general question of how best to organize economic activity, answering that question in part by reference to existing forms.⁴⁵

Theories of the firm generally assume that economic activity, to the extent it is organized at all, ought to be organized to optimize operational or productive efficiency.⁴⁶ They then go about answering that more general question by (at least as a first cut) posing a binary choice: firm or market; make or buy. Units of economic coordination—all the while optimizing operational efficiency—thus exists as islands or nodes in a market regulated by competition—all the while optimizing allocative efficiency. The primary question then becomes to further specify the essential aspects of these islands of coordination in a way that would help predict (or recommend) where the islands should end, and where the sea of contracts and competition should begin.

It is precisely this combination of three things—aiming at a general theory about economic coordination; limiting the values around which that theory revolved around operational efficiency; and drawing on limited and contingent attributes of the world in answering even that narrow question—that later leads to problems when this thinking is deployed in a normative vein to guide law and policy (as it has been in competition law, for example).

Assuming for the moment the singular aim of efficiency, the theory of the firm generally narrows the economic coordination mechanisms among which to choose to firms and markets—hence the shortening of the general query to “make or buy?”. Given the understanding of markets as devoid of coordination implicit in most theorists’ views, this could also be restated as “to coordinate or not?” Contained in this binary, in other words, are two important background assumptions: that unconditioned competition can be posed as an alternative to coordination, on the one hand, and that

⁴⁵ This is why some of its conclusions have been extended to forms of economic coordination beyond the firm (notably to contractual relations often known as “vertical restraints” in competition law). See, e.g., Blair & Kaserman, *supra*.

⁴⁶ There are certainly discussions of the firm in the academic literature that don't do this, or in other ways may not bear out the other general attributes I discuss here. I refer here to a relatively well-defined stream of thought sometimes known as “neo-institutionalist” or “transaction cost” theory that begins with the work of Ronald Coase and is continued by key figures such as Oliver Williamson, Armen Alchian and Harold Demsetz, Oliver Hart, and others. Certainly, boundaries of conversations are not fixed, and this is particularly true of a stream of thought that that -- while it does have defining features -- is not defined by deductively binding axioms in the same way that some other aspects of economic theory (to which it is nevertheless linked) are. Yet these theorists have influenced law, and legal thinking, in enduring ways that I would argue flow from the features identified here.

firms (and perhaps units of coordination generally) represent hierarchies or coordination-by-command, on the other.

The “make or buy” binary may make some sense for household economic decisions (thus motivating our intuitions), but at the level of firm-to-firm relations, it belies the networks of coordination between commercial and industrial enterprises that exist and have existed—to say nothing of alternative forms of coordination and production that are *possible*. What if “cartels” of smaller producers and service-providers were permitted to coordinate above-ground in a way that made more robust inter-firm investments and collaboration (including on operational matters) feasible⁴⁷—while price leadership by dominant firms and coordinating activity by commodities exchanges in more dispersed markets was more strictly scrutinized by competition law? What if labor unions were permitted by labor law and by competition law—and had built up institutional memory, through the institutional structure that the law helps to create—to coordinate not only regarding wages and working conditions but also regarding prices, operational decisions, and more? In short, how would the literature on the “efficient” balance between the costs of various forms of association look different if the law governing them was different? By the time the conversation we now know in terms of “the theory of the firm” began to get off the ground, this basic structure was essentially in place (in American law).⁴⁸

It is understandable to look around oneself at the existing world in an attempt to understand why it is the way it is, and even, in various ways, why it should be that way. I do not dispute that the theory of firm literature (or the “neo-institutionalist” literature, or transaction cost literature) has yielded many useful insights that can be carried forward as we think about how to organize economic activity. I also do not, however, think it provides a basic justification for the legal and social structure of economic coordination it assumes—centrally, the analytic and regulatory primacy of firm-based, hierarchical coordination.

⁴⁷ I am not necessarily endorsing this, particularly not among already powerful firms, but am suggesting that assessment of operational efficiency tradeoffs could look very different if such coordination were permitted. For a case in favor of permitting joint bargaining among small and medium-sized businesses from a labor perspective, see Tess Hardy & Shae McCrystal, “Bargaining in a Vacuum? An Examination of the Proposed Class Exemption for Collective Bargaining for Small Businesses” (2020) 42(3) *Sydney Law Review* 311 (on the new rule by the Australian competition agency providing for such joint bargaining).

⁴⁸ I develop these points in detail in separate work (*SOLIDARITY IN THE SHADOW OF ANTITRUST: LABOR & THE LEGAL IDEA OF COMPETITION*), calling this structure the New Deal settlement on economic coordination. I also argue that the theory of the firm is a vein of theorizing economic coordination that grew up in the soil of the New Deal legal settlement on economic coordination while eventually underwriting normative recommendations that helped to undo that settlement. I am not telling either aspect of that developmental story here. Here, I am focusing on these theories of economic coordination on their own terms.

The remaining discussion attempts to survey key moves in this literature in terms of whether they may provide justification for the preferred legal (and analytic) treatment given to coordination within firms. This will necessary cover well-trodden ground, and far too briefly and summarily for anyone's liking. Yet it is necessary to take a partial step back in order to see the picture this essay seeks to bring into view. Following a bit of context, I consider three primary justifications for the primacy of hierarchical, firm-based coordination: the costs of too-little labor effort ("shirking" or "malingering"); the relative costs of democratic association; and the costs of "hold-ups" (insufficiently large markets, for instance in case of complementary assets or simply intermediate inputs, leading to opportunistic bottlenecks). I conclude that these reasons are insufficient, and that an attempt to leverage the theory of the firm literature in favor of the primacy of firm-based coordination fails to consider other important reasons.

