A Model of What?

Exploring Boldren's and Levine's "A Model of Discovery"

Working Paper

Please do not quote without permission.

Stephen E. Margolis

Department of Economics
North Carolina State University

June 19, 2013

In "A Model of Discovery"[1], Boldrin and Levine continue their broad attack on intellectual property. They present a model that purportedly shows that competition in the creation of "knowledge" yields more of it than monopoly would. What they mean by competition amounts to absence of any prohibitions on copying—the absence of any intellectual property law. In contrast, simple monopoly stands in for intellectual property institutions. In short, their model is a rhetorical effort that would persuade us that patent reduces invention. A premise in their argument, one that is not particularly emphasized, is that the assumptions on which their model is built are conventional and therefore quite reasonable, while the more standard assumptions employed in economic models of intellectual property are not. The gist of what follows is that an essential feature of their model—than marginal cost is everywhere increasing—is indeed conventional in much of economics, where it can be applied "without loss of generality," but it is decisively wrong for invention and creative work.

MONOPOLY, COMPETITION, and DISCOVERY

Boldren and Levine use the term "discovery" in their title and throughout their paper to denote the intended product of intellectual efforts. In so doing, they avoid the terms invention or creative work, as would correspond to patent or copyright, respectively. This choice is apt, as the type of knowledge creation they discuss seems does not fit either of these legal institutions particularly well, although it seems to come closer to copyright than patent. In their introduction, they do present their argument as critique of patent system, but subsequently they mostly steer clear of the particulars of any actual intellectual property institutions. Interestingly though, their model does use "blueprints" as the unit of output. Blueprints, of course, as architecture, are protected by copyright.

[1] Michelle Boldrin and David K. Levine, *American Economic Review*, May 2009, p. 337

The particular meanings in their paper of competition, knowledge, and even monopoly are idiosyncratic, though perhaps not as strained as their characterization of the production of knowledge. The critical feature of their model is the assumption that the marginal product of effort in the production of knowledge is everywhere diminishing. Many economic models do assume diminishing returns, but in most instances it is an inconsequential simplification. It does not determine the results. For Boldren's and Levine's model, however, the assumption is critical. It is, however, inappropriate for the economic questions that they seek to address. In what follows, after providing some context, I discuss the problem with this key assumption as a basis for considering invention. Following that, I discuss the various ways in which their model conflicts with the options available to inventors in the absence of patents, and the general incongruity of their model as a model of the institutions of intellectual property.

One should probably not make too much of any title, however Boldrin's and Levine's title is a clue to a key shortcoming of their paper in so far as it might be taken as a model with relevance to patent. Legal scholarship on patent often distinguishes between invention and discovery. Discovery typically applies to human findings regarding states of nature that were always there, while invention is figuring out something new. America was discovered, the light bulb was invented. Invention and discovery are related, of course. Discoveries can lead to invention, so Fleming's discovery that bread mold attacks certain bacteria led to the invention of penicillin. Certain discoveries of the properties of electricity led to many inventions of electrical devices. Interestingly, the word invention does not appear anywhere in "A Model of Discovery." I will observe below that their characterization of the technology of knowledge creation is ill suited to invention *or* discovery, but it is particularly inappropriate for patentable invention.

INCREASING RETURNS

Boldrin and Levine present the conventional model of intellectual property as one in which some fixed cost is incurred to develop a new product, after which individual copies are made at low cost. The availability of cheap copying means that without protection, those who develop new products—new knowledge to be faithful to Boldrin and Levine—will have higher costs than those who merely copy, so that without restrictions on copying, those who develop new products will be unable to recover their costs. They characterize this sort of production as imbedding increasing returns to scale. All of this is unexceptional. They distinguish this standard representation from their own model, in which decreasing returns are central.

