UNIFYING COPYRIGHT: AN INSTRUMENTALIST’S RESPONSE TO SHYAMKRISHNA BALGANESH

Richard A. Epstein∗

INTRODUCTION: PUBLIC AND PRIVATE LAW

It is altogether appropriate in a symposium devoted to the revival of private law that Professor Shyamkrishna Balganesh should write a paper to explain how the law of copyright is in fact moored deeply in the private law.1 In one sense, his claim sounds a bit like a paradox, for the origin of any copyright rests solely on the statutory system that confers on the holder of a copyright the exclusive right to use, license or sell any copyrighted work.2 All the requirements that are needed to generate a copyright — that it be, for example, an original work fixed in a tangible medium — derive from statute.3 In this regard, the law of copyright differs sharply from the law that governs land and chattels, in which common law largely created the exclusive rights of use, and the limitations thereon, even if they are subject to statutory modification.

The explanation as to why copyright falls under the rubric of private law comes in two parts. First, the definition of private law does not depend on the origin of the rights in question, but only on the parties to a particular dispute. Private law involves suits between two private individuals, neither of whom receives any privilege or advantage from the state. In this regard, private law differs sharply from the two most common forms of public law. The first involves situations in which the state imposes sanctions through the criminal law or in which it seeks to exercise the powers of taxation or eminent domain, which necessarily lodge only in public bodies or private indi-

∗ Laurence A. Tisch Professor of Law, New York University School of Law; Peter and Kirsten Bedford Senior Fellow, The Hoover Institution; James Parker Hall Distinguished Service Professor Emeritus of Law and Senior Lecturer, The University of Chicago. My thanks to Samuel Eckman and Taylor Rouse, University of Chicago Law School, Class of 2013, for their valuable research assistance.


3 17 U.S.C. § 102(a) (“Copyright protection subsists, in accordance with this title, in original works of authorship fixed in any tangible medium of expression, now known or later developed, from which they can be perceived, reproduced, or otherwise communicated, either directly or with the aid of a machine or device.”).
viduals who are authorized to act on their behalf. The second part of the explanation considers situations in which the state seeks to defend itself — for example, by some notion of sovereign immunity, which it claims provides it an absolute or qualified immunity against suit for any form of government misconduct.4

This simple topology precludes creating as a structural matter any hard distinctions between the private law of land and the private law of copyright or, indeed, patents or any of the other forms of intellectual property. The same rough equivalence holds with the public law as it relates to copyright and other forms of intellectual property. These forms are subject to criminalization, taxation, condemnation, or even expropriation. In all cases, the task at hand is to articulate a general analysis to resolve two problems. The first is to explain the similarities and differences among the various fields. The second is to explain how the various forms of property can be subsumed in a single general theory that respects these similarities and differences among the various forms of property.

In dealing with these issues, Balganesh is of two minds. On occasion, he speaks about the common elements between the various forms of property; other times, he stresses their differences. But what is missing in his account is a unified theory of property that links the various strands together. In this short response, I shall have time only to point out some of the gaps and the tension in Balganesh’s general approach. In Part I, I examine whether property rights can be addressed solely in instrumental terms or whether some independent moral substrate lies beneath those admitted consequentialist virtues, and I conclude that there is no such independent basis. In Part II, I ask whether it is possible to ignore the strict correlative nature of rights and duties in understanding how various systems of property rights work and conclude that it is not. In Part III, I address the role of defeasibility in fleshing out any complete system of obligations under the private law and again conclude that they operate in the same fashion in both domains. Finally, in Part IV, I examine briefly the systematic interconnections between free speech and copyright law on the one hand and ordinary property rights and patents on the other. I conclude that both copyright and patent law are justified deviations from a pure libertarian theory of property rights.

I. PROPERTY AS AN INSTRUMENTAL VIRTUE

Balganesh is of two minds when he speaks about the instrumental virtue of copyright law. The reasons to support it are clear — namely,

---

the view that the creation of exclusive rights in various kinds of writings promotes their long-term creation. But he is uneasy about casting copyright law uniquely in such consequentialist terms. Taking a cue from his Oxford mentor, the late Professor Peter Birks, Balganesh also thinks that it is important to take into account “a uniquely legal normativity that instrumental accounts find difficult to capture.” In so doing, he clearly relies heavily on Professor Ernest Weinrib’s influential effort to develop a system of private law that rests upon the notion of “correlativity,” which Weinrib contends is founded on “a bipolar conception of the remedy as the annulment of the parties’ correlative gain and loss.”