A note on genesis & context

When the institutionalist economist Walton Hamilton catalogued the relationship between schools of economic thought for the *American Economic Review* in 1919, he identified both neoclassical economics and (early) institutionalist economics as heirs to different aspects of the classical economics of the previous century: neoclassical theory was essentially a refinement and formalization of the concept of the self-regulating market in classical economics, while institutionalism was the continuation of its focus upon elucidating the legal and social—institutional—structure of markets, and their resultant qualitative characteristics.⁴⁹ These two streams of thinking and research were distinct but overlapping (in terms of social ties and to some extent in terms of political commitments) in the early twentieth century. As institutionalism's influence and presence waned in the United States by mid-century, a new current of thought staked out a place within institutionalism's erstwhile domain as Hamilton had earlier identified it.⁵⁰ But that newcomer took a much less ambivalent attitude toward neoclassical theory than the old institutionalists had.

⁴⁹ Hamilton, Walton H. "The Institutional Approach to Economic Theory," 9(1) *The American Economic Review* 309 (1919).

⁵⁰ In fact, many strands of it came to be known as "neo-institutionalism." It was not the only current within this domain: as Beth Popp-Berman and William Kovacic both describe, the sub-field of "industrial organizations," then dominated by the "Harvard School," already incorporated some of the features of the early institutionalists as well as neoclassical analysis in their accounts of markets and organizations. William E. Kovacic, *The Intellectual DNA of Modern U.S. Competition Law for Dominant Firm Conduct: The Chicago/Harvard Double Helix* (2007); Elizabeth Popp Berman, *THINKING LIKE AN ECONOMIST: HOW ECONOMICS BECAME THE LANGUAGE OF U.S. PUBLIC POLICY* (2022). Ironically, thinkers like Coase and Williamson pushed further beyond neoclassical analysis into essentially internal and qualitative features of economic organization, even as they understood their project as complementary to neoclassical analysis.

As previously noted, neoclassical price theory itself at best does not define the internal organization of its fundamental units of analysis—firms—and at worst assumes them away as infinitely small or effectively individual producers.⁵¹ It is this black box, or black hole, at the center of price theory into which Ronald Coase’s germinal 1937 paper enters. It is the resultant stream of thought—the theory of the firm, focusing on transaction costs and productive efficiency—that forms much of the intellectual core of “Chicago School” developments in competition law.⁵² Even though both price theory and its accompanying organizational theory employ the term of art “efficiency,” the productive efficiency invoked by the theory of the firm and the allocative efficiency of neoclassical price theory are far from equivalent, and how the concepts are to relate to each other is in many ways ambiguous.⁵³

Very broadly, the paradigmatic firm of theory *centralizes* or concentrates both decision-making and flows of incomes (whether by centralizing ownership of assets, or in some other way). The degree to which this is true will obviously vary (and individual situations of reversal will arise), but overall this is what firm-based coordination seems to have meant in contrast to other broad visions of economic coordination, both in transition (see last section) and in the twentieth-century literature that conceptualizes it. Alchian and Demsetz defined the “classical firm,” the organization they wanted to explain, in the following terms: joint production, coordinated by a central party who enters into contracts with all “input owners,” who holds the residual claim, and whose

⁵¹ See, e.g., Oliver Williamson, *The Organization of Work: A Comparative Institutional Assessment*, 1 *J. Econ. Behavior & Org.* 5 (1980) (“questions regarding alternative modes of internal organization do not arise naturally within, and in some respects are even alien to, the neoclassical tradition”); Aaron Gordon, Address to the American Economic Ass’n (quoted in Williamson, *id.*) (1975) (“Nor...should we forget the extent to which conventional theory ignores how and why work is organized within the firm and establishment the way it is”); Stephen A. Marglin, *What Do Bosses Do?: The Origins and Functions of Hierarchy in Capitalist Production*, 6 *Rev. Rad. Pol. Econ.* 60, 83-84 (1974) (“In the competitive model, there is no scope for supervision and discipline except for that imposed by the market mechanism.”).

⁵² An assumption of robustness of potential competition, as a disciplining factor on current market participants, was the other key intellectual component of this policy influence. The study of economic governance within both firms and markets that is usually identified as “industrial organizations economics” predates Williamson, though his work transformed it. Broadly speaking, prior to his influence the field was both defined more by neoclassical price theory than by Coasean transaction costs considerations, and focused more on market structure than enterprise governance. For a broader discussion of the origins and history of industrial organizations, or “I/O,” Popp Berman, *supra*, at Chapter 4.

⁵³ See *Antitrust as Allocator* for a discussion in the context of Bork. It is my broader contention that both wings have worked together to favor hierarchical coordination: price theory by condemning various forms of democratic association as anti-competitive, and the theory of the firm by elevating firm-based coordination (and some other forms of hierarchy) as offering productive efficiency benefits that often outweigh competition considerations. (This is the famous Williamson diagram showing how merger efficiencies can outweigh “deadweight loss,” later adopted by Bork.) The applicability of neoclassical price theory to bar horizontal coordination beyond firm boundaries, and thus many relatively democratic forms of economic coordination, is straightforward on its own terms. The force of transaction cost theory in justifying the classical firm (as well as its relationship to price theory), which is discussed here, is less deductively straightforward.

bundle of rights is alienable.⁵⁴ They contrasted what they dubbed this “classical firm” to various other sorts of economic organizations, including, in their words: partnerships, nonprofit enterprises, labor unions, and “socialist firms.”⁵⁵ But it would be a mistake to pin this particular configuration of economic coordination—and its accompanying privileged place in law—*only* upon transaction cost theorists or upon Chicago School-associated thinkers. A version of it was baked into the New Deal settlement itself, a fact that is evident within the key legal developments that constituted the foundations of that settlement, and also in contemporaneous intellectual work that was influential in New Deal reform.⁵⁶ Transaction cost analysis can be understood as an important influence in the latter half of the century that effectively extended and purified this already-existing legal and economic preference for firm-based coordination, and also helped to discourage competing forms of economic organization as much as possible.

Perfect competition as a normative benchmark, though it does not provide a logical *foundation* for transaction cost analysis or for the concept of efficiency it elevates, continues to shape this line of thinking. Most obviously, neoclassical price theory’s ideal of a self-regulating market, governed only by competition, naturally suggests forms of economic coordination that are contained rather than diffuse—*units* of coordination, like firms, rather than broader patterns of coordination that could be market-wide in scope. More, it may lead us to overlook the price coordination rights allocated to particular forms of organization, because in perfect competition the price that all firms ‘choose’ just is the market-clearing price.