I confess that I find this overall approach an intellectual nonstarter. The first difficulty with it is common to all efforts that seek to unify two distinct approaches into a coherent whole. That task cannot be done if each of the two notions is irreducible to the other. At that point, we are in an “ordinal universe” in which the only cases with clear results are those in which the two tests point in the same direction — at which point we can accept either and ignore the other without having to choose between them. Yet whenever the two variables point in the opposite direction, there is no principled way to give one lexical preference over the other or to reduce the two to a common denominator. Balganesh claims an effort to spell out the architecture of copyright is more than just of theoretical significance but does not explain why or offer an example of any outstanding dispute in copyright law that is clarified by resort to this two-part procedure.

A more uniform, consequentialist view of property law avoids these analytical traps. That approach develops legal rules and legal institutions — both are key — in a systematic fashion that maximizes a specific conception of human well-being. That substantive view proceeds by comparing alternative legal regimes in order to see whether at least one person is better off in one regime when no person is worse off and then moves to that position. This conscious effort to use legal rules to create Pareto improvements creates a unified theory because it compares various costs in a systematic, incremental fashion. It starts with any baseline and examines its properties. Once the defect in those arrangements is identified, it introduces a specific proposal for an overall improvement and examines whether the proposal achieves its intended result.


6 Balganesh, supra note 1, at 1664.


8 Id. at 66.
It is just this theory that props up the general social contract theory that derives from Thomas Hobbes, John Locke, David Hume, and Adam Smith. It starts with the state of nature and asks what inconveniences exist in a world without property. Chaos and tumult supply a sufficient answer. The recognition of personal rights of autonomy improves the overall situation markedly if all persons are secure against the use of force and fraud by others. That result cannot be created by voluntary exchanges that bind all persons over all times, so they are imposed through a “social contract” whose terms are justified not by any form of individualized consent, but by the overall improvements that it creates for all persons who live under that regime. Next, these individuals may acquire property by occupation or capture, which again produces an overall improvement by generating security of possession for all persons. These institutions are not self-enforcing, so the state then taxes all individuals such that the revenues they pay to the state cost less than the increase in the security of their persons and possessions. How to operate at each of these stages creates many midlevel questions — how does the first possessor demarcate his claim, for example — which give rise to additional questions. And this regime is subject to further qualifications — should the first possession rules be limited in operation if the creation of common pool problems leads to premature extinction of various species? But in each specific case we can ask questions such as whether the increased administrative costs of government are justified by the gains in output that they create; and we can use, if possible, side payments (just compensation) to deal with any imbalance that the shift in legal regime creates. The full process requires considerable conceptual elaboration and attention to institutional detail. But it does not run into the theoretical difficulties that plague any effort to posit two independent values that are not reducible to each.

II. RIGHTS AND THEIR CORRELATIVE DUTIES

Balganesh’s desire to isolate some distinctive noninstrumental value leads to a second major confusion in his attitude toward rights and duties. At some points, he takes the traditional, or Hohfeldian,9 notion that rights and duties are strictly correlative.10 But at other times, he writes as if there is some deep mistake in the common accounts of copyright law ignoring this element because “copyright law is rarely, if ever, conceptualized as a duty-imposing system.”11 He seeks to bolster

10 Balganesh, supra note 1, at 1667.
11 Id. at 1665.
this position by an appeal to the writing of Professor Jeremy Waldron, who observed that it is often productive to start with legal duties because “legal duties are hard things for people to have — since they constrain conduct and in that sense limit freedom.”12 That argument can then be bolstered by the claim that the “internal” aspect of law, as H.L.A. Hart defines it, deals with the sense of obligation as well.13

With respect, the point gets us very little. Rights are attractive because they expand freedom. There is no more reason to start with duties because they hurt than with rights because they benefit. Nor should we think that the internal aspect of law only applies to duties. It also emboldens individuals to assert their claims because they have been legitimated by the same method of norm internalization that is so critical for the acceptance of duties. In all cases, the question is not where we start, but where we conclude. The key question is whether, at a minimum, the creation of the new set of right-duty relationships creates more wealth (or utility) than the one it displaces, which means that no matter where the inquiry starts, the conscientious evaluation of the comparison is strictly needed to get matters right. Yet Balganesh does not go back to the traditional conception, but takes a different path. Although his discussion on the point is not entirely clear, he seems to think that “power conferring rules” give more play to consequentialist considerations than do “obligatory rules,”14 without explaining how both can be done simultaneously in a legal regime in which rights and duties are strictly correlative.

On balance, therefore, Balganesh is on far firmer ground when he insists on the strict “correlativity”15 of rights and duties than when he strays from that basic conception in his effort to undermine the instrumental conception of property that he questions from his opening paragraph, which weakly admits that “this instrumental focus may have the beneficial effect of limiting copyright’s unending expansion.”16 The respect for correlativity, however, does not depend on embracing any particular normative theory of property rights. It rests on nothing more than the simple proposition that this strict correlativity is an inevitable consequence of scarcity. Any ability to create rights without imposing correlative duties is the legal equivalent of a perpetual motion machine. But given scarcity, all rights claims will necessarily create conflicts, such that the social task is to find that mixture of rights and duties that minimizes those conflicts in light of the con-

14 Balganesh, *supra* note 1, at 1671–73.
15 *Id.* at 2.
16 *Id.* at 1.
sequentialist framework outlined above. The deontic approach to which Balganesh shows limited fealty never explains how the scope of copyright protection turns “on the connection between an author and his or her creative expression.”17 Balganesh then touts this Kantian argument: “[c]opying is morally wrong not because it operates as an infraction of an ownership or property interest, but because it directly interferes with an individual’s ability to perform a speech act (that is, expression) and communicate with the public as speaker.”18 But the point is surely incorrect. Copyright does not protect the right of the speaker to perform any speech act. I can read my novel aloud happily long after it loses copyright protection. The issue is of course exclusivity, which cannot simply be achieved by the speaker, but by the law, which can stop others from speaking without the copyright holder’s permission. Balganesh at one point is puzzled about the difference between a positive right that “actively enables” speech and the use of copyright that “disables” the right of speech in others.19 But there is no mystery here. No one needs help from the law to talk. It is needed to suppress the speech of others who are quite able to speak for themselves. The fatal weakness of the supposed deontic approach is that it never addresses the question of why an author is able to prevent others from copying that work.

In order to avoid the solipsism of the deontic approach, it is necessary to think of rights in two distinct ways. The first sets out the rights that all persons have vis-à-vis the rest of the world over their person and their property. The second asks what should be done when one person invades the rights of another. The first is a question of property. The second is a question of torts.

On the first point, the standard rules of property require each person to abstain from the use of force and fraud against the person or property of strangers. That system only works because its obligations are simple enough for ordinary people everywhere to understand without some special system of notice: keep your hands to yourself, as it were. This system is also scalable, such that its basic commitments remain constant regardless of the number or identity of people it governs. Finally, that system is largely invariant in relation to social levels of wealth, lest the content of the rights constantly shift with the ebb and flow of economic fortune.20 For these purposes, moreover, it is not important that the people governed by this legal regime understand its instrumental foundations, so long as they know how to conform their

17 Id. at 1677.
18 Id. at 1679; see also Immanuel Kant, On the Wrongfulness of Unauthorized Publication of Books, in PRACTICAL PHILOSOPHY 23, 30, 33 (Mary J. Gregor ed., 1996).
19 Balganesh, supra note 1, at 1670.
20 For further discussion, see RICHARD A. EPSTEIN, DESIGN FOR LIBERTY 74–76 (2011).
conduct to its basic rules. For that reason it is useful in everyday life for the “rights based,” or deontological approach, to exert a strong hold. Insisting on the moral basis for these commands increases the rate of compliance by ordinary people, which in turn allows the social system to flourish.