The two analytical frameworks (price theory and transaction cost theory) are sometimes simply assimilated into one another within descriptive accounts of the development of economic thought or its influence upon policy. For example, Williamson’s influence upon antitrust is sometimes assimilated to the ascent of “neoclassical” economics’ policy influence. Yet Williamson himself characterized his influence in the field as bringing in questions about “organizational design” that were

⁵⁴ Armen A. Alchian and Harold Demsetz, *Production, Information Costs, and Economic Organization*, __ Amer. Econ. Rev. 777, 794 (1972).

⁵⁵ *Id.* at __.

⁵⁶ Defining New Deal legal developments, including the Wagner Act, internalized the firm exemption in structural ways, and influential contemporaneous commentary (e.g., Berle and Means, *supra*) dramatized a kind of transition from antimonopolist contestation about the basic forms of economic coordination to the acceptance of hierarchical firm-based coordination as part of the New Deal bargain. This theme is discussed at greater length in forthcoming work (*Solidarity*).

“outside the canon” of the “price-theoretic” approach that dominated antitrust agencies when he came to them in the mid- to late 1960’s.⁵⁷

Transaction cost analysis overlaps with price theory insofar as imperfect competition is typically the opening that makes it relevant in the first place: “bilateral monopoly” or small-numbers dealing provide one of the primary reasons for coordination through ownership, employment, or other ‘non-market’ association rather than through contracts ‘in the market.’ But from there, as it begins comparing types of association—and even to some extent as it compares association with market contracting—this current of economic theory mostly departs from the realm of neoclassical price theory altogether, drawing upon a variety of ultimately qualitative observations and inferences regarding human behavior, group dynamics, economic and social history, and the comparative characteristics of various occupations, trades, and types of economic activity in order to compare solutions to ‘coordination problems’. Its key conclusions certainly do not follow deductively from price theory or even from some distinct set of shared commitments.⁵⁸

Nevertheless, important connections to the perfect competition framework exist, and other defining tendencies exist. In Coase’s foundational account of the firm, he undertook to explain the presence of firm-based coordination in a market economy. He began by observing that organizing production through “the price mechanism” on the open market has costs: specifically, the cost of discovering the ‘true’ prices of inputs, and the time/effort cost of negotiating contracts for each individual input.⁵⁹ This is his original definition of transaction costs. The second type of transaction cost is an empirically observable one that does not depend upon any given theoretical framework—it may take more time and energy to negotiate individual contracts for every step of production than it does to employ people who can handle or alternate between several functions, for example. But the first type of transaction cost (discovering the prices of inputs) has no existence at all apart from the theoretical perfect competition benchmark. That is, if there are no “true” equilibrium prices given

⁵⁷ “Although the leadership and staff of the Antitrust Division in the late 1960s were both superlative, the prevailing attitude toward nonstandard and unfamiliar contractual practices and organization structures was that such “abnormalities could be presumed to have anticompetitive purpose and effect.” Indeed, given that the prevailing price theoretic orientation effectively disallowed economies of a non-technological kind, it could hardly have been otherwise. That economies could result from organizational and contractual design was simply outside the canon.” <https://www.nobelprize.org/prizes/economic-sciences/2009/williamson/biographical/>

⁵⁸ This is evident in the nature of the disagreements among theorists in this tradition: for instance, Hansmann draws very different conclusions regarding the management of “shirking” within various organizational forms, and across different types of work, than do Alchian and Demsetz.

⁵⁹ Ronald Coase, *The Nature of the Firm*, 4 *ECONOMICA* 386, 390-91 (1937).

by perfect competition, then discovering the true price of an input is no longer a cost at all.

Moreover, the focus upon maximizing output and, somewhat less straightforwardly, upon minimizing costs is a shared focus connecting the two frameworks. But this shared focus also highlights a basic divergence. In a sense, the defining feature of neoclassical price theory, both on its own technical terms and in terms of its wider influence beyond strict professional borders, is its explanatory focus upon market dynamics *external* to the enterprise or individual—competition—in determining certain outcomes, whether those outcomes are framed in terms of output or costs or prices or something else. The broader, common-sense interpretation of this general idea is fundamentally a causal one: the discipline of market competition induces participants in the market to organize and manage themselves—whether through technological innovation, internal organization, or simply self-cultivation—in ways that save costs, increase quality, and/or increase output. (In this ordinary-language sense, many who do not necessarily subscribe to neoclassical price theory do believe that economic competition has great value, even if it is not the *only* important organizing principle for economic systems.) Yet the formalized version of this idea, within price theory, also collapses time horizons in a way that tends to erase any space for causality. That is, in a perfectly competitive market, managers or workers in firms do not really innovate, manage, or work their way to particular price or even quality outcomes.⁶⁰ Instead, there simply *is* a market-clearing price, and all firms are price-takers. To the extent that quality can be translated into price terms (or just understood as product differentiation, putting the product in a different market altogether), it is also just another given within this analytical framework. Technological development, too, is considered “exogenous.” Even putting this aside and adopting a dynamic conception of competition, it is not clear why we *need* to solve internal coordination problems through organizational choices if competition determines them.

The analysis of qualitative forms of organization in the tradition of Coase and Williamson *does*, of course, delve into these negative spaces left by perfect competition. That does not automatically mean that they constitute complementary halves of a coherent framework, though the structure of our regulatory frameworks can be read as assuming so: *inside* the enterprise, the key goal is to optimize productive efficiency and solve ‘the agency problem’, while *outside* the enterprise the key goal is to promote

⁶⁰ This point is nicely made by Howard Botwinick. PERSISTENT INEQUALITIES: WAGE DISPARITY UNDER CAPITALIST COMPETITION 25 (1993) (quoting a “widely used micro textbook” that overtly contrasts “business executives’” notion of a competitive market and economists’, concluding that the latter involves “no competition at all”).

competition. The prominence of the firm-market divide in our basic map of regulatory frameworks—the divide between intra-enterprise and inter-enterprise coordination—also then naturally accompanies the conjunction of these two influential modes of thought.