The second question goes to what kinds of conduct count as invasions of that right. Here, the instrumental and the deontic approaches surely coalesce so that virtually everyone, regardless of philosophical preference, agrees that this duty of uniform noninterference in the sense specified applies to deliberate use or threat of force. The great fruits of this rule are found in the harms that never occur precisely because the norm is so strong. But differences arise when asking whether to extend that protection to all accidental harms or only those caused by negligence. Most debates over this choice resort to the implicit logic of the social contract by insisting first that property rights are not “absolute” and then asking which rights against invasion all property owners surrender — shades of the Lockean theory — to promote social progress. In some instances, lawmakers have used this argument to reject a system of strict liability and thus limit liability for accidental harms to instances of negligence. But although this approach adopts the proper methodology, it need not lead to a negligence rule for accidental harms. What remains is the empirical issue of whether a person receives “more than a compensation” when surrendering the greater protection afforded under a strict liability system for the lesser protection afforded under negligence. That conclusion is, to say the least, contestable. Since rights and duties are strictly correlative, why assume that the greater freedom of action is worth the loss in security? The willingness to make extensive investments in real estate depends critically on protection from damages resulting from accidental harm, such that the development of factories could benefit from the use of strict liability rules.

21 See, e.g., Losee v. Buchanan, 51 N.Y. 476, 484–485 (1873) (“By becoming a member of civilized society, I am compelled to give up many of my natural rights, but I receive more than a compensation from the surrender by every other man of the same rights, and the security, advantage and protection which the laws give me. So, too, the general rules that I may have the exclusive and undisturbed use and possession of my real estate, and that I must so use my real estate as not to injure my neighbor, are much modified by the exigencies of the social state. We must have factories, machinery, dams, canals and railroads. They are demanded by the manifold wants of mankind, and lay at the basis of all our civilization. If I have any of these upon my lands, and they are not a nuisance and are not so managed as to become such, I am not responsible for any damage they accidentally and unavoidably do my neighbor. He receives his compensation for such damage by the general good, in which he shares, and the right which he has to place the same things upon his lands.”).

For similar sentiments, see Brown v. Collins, 53 N.H. 442, 445–50 (1873).
In general, I come down in favor of the strict liability result, but for these purposes, it is key to avoid the tendency to over-intellectualize the relationship between property law (that sets out the general relationships) and tort law (that deals with particular invasions). Balganesh never answers that question directly. Nor, ironically, does he offer a clear answer to the question of why the rights of redress are lodged in the owner, which is the standard conclusion of theories of corrective justice. But it is possible to give a sensible instrumental account of the moral intuition. Surely the identity of the owner matters; otherwise, it is impossible to know who may sue for property damage or destruction. The utilitarian foundation for the standard theories of corrective justice is that, as a first approximation, the owner is both in the best position to maintain the suit and most in need of recovery. After all, who will be willing to invest in real estate development if he receives no compensation for its destruction? The analysis, moreover, is exactly the same for copyright and patents, which again couple generalized duties of noninterference with specific rights of action vested in the owner whose rights are infringed. To be sure, copyrights and patents do not have spatial locations, which is why, in any given jurisdiction, they have to be protected by being listed in a registry of some type accessible to all. But so long as the right to exclude is good against the rest of the world, does its protection cover only deliberate infringement, or also those made accidentally or by negligence?

On this question, copyright and patent follow different paths, and for good reason. The course of scientific progress is such that it would be virtually impossible to prove copying for new inventions that depend largely on the same set of scientific laws and basic technologies. So the liability remains strict lest the protection prove worthless. Copyright presents a different set of issues, for the forms of words that are truly distinctive are not defined by the state of technology. “Shall I compare thee to a summer’s day” is not likely to be written by anyone, such that in general it is easy enough to establish that extensive passages have been lifted. And Balganesh is surely right that this presumption is in fact reversed for common phrases of words or music in general circulation. The liability rule for copyright thus requires, as the name suggests, copying. In both cases, however, the principle of strict correlativity is observed. If liability depends on copying, the independent creation of a writing or a song is a protected liberty, not a legal wrong.