Labor effort, technical efficiencies, and outputism

Coase posited that a particular type of economic coordination—the type embodied in the employment relationship, as structured by master-servant law—arose to solve the problem of transaction costs. Coase’s explanation of firms therefore relied crucially upon labor regulation. In so doing, it also took a particular, contingent legal form for managing labor as given. The employment relationship, in contrast to contract, defines the essence of what firm-based coordination (in contrast to market-based coordination) *is* for Coase.⁶¹ But Coase’s account assumed a selectively simplified picture of the legal form of employment—essentially a version of common law employment.⁶² The essential element of this legal form for Coase’s purposes was the command relation inherent to master-servant law, and thus to the legal relationship of principals to agents. In short, for Coase the employment relationship represented economic coordination through command, which in turn defined the firm.

The central role of labor within transaction cost analysis continued in the formative work of Oliver Williamson, Coase’s intellectual heir. Williamson argued for the indispensability of hierarchical economic organization (even while later recalling quite fondly his own upbringing in the “most democratic” milieu of a prairie town shaped by populist and egalitarian traditions).⁶³ Williamson began where Coase left off, first restating the Coasean point that market-based coordination incurs costs. The limited “computational capacity” of humans entails that contracting can be costly.⁶⁴ Moreover, humans may not always be honest, and they may be influenced or motivated by non-pecuniary factors, which Williamson sometimes interestingly refers to as “atmosphere” or “energies.”⁶⁵ Partly as a result of these human factors, market contracting is also often subject to asymmetric information and to the “small numbers”

⁶¹ Coase, *id.*

⁶² *Id.* For a more recent and even more express endorsement of taking this simplified version of common law labor regulation as fundamental to structuring markets, see Richard Epstein, “The Application of Antitrust Law to Labor Markets — Then and Now,” *NYU J. Law & Lib.* (2022).

⁶³ Nobel bio.

⁶⁴ *Markets and Hierarchies: Some Elementary Considerations*, 63 *Amer. Econ. Rev.* 316, 317 (1973).

⁶⁵ *Id.* at 317; Williamson, *The Organization of Work: A Comparative Institutional Assessment*, 1 *J. Econ. Behav. & Org.* 5, — (1980) (“energies”).

problem (another way of saying that situational monopoly or market power can arise through relationship-specific or transaction-specific knowledge or capital).⁶⁶

Where Williamson went beyond Coase is that he did not take hierarchy, as embodied for instance in master-servant law, *for granted* as a solution to the “transaction costs” that attend market-based coordination. Instead, he expressly named and considered the possibility of “non-hierarchical associations of workers” as an alternate solution to the problem of transaction costs posed by market-based coordination.⁶⁷ However, Williamson importantly posited that just as transaction costs may attend market-based coordination, coordination through *association* risks incurring certain costs as well—a problem better solved by some forms of association than others. He then argued that hierarchical organization is ultimately superior to non-hierarchical organization in solving the costs that attend association, and thus is usually more productively efficient.

For Williamson, one of the central problems of economic associations is the prospect of “malingering” by workers.⁶⁸ (While he posits other types of “opportunism” that may arise in multi-lateral economic activity as well, this particular problem is essential for generating the particular solution of *hierarchical* organization associated with the classical firm.) He argues that the superiority of hierarchy over “nonhierarchical association of workers,” for purposes such as pre-screening workers (with respect to skills and propensity for effort) and for policing “malingering and other *ex post* manifestations of moral hazard,” accounts for hierarchy’s observable frequency and explains its superior productive efficiency.⁶⁹ Importantly, this fundamentally empirical argument about *work and workers* is at the heart of Williamson’s broader views about economic organizations and markets, and of its implications for antitrust and related areas of law.

Williamson went to lengths to emphasize that questions of human effort and human decision-making (including shirking), rather than technology or technological change, were the ultimate basis for the transaction cost analysis of the firm (and other

⁶⁶ Id. at 318.

⁶⁷ Id. at 321.

⁶⁸ Id. He characterizes the more general problem of association as “opportunism,” which may be exhibited both by controllers of firms, in their dealings with each other, and by individual workers toward the enterprise. Note that “opportunism” can arise for Williamson both in market dealings (thanks to situational market power) and in associations.

⁶⁹ Id. at 321-24.

economic organizations).⁷⁰ Fellow travelers like Alchian and Demsetz — while they began from a critique of the Coase-Williamson account, emphasizing the contractual nature of the firm — ultimately also concurred in the focus upon “shirking” and insufficient effort by workers, and specifically in explaining the superior productive efficiency of the organizational structure associated with the “classical firm” in most circumstances.⁷¹

The prospect of shirking is, ultimately, one of the key reasons for favoring hierarchy over democracy in industrial organization, on these accounts. Yet there is no clear conceptual space for the converse problem: *too much*, or greater-than-optimal, effort. This is partly due to the lack of any developed concept of the real, objective, and often physical costs of labor effort. Instead, the costs of labor, to the extent they are acknowledged at all, are typically acknowledged in terms of purely subjective “disutility,” as opposed to the objective sense of cost that is at least available when discussing business costs.⁷²

The prominence of shirking costs and the lack of space for overwork costs is also shaped by the focus on *output* that transaction cost analysis shares with ideal price theory, and that links them in them in various areas of policy analysis, just as antitrust.⁷³ John Newman has gone so far as to label contemporary antitrust’s fixation on output effects “outputism.”⁷⁴ In fact, while there might, just, be a way to make sense of *too much output* in a perfect competition framework — allocating too many social resources to a particular sector of production — there is no straightforward way to accommodate the idea of excessive output in a transaction cost framework focusing on a particular firm or sector, which largely revolves around output as an indication of productive efficiency. It is no surprise, then, that the current antitrust framework, shaped as it has been by transaction cost theory, has had so much trouble making sense of too much work for too little remuneration as a problem.⁷⁵

⁷⁰ Id. at 316 (“the interesting problems of economic organization are mainly to be explained by reference to the conjunction of a set of human attributes with a related set of (largely nontechnological) transactional factors ... Discussions of economic organization nevertheless are frequently dominated by references to technology.”)

⁷¹ Alchian and Demsetz, earlier, 784. They do predict that worker ownership will be more likely where the costs of supervision (policing shirking) are high, which they identify as the situation of many professional service firms—where worker ownership in fact predominates—in contrast to, e.g., supervising loading work by dockworkers. See discussion of Hansmann, below, who inverts this supposition.

⁷² See, e.g., Williamson, *The Organization of Work*, earlier, 10.

⁷³ Robert Bork, THE ANTITRUST PARADOX, 107–09 (citing Oliver E. Williamson, Economies as an Antitrust Defense: The Welfare Tradeoffs, 58 AM. ECON. REV. 18, 21 (1968)); see also Oliver E. Williamson, Allocative Efficiency and the Limits of Antitrust, 59 AM. ECON. REV. 105 (1969).

⁷⁴ John Newman, *The Output-Welfare Fallacy: A Modern Antitrust Paradox*, 107 Iowa L. Rev. 563 (2022).

⁷⁵ On this tendency in the law, see, e.g., Hiba Hafiz, *Labor’s Antitrust Paradox*, 86 U. Chi. L. Rev. 381 (2019)

This underlying approach to output and labor effort does not change fundamentally even within strands of the neo-institutionalist tradition that are much friendlier to democratic and worker-controlled organizations. Henry Hansmann's work (discussed further below) ultimately concluded that despite the time and effort costs of democratic decision-making, worker-owned firms will be productively efficient in a wider range of circumstances than theory had previously assumed, but the conceptual role of labor effort in his account remains similar. Hansmann noted the advantages of worker-run firms in monitoring shirking, but the primary associational problem where work is concerned is still shirking—and not overwork. The idea of too much effort or too much work does not enter in.⁷⁶ This feature seems to be endemic in the transaction cost strand of thinking.

To be sure, alternative conceptions of productive efficiency are possible, and indeed desirable as policy goals. True technical efficiencies consist in deriving more output while holding inputs—including labor effort—fixed. They do not consist in increasing output simply by *increasing inputs*. Yet our dominant stream of thought for conceptualizing economic coordination frequently runs together (particularly in applied or policy discussion) the question of technical (or productive, or operational) efficiency with the question of maximizing output by increasing input. This analytic tendency seems to be exaggerated in the case of labor effort, given the evident difficulty of imagining too much of it. Imagine if more labor effort were substituted for more of some other input: land, or raw materials, or investment capital. While outputism might still lead us to be overly sanguine about such deployments of resources, we are somewhat less likely to mistake increases in those inputs (even where they ultimately increase output) with technical efficiency. Yet we seem to do just this frequently in the case of labor effort.

It may be high time for theories of the firm to recognize, and be guided by, 'the sweatshop problem'—the persistent occurrence of overwork—as a problem for law and policy to navigate, right alongside the agency problem and the shirking problem. But it appears that conceptual slippage between technical efficiency and increased labor effort, together with the failure to account for the possibility of too much labor effort, is woven into influential streams of thinking about how to choose between forms of coordination.

⁷⁶ Contrast the neo-institutionalist approach to labor effort—where it is at best subjective dis-utility—with both Brandeis and the original institutionalists, who recognized the substantive efficiency implications of *too much* labor effort, particularly on a social scale. See, e.g., Sidney and Beatrice Webb, *INDUSTRIAL DEMOCRACY* (1897); Louis D. Brandeis [].

Costs (and benefits) of democratic association

Some recent thinking takes a more friendly view toward democratic governance of enterprise. Henry Hansmann’s study of worker- and producer-owned enterprises focused on ownership as a proxy for control, concluding that worker ownership may be productively efficient in a number of circumstances.⁷⁷ Capital, like labor and like other inputs, is on this framework just another factor of production: shareholders are providing capital to the firm, just as workers are providing labor.⁷⁸

Hansmann then proceeded by considering whether one mode of organization may have more productive efficiencies than others, depending on the circumstances.⁷⁹ Like many other theorists in this tradition, he tends to infer that existence, persistence, and frequency at least roughly indicate superior productive efficiency, absent legal or other bars. Like Williamson, Hansmann supposed that association, and specifically association through ownership, is most viable where competitive markets are not. And for him, as for Williamson, the costs of a particular type of association in comparison with another type, and with the costs of market-based transactions, are a key factor.⁸⁰ Also similar to others, his approach predicts that worker-owned firms are superior and will therefore be more frequent in those situations where the assignment of ownership to workers (as opposed to some other set of the firm’s “patrons”) results in a net savings in costs, considering both the transaction costs of contracting and the costs of ownership (or association). In other words, worker governance of enterprises is most viable where contracting with workers in the market is particularly costly (in comparison with contracting with other firm “patrons”) or where the associational costs of worker ownership are particularly low.⁸¹

Hansmann ultimately put less stock in “shirking” and “malingering” as determinative problems of association than did Williamson, Alchian, and Demsetz, and more stock in the simple time and effort costs of democratic and horizontal decision-making. He in fact starts by supposing that worker-owned firms actually offer

⁷⁷ Hansmann largely assumes that ordinary business corporations are owned by shareholders, and also that these shareholders have “formal” though often not “effective” rights of control. *But see* Lynn Stout (even in investor-oriented firm, shareholders do not in fact own the corporation). Hansmann also treats the concept of “ownership” as coextensive with, or even defined by, internal governance rights. Hansmann 19-20.

⁷⁸ In a way, Hansmann actually recedes further than Williamson—who is quite frank about hierarchical relations, and expressly engages the question of power, though he eventually rejects it as explanatory—from the consideration of power as determinative of organizational relations.

⁷⁹ Hansmann largely dismisses the possibility of multiple classes of firm stakeholders sharing internal governance rights, on the ground that the costs of collective decision-making would be too immense given the divergent interests involved. Hansmann at 44.

⁸⁰ Henry Hansmann, *THE OWNERSHIP OF ENTERPRISE* (HUP 1996).