23 Balganesh, supra note 1, at 1683–84.
III. DEFEASIBILITY

Thus far I have stressed the two initial stages of any legal inquiry: the delineation of the basic rules of property good against the rest of the world and the articulation of the prima facie case for their invasion. But any prima facie case admits to defenses that can “defeat” liability in a particular case. It is this idea of defeasibility, on which Balganesh rightly relies, that links property law to tort. In dealing with this notion, however, it is critical not to repeat the mistake that H.L.A. Hart made when he claimed that defeasibility was the key to understanding the meaning of various key terms, such as contract. Defeasibility is not definitional: a contract is an agreement between two or more people to abstain or perform specified actions. Instead, defeasibility serves a different office by expressing the relationship between action and responsibility. Thus, if A does not keep his promise, is he justified if B, who was to perform first, did not keep his? Similarly, the proposition that A entered B’s land states a prima facie case of trespass, but that inference can be defeated in a number of ways. Some are consensual, as by showing that A was a guest or customer of B. Others are nonconsensual, as by showing that the entry was under conditions of necessity or pursuant to higher legal authority. It is thus instructive that the English language has these pairs: entry/trespass, killing/murder, and nonperformance/breach. The first term in each pair states the prima facie case of the latter. But a complex array of defenses, which are themselves subject to further qualifications, prevents any easy equation of the two.

The exact details of these defenses are not easy to summarize in a few short sentences. But the entire project is in line with the larger instrumentalist effort to maximize the appropriate (Paretian) standard of social welfare by a system of successive approximations to some ideal distribution of rights. This strategy is very much in play in the law of copyright, which has its own extensive set of consensual and nonconsensual defenses, where the first is contained in the law of copyright licenses and the latter in the amalgam of “fair use” defenses to which Balganesh refers. The precise contours of these defenses are beyond the scope of this short article, but the basic pattern of their implementation is necessarily subject to the law of diminishing returns as it ap-

---

24 Id. at 1684–86.
26 For a systematic implementation of this project, see Richard A. Epstein, Defenses and Subsequent Pleas in a System of Strict Liability, 3 J. LEGAL STUD. 165 (1974).
plies to the design of legal systems. The initial cut into the system, which creates the basic system of entitlements, promises large rates of return over the previous state of affairs and does so at relatively low cost. The protection of individual autonomy and private property are key conditions for civic peace. Each new refinement in the system, however, produces smaller benefits at increasing costs, as is surely the case with the complex set of public and private necessity doctrines that authorize a narrow class of nonconsensual entries into land. The decision over whether to supply compensation for entries that cause damages offers still smaller gains but does so at increased cost. Sooner or later, therefore, the costs of further refinement exceed the social gains, so that it becomes futile to capture each novel variation with its own special rule. At this point, for want of a better alternative, legal systems often resort to basic conceptions of reasonableness, which operate as an invitation for triers of fact to continue the process of refinement on an ad hoc basis in individual cases. But that system works best as a backstop to the more specific process of introducing new arguments step by step in a general pleading system.

IV. COPYRIGHT AND FREE SPEECH

It is just this process of successive approximations that helps explain how intellectual property rights become incorporated into a more general legal system. To set the stage, note that our basic system of property rights starts with the protection of personal autonomy and property rights in tangible things. If that law were treated as complete, it would leave no space for either patent or copyright protection. A patent gives its owner the exclusive right, good against the world, to make or vend a particular thing. That right necessarily limits the ability of all other individuals to fashion their tangible possessions into devices of their own choosing even when their conduct neither poses a threat of the use of force against a third person nor counts as a breach of contract with any other person. Similarly, the ordinary person has rights to freedom of speech so long as he does not engage in force (threats) or fraud (deceit and defamation). Reading or reproducing the words of other individuals does not violate either of these fundamental prohibitions and thus counts as a protected version of free speech.

It is clear, therefore, that any system of intellectual property is in marked tension with the apparently complete system of liberty and property generated under orthodox libertarian theories pertaining to force and fraud, which shape so much of the law. One possibility is to

---

28 For further discussion, see Epstein, supra note 20, at 31–42.
conclude that both systems of intellectual property are illegitimate, but that option seems foolish in light of the enormous across-the-board gains that a sound system of copyright and patents can generate. The basic social contract thus uses the same strategy that first generated rights in liberty and property to add yet one more key refinement into the overall system. Recall that the rules of individual autonomy and first possession necessarily limit the rights of others to kill or maim or to wander freely over the land. Due to the large number of individuals, there is no feasible path by which any one person can negotiate an elaborate set of private agreements that would move the world to a new equilibrium that protects liberty and property against all invasions from any quarter. So the social contract amounts to a gigantic forced exchange whereby everyone surrenders his or her right to invade the person and property of others in exchange for similar security. The change works because the new constellation of rights yields higher values to all concerned than the older one — an empirical truth, to be sure, but a truth nonetheless.

Instrumentalism is so crucial for understanding intellectual property law because it offers the only explanation for how the acceptance of both patents and copyrights generates widely diffused social gains that justify the nontrivial costs of their implementation. That conclusion necessarily depends first on the ability of copyrights and patents to create incentives for production and second on their ability to facilitate the commercialization of these writings and inventions once created. So understood, copyright and patent law function as important Pareto improvements over a system of liberty and (tangible) property that makes no allowance for their creation or enforcement.

In general these basic rights work best when they follow the pattern of other forms of property with respect to both protection against infringement and licensing for use. But they do differ in one key respect. The gains from privatizing writings and inventions should be, and routinely are, subject to a time limitation that is not found in the law of real property. Here is the functional explanation. Suppose that the first possession rule gave the initial possessor title for only twenty years: what would happen at the expiration? At that time, the property either goes back into the common pool or it becomes subject to government ownership. Either way, the prior owner of the property will have few incentives to make improvements that will outlast his ownership of the property. Yet nothing is gained by putting this title up for grabs a second time. Hence, it is better to create an ownership system with infinite terms out of which shorter terms can be created by voluntary leasehold or other conveyance.

Creating property rights in writings and inventions of unlimited duration has the advantage of longer time horizons for coherent property development, but only at a far higher cost. Given that patents and copyrights encode information, the original copyright or patent
holder can continue to use that information, just as others use it, after the copyright falls into the public domain upon the expiration of its term. The great advantage of the open-access regime, which cannot be obtained with land, is that it increases the level of dissemination of the patented or copyrighted work in the later period by allowing many users access to the same basic writing or invention without charging fees. Those fees otherwise would count as an additional cost, which would necessarily block usage by those who value the invention or writing less than the licensing fee. Those losses are virtually impossible with land, which under most circumstances cannot be used simultaneously and productively by its owner and a legion of outsiders.

At this point, the fundamental question is just how long these rights should extend. Everyone agrees that a longer period is proper for copyright than for patents, a fact that constrains but does not define the proper term for either class of property rights. Getting the right answer is soluble, at least in principle, through a search that seeks to maximize net social value from the creation of new inventions or writings. For our purposes, however, the key point is not to identify these optimal terms, which should not reach the outrageous lengths found in the Copyright Term Extension Act, but to show that this inquiry is part of a continuous enterprise that allows for the orderly expansion of property law.

CONCLUSION

In this short comment, I have tried to defend what I regard as a traditional instrumentalist account of copyright law against Balganesh’s revisionist position. The basis for that defense is easy to convey. Generally speaking, it is best to rest on a simple and coherent intellectual base in order to develop the law of copyright and to show its linkage to the rest of property law. Although I have given the general outlines of this position, I have not developed its details. But I think that I have said enough to warn people away from the ambitious and ingenious effort to couple the standard instrumental theories of copyright with an alternative deontological, if mystifying, account. The disagreement with Balganesh is not with regard to particular rules but with regard to overall approach. I do not see any case in which his elaborate dualism clarifies the operation of copyright law, as his resort to elusive notions of “legal normativity” poses the risk of leading copyright doctrine away from its most desirable path. It is not that Balganesh’s flawed approach has generated results that are counterin-

---

tuitive. It is just that the entire edifice can rest on firmer foundation if the law hews to a consistent instrumentalist interpretation. Here, as in so many other areas, simple is best.