⁸¹ *Id.* at 20-21.

comparative productive efficiencies where shirking is concerned, in that they give workers an incentive to police each other's shirking unlike in investor-owned firms. He also argues, in opposition to Alchian and Demsetz, that the difficulties of policing shirking are actually much *more* pronounced in the context of complex, capital-intensive production (e.g., traditional manufacturing operations) than in the context of services, where an individual's contribution to the firm's earnings is much easier to isolate.⁸² Although these two inversions put Hansmann in a position to claim that minimizing shirking is an explanation for the incidence of worker-owned firms (given their frequency in the context of professional service firms), he primarily emphasizes worker ownership is most efficient, and most likely to arise, where the costs of collective decision-making among workers is the lowest. These he identifies as contexts in which workers within the firm are relatively homogenous in terms of the type of work they do, where their status within the firm (and one might add, in society more broadly) is relatively equal, and where there is a relative lack of hierarchy or supervisory relationships between them.⁸³

Notably, the second and third of these three factors (flatness of hierarchy and status among workers) and arguably also the first (which implicates the division of labor) would seem to be constitutive of the very questions at hand. In many ways, the professions (such as law), where worker ownership is still relatively common, still *do* involve relatively horizontal relations, both within firms and within the profession as a whole, with differences largely defined in terms of experience and seniority (i.e., largely across the life cycle rather than across persons, broadly speaking). But this was also once true in other contexts too—as Wythe Holt said of American apparel makers in the early nineteenth century, for example.⁸⁴ One can quite easily imagine things being much more hierarchical and specialized in the legal services sector—entailing further specialization and greater status distinction both within firms and across the profession, and a more unequal distribution of revenues within enterprises. And one can also imagine things being much less stratified in apparel making. There is no natural or technological inevitability to the *vastly* different social and economic statuses currently occupied by fashion designers and garment workers, for example; we certainly should not assume this is any more inevitable than the relative social parity between law partners and associates is.

⁸² Id. at 70-71.

⁸³ Id. at 91-92.

⁸⁴ See *supra*.

In fact, the professions today represent in a way the limited survival of guild-like economic organization, which once obtained throughout manufacture as well.⁸⁵ It is far from clear that the relative prevalence of worker-owned (or controlled) firms in the context of professional services today should be explained in terms of neutral, objective differences in the heterogeneity or homogeneity of the work involved, rather than by independent social and legal evolution. It is also worth considering whether the relatively privileged strata of society, who retain more power to choose or at least influence the economic organization of our work, also tend to choose relatively horizontal organization for ourselves, our peers, and those we view as our descendants or successors in social terms.⁸⁶

Finally, one wonders whether the more direct benefits of more democratic economic governance are overly discounted in the neo-institutionalist stream of thinking. Williamson at one point refers to these as “energies” that some workers (or participants in the economic activity, more broadly) may prefer. But there are also often objective, substantive benefits for productive efficiency when all participants in a productive process are able to contribute their insights and experiences to decision-making that will direct that process.⁸⁷ Moreover, even the “subjective” benefits of shared governance are also likely understated in a framework that views labor effort mainly in terms of subjective dis-utility, rather than something anyone in society (rather than only a select few) can under the right circumstances find fulfillment in.⁸⁸

⁸⁵ Tomlins, Epstein, *supra*.

⁸⁶ One might even think of this retention of flatter decision structures and income distributions in the professions as an aspect of what David Graeber called “the communism of the rich,” by which he just meant that the sort of moral economy framework—wherein economic activity was mediated through thick social ties, and in which a ‘social safety net’ was woven in—that once suffused traditional societies effectively persisted in many ways in the upper strata of modern commercial societies. David Graeber, *DEBT: THE FIRST 5,000 YEARS* (2011).

⁸⁷ A number of lines of existing empirical inquiry, particularly comparative, can help inform a broader consideration. See, e.g., Jan Ekke Wigboldus, Jan Kees Looise and André Nijhof, *Understanding the Effects of Works Councils on Organizational Performance: A Theoretical Model and Results from Initial Case Studies from the Netherlands*, 19 *MANAGEMENT REVUE* (Special Issue: Industrial Democracy) 307 (2008) (emphasizing the efficiency effects of workers’ insights about enterprise processes themselves, i.e., efficiency effects do not run only via greater worker satisfaction or fulfillment).

⁸⁸ This too goes back at least to Smith. As Marx observed, “And this is labor for Smith, a curse. ‘Tranquillity’ appears as ... identical with ‘freedom’ and ‘happiness’. It seems quite far from [his] mind that the individual, ‘in his normal state of health, strength, activity, skill, facility’, *also needs a normal portion of work* ... But Smith has no inkling whatever that this overcoming of obstacles is in itself a liberating activity—and that, further, the external aims become stripped of the semblance of merely external natural urgencies, and become posited as aims which the individual himself posits – hence as self-realization ... hence real freedom...” Karl Marx, *Grundrisse*, Chapter on Capital, Section two. This is Marx criticizing Smith for failing to recognize that—under the right conditions—work can represent one of the “real freedoms” of being human, rather than being counted only in negative utilities.

Solving holdup problems

Holdup problems are a key reason for firm-based coordination from Coase to Williamson to Hart, and they are worth our attention. While labor effort may be overvalued in this stream of thought, and the benefits of democratic association may be undervalued, economic bottlenecks where one or a few actors have almost unilateral power to hold up a transaction are a real economic prospect independently of how we value these other things. Hart and Williamson (among others) both emphasize relationship-specific investments and the potential they create for one or the other party to hold up a transaction (for unreasonable terms), as an important impetus for firm-based coordination (or ownership integration of the relevant assets). While these problems are real, it is not clear that one would need *hierarchical* (as opposed to democratic) association to solve them. Moreover, various types of fair contracting rules would help to ameliorate, though perhaps not eliminate, such problems.

A key problem for Williamson was the kind of opportunism that may arise given the tendency of markets to shrink or disappear under conditions of non-homogeneity and transaction-specific investments.⁸⁹ A famous example is that of the railroad owner who invests in a spur that goes nowhere (of economic relevance) other than to a particular coal mine at the top of a particular mountain, owned by another person. Once he builds the spur, the railroad owner is effectively at the mercy of the coal mine operator, who—in the absence of competing buyers for the use of the spur—can bargain him down, perhaps below the costs incurred for the project. The railroad operator can no longer threaten not to build the spur, and moreover, he now has incurred costs that he will want to mitigate to whatever extent he can. Of course, there will usually be a contract in place before the spur is built. Williamson supposed that contracts would often be insufficient to police opportunistic behavior (at least without symmetrical investments by the counter-party, or “hostages”) while others supposed that contracts would generally be sufficient.⁹⁰

While Williamson’s solution to the holdup problem ultimately emphasized hierarchy (insofar as he expressly understands the character of firm-based coordination this way), Oliver Hart’s solution to the problem of relationship-specific investments (as for example arises in case of complementary assets—such as the railroad spur and the mine) instead emphasized centralization of ownership. Restating this slightly, we can understand the first approach as centralizing decision-making rights (over the economic

⁸⁹ See, e.g., Williamson, *Markets and Hierarchies*, 318.

⁹⁰ Russell Pittman, *Specific Investments, Contracts, and Opportunism: The Evolution of Railroad Sidetrack Agreements*, 34 *J. L. & Econ.* 565 (1991) (discussing divergence between Coase’s heirs).

activity at issue) and the second as centralizing the income streams that flow from the activity (through ownership, or residual rights in the asset)—which will in turn result in centralizing decision-making rights, by investing a particular party with bargaining power relative to other participants.⁹¹ (Hart also suggested that the person with the clearest incentive to ‘improve’ the asset, or more generally to take productive efficiency-enhancing action with respect to the process or activity as a whole, ought to be invested with ownership.)

Holdup problems seem to militate in favor of vertical integration. But the reason that ‘centralization’ is preferred in these instances is because the dispersed decision-making scenario to which it is being compared is one in which the dispersed coordination rights *correspond to the complementary assets* (i.e., we each control one of the complementary assets). If instead we disperse decision-making rights to the same numerical extent but *across* those assets, the holdup problem is also solved. What we may not want is one party (or faction) with veto power over a given (complementary) asset—but we can set up various coordination mechanisms (within or without firms) that avoid that, without centralizing decision-making rights overall. One example would be an integrated enterprise with internally democratic governance mechanisms,⁹² but various forms of market-wide coordination can play this role without requiring integration. For instance, we can imagine fair contracting and even pricing norms (enforced by law, regulation, or some kind of public-private body depending upon the context) that would constrain the sort of opportunistic holdups that vertical integration is meant to solve. Holdup problems are not limited to strictly complementary assets; they may extend to intermediate input markets generally under conditions of either demand or supply shocks. In these contexts, over-reliance upon firms as units of economic coordination may actually worsen holdup problems by encouraging firm-

⁹¹ See generally Grossman, Sanford J., and Oliver D. Hart. 1986. The costs and benefits of ownership: A theory of vertical and lateral integration, *Journal of Political Economy* 94(4): 691-719. For a nice discussion, see Kevin Bryan, “Oliver Hart and the Nature of the Firm,” *Voxeu* (blog) (November 1, 2016). Hart’s framework formalizes and sharpens certain questions about how particular ownership and contracting structures—ultimately, the allocation of decision-making rights and responsibilities, and financial rights and responsibilities—will create incentives and opportunities to do (or not do) certain things that have implications for efficiency and output. Those include: labor effort, careful supervision, risk-taking, and use of expertise, based on various actors’ incentives, capabilities, and preferences. See also Oliver Hart and John Moore, *Property Rights and the Nature of the Firm*, 98 *J. Pol. Econ.* 1119 (1990). Hart’s framework may be useful as a way of investigating specific coordination questions, once we have already delimited the available legal structures of economic coordination. But it does not seem to introduce a new normative or conceptual basis for centralizing decision-making rights in economic activity.

⁹² One could articulate a version of the holdup problem even with some very strong forms of ‘internal’ democratic governance, of course (even aside from the issue of complementary assets). But all but the most extreme consensus-based mechanisms should avoid this problem.

level hoarding (that in turn intensifies bottlenecks), while market-wide coordination might better solve these problems.⁹³

In short, holdup problems are real, but it is unlikely that vertical integration (i.e., an expansion of traditional firm-based coordination) is the only coordination mechanism that can handle them.

Implications

This paper has discussed, briefly, an influential stream of thinking about economic coordination and specifically about firm-based coordination. That stream of thought has also influenced the legal organization of the economy. It has done this by acting as both explanation and tacit justification for the de facto foundational unit of economic coordination that law favors and new policy is often built around (the traditionally organized firm, with both centralized coordination rights and flows of income); by acting as a key ‘linking theory’ for neoclassical perfect competition as a goal of law and policy⁹⁴; and in the U.S., by forming the primary intellectual groundwork for a program of legal and policy changes in competition law that have collectively tended to further concentrate both coordination rights and the flow of incomes (and a result, economic power) in markets more broadly.⁹⁵

I have discussed the preferred features of economic coordination in this approach in broad strokes. There is of course variation and debate—more than captured in this essay—but the debate is bounded by some common features. In identifying preferred forms of economic coordination, the most influential stream of thought has sought to ask how they solve for problems of productive or operational efficiency, confronting both “transaction costs” (of marketing) and “associational costs” (of intra-enterprise coordination), effectively trading these off against each other in various circumstances. Functionally, much of the variation can be described in terms of the

⁹³ These issues have highlighted lately given the effects of the COVID pandemic upon both production and demand patterns in various sectors, at various points in time. See, e.g., Nathan Tankus and Joe Weisenthal, *Notes on the Crises* (interview) (November 3, 2021), <https://www.crisisnotes.com/notes-on-the-crises-podcast-1-joe-weisenthal-on-supply-chains/>. For one approach to begin thinking about fair contracting norms more generally, see Sandeep Vaheesan, *The Morality of Monopolization Law*, 63 W. & M. L. Rev. (2022). And while pricing and contracting norms to anticipate every instance of specialized commercial dealing may not be realistic, even in the original railroad spur scenario, for example, the original contract would serve as a quite serviceable basis to work out a fair price. (If specific commodity prices impacting one party or the other have changed, that is also fairly easily taken into account.)

⁹⁴ David Grewal.

⁹⁵ The mechanisms for the latter effects (via changes in competition law) have in turn been both direct—expanding the scope of firm-based coordination through lax merger and monopolization policy—and indirect—by legitimizing inter-firm coordination that mirrors the preferred features of intra-firm coordination. See e.g., Callaci, Steinbaum, et al, “Vertical Restraints and Labor Markets in Franchised Industries” (working paper, July 19, 2022) (discussing Blair & Kaserman).

allocation of coordination rights over the economic activity itself, and the allocation of material or pecuniary benefits (or simply flows of income), from that economic activity. Broadly speaking, centralized coordination rights are presumed to solve both holdup problems and associational costs better than dispersed ones (Williamson), except perhaps in case of functional and status homogeneity among workers (Hansmann). Centralized flows of economic benefits are also presumed to conduce to productive efficiency, by incentivizing efficiency-conducting actions by those best positioned to undertake them (Hart, others).

However, this approach in the end does not really justify or explain the basic pattern of economic coordination it assumes. Often, it simply assumes hierarchy is the only alternative to “the market” (Coase, Smith), or fails to consider forms of broader and more participatory decision-making—including fair contracting rules and various forms of public pricing norms (all). Overall, the dominant accounts are also inflected with “outputism,” a bias toward more output that is baked in at the level of organizational design. This tendency becomes exaggerated where labor effort is concerned, because the neo-institutionalist stream of thinking has few conceptual resources for making sense of *too much* labor effort rather than too little—for conceiving of it as a productive inefficiency in itself, particularly on a social scale. By inscribing these tendencies into the ideas about organizational design that aim to fill in the black box at the centre of neoclassical price theory, analysis in this vein will always be hobbled in dealing with ‘the sweatshop problem’—one that casual observation tells us is endemic in our world. Bringing this in after the fact at the level of features of markets, as the consequence of monopsony or some other ‘market failure’, will not quite address a tendency that has been baked into the theorization of economic units themselves. Similarly, describing these problems simply as special, post-hoc humanitarian exceptions projects them deviations from a rational market order, hampering efforts to truly account for them.

Finally, while I have argued that neo-institutionalist approaches to economic coordination have in many cases ignored possible or alternative legal rules, instead taking status quo rules as given when solving productive efficiency problems, in other ways these accounts also ignore how actual economic actors making decisions about firm boundaries are often motivated to maximize benefits to themselves under *actual* legal rules. One example of this is control groups within firms that engage in merger & acquisition activity not in order to realize pure operational efficiencies, but in order to realize the pecuniary benefits to themselves (and to some or all shareholders) that so often flow from merger activity but do not (necessarily) reflect any particular

operational business reality at all.⁹⁶ Another example is decisions made by “lead firms” to shed divisions into sub-contracting or “independent contractor” relationships as long as they can still effectively control the actors involved through contract, while shedding the responsibilities and countervailing rights implied by the employment relationship under the New Deal settlement.⁹⁷ The phenomenon of fissuring, in fact, poses counter-examples to some of the predictions flowing from accounts assuming that “make or buy” decisions reflect optimizations of productive efficiency.⁹⁸ Fissuring—frequently discussed by labor scholars—directly implicates the explanatory domain of the theory of the firm.⁹⁹ Overall, it is better explained by the tendency of existing differentials in coordination rights and flows of income to intensify themselves—a power-perpetuation explanation rather than a technical efficiency explanation—just as the advent of firm-based coordination itself is better explained this way.

This doesn’t mean, of course, that real differences in productive efficiency do not ever explain the advent or persistence of various economic arrangements. But overall, we have undervalued the tendency of patterns of economic power to reproduce themselves through institutional design. When studying these arrangements, we only intensify that tendency by assuming that status quo arrangements tend naturally toward neutral solutions of productive efficiency problems. The argument in this paper encourages us not to make this assumption, but it also encourages us to consider values in addition to productive efficiency—fairness, democratic governance, ecological considerations, and egalitarian outcomes, *alongside* an appropriate conception of productive efficiency (cleansed of a bias toward outputism and incorporating an accounting of the efficiency costs of too much labor effort)—from the ground-up in constructing forms of economic coordination. It is legitimate to incorporate these

⁹⁶ See, e.g., J.W. Mason on merger activity as a “corporate money hose” funneling funds from productive spheres to the financial sector.

⁹⁷ David Weil, *THE FISSURED WORKPLACE* (2014); Brian Callaci, Control Without Responsibility: The Legal Creation of Franchising, 1960-1980, 22 *ENTER. & SOC’Y* 156, 178 (2020); Sanjukta Paul, Fissuring and the Firm Exemption, 82 *L. & CONTEMP.PROBS.* 65, 67-72 (2019); Marshall Steinbaum, Antitrust, the Gig Economy, and Labor Market Power, 82 *L. & CONTEMP. PROBS.* 45, 62 (2019); Callaci, et al, “Vertical Restraints and Labor Markets in Franchised Industries” (working paper, July 19, 2022)

⁹⁸ For instance, George P. Baker and Thomas N. Hubbard, *CONTRACTIBILITY AND ASSET OWNERSHIP: ON-BOARD COMPUTERS AND GOVERNANCE IN U.S. TRUCKING* (June 2004) (arguing that the advent of surveillance tech in the trucking industry would lead to ownership integration of trucks by trucking companies). Generally speaking the trucking industry in the United States remains “fissured,” with drivers designated as independent contractors; where trucking companies do own the trucks it is often because drivers in a given market generally don’t have access to capital.

⁹⁹ David Weil, *THE FISSURED WORKPLACE* (2014) (originating the concept of “fissuring,” and showing that surveillance technology spurred vertical *dis*integration across a number of sectors, with control reproduced through contract (vertical restraints) instead); see also Brishen Rogers, *The Law and Political Economy of Workplace Technological Change*, 569-70 (on “data-driven fissuring”).

Paul:

DRAFT, U. Chi. L. Rev. (2023, forthcoming)

considerations from the ground up rather than as extrinsic values to be balanced with “the market.” These values have just as much or as little to do with markets and competition as productive efficiency does.