However, the decreasing returns that Boldrin and Levine incorporate actually occur entirely in knowledge creation. Their copying technology is linear, conventional, and, depending on the cost of copying, quite possibly a source of increasing returns in the production of copies, the same sense that increasing returns are present in the standard model.[2]

This brings us to the specific form of decreasing returns that Boldrin and Levine do incorporate in their model; decreasing returns in the creation of new knowledge. In their model, they model the "discovery" of a particular good, not the progress of general knowledge. The discovery process is presented as a perturbed Cobb-Douglas function. The perturbation is incorporated so that capital (blueprints) can be initially zero while allowing initial output to be positive. Of this they write, "[W]e assume that the very first bits of knowledge—incomplete sketches, intuitions, and so forth—are extremely valuable in the production of

[2] For B&L, it is the requirement in the standard model of some fixed cost that must be incurred before any production can take place that embeds decreasing returns to scale. However, the standard model is generally non-specific about the technology of invention itself. I will argue below that initially-increasing returns is a reasonable characterization. One problem with a discussion of returns to scale relates to the units involved. In the accumulation of knowledge ordinal, rather than cardinal, valuation probably makes more sense.

additional original knowledge." Initial efforts are enormously productive; everything after that is down hill. In such a world, we would each do a little bit of everything. We should each do a bit of "discovery." Or if not that those who are capable of discovery should apply themselves to many distinct "initial efforts."

Boldrin and Levine support their assumption by observing that decreasing returns is a conventional assumption in much of economics. However, in the conventional setting, economists are concerned with the determination of the level of output or the ultimate limits of the size of the firm. In short, we are usually asking how much to produce, not whether to produce. So in the usual context, the assumption of diminishing returns is an artful simplification that is accepted with some loss of generality, but without loss of meaning. One of the useful things that economics offers is a clear characterization of short-run supply, which we get to by way of the law of diminishing marginal product. Nevertheless, brushing past the likelihood that production is subject to increasing returns at very low levels of output is not innocuous. Neither is it intuitively appealing or empirically supported.

While special cases are of interest to economists, there nevertheless remains some consensus about what we think are the important ideas and useful tools. This consensus is perhaps best evidenced by what we keep in our economics principles courses. And there, with few exceptions, the textbooks present production technologies that reflect increasing returns at very low levels of output, transitioning eventually to diminishing returns. We don't really need the increasing returns portion for what comes later, but we do this for a good reason. If we didn't, the kids wouldn't believe us. While complete implausibility is not generally regarded as a fatal weakness in economics, it might handicap us in finding new recruits. Our novice students may not come to us with a lot of good intuition about economics, but they know from experience that very small scale production is inefficient. A one-off umbrella would be expensive and probably ineffective. So is one-off almost everything else. Bachelors are known to dine out

more often than families of four. People may root for the very-little-guy, but they usually buy from sellers who can at least support a storefront.

In economics classes, we work to get our novices to understand the factors that limit the sizes of firms. To get there, however, we must acknowledge what is going on at left hand side of the quantity axis. We teach U shaped marginal and average cost curves. There are exceptions, of course, but they are covered as such.

There are compelling explanations for initially-increasing returns. A certain amount of resources are used just to start a process—whether we are initiating production for the day or developing an entire new business. These are initial costs that may yield no or little output. Noting Alchian's distinction, these effects are present whether we are concerned with the rate or the volume of output.

There are other places in economics that give considerable attention to the downward sloping portion of the long-run-average-cost curve. Industrial organization makes frequent reference to the concept of minimum efficient scale and there is a body of empirical work aimed at estimating MES by industry. Contemporary international trade also relies on the influence of a range of outputs in which there are economies to scale.

Economies to scale have long been recognized as central to the economics of urban areas. In one of the seminal papers in the field, Edwin Mills[3] argues that increasing returns to production and consumption, over at least some range of activity levels, are necessary to explain concentrated settlement. Spatial concentration of people makes some locations expensive even when land itself isn't particularly scarce. People could avoid costs by remaining spread out. But we do concentrate spatially, bidding up the price of land against one another, and

[3] Edwin S. Mills. "An Aggregative Model of Resource Allocation in a Metropolitan Area." *American Economic Review*. 57 (2) May 1967. 197

bidding up the prices at the centers of settlement most of all. But absent all scale economies, we could avoid paying elevated land rents by spreading ourselves out.

Absent some range of scale economies, Mills observes, each household would be self sufficient, producing everything it needs for itself at miniscule scale. With the stronger of assumption that decreasing returns to scale that begin at zero output, the case for autarky is still stronger. We are jacks of all trades, yet masters of them all.

It sounds ridiculous and it is, and that is precisely Mills's point. Initially-increasing returns to scale are central to many features of social organization. Without them we would each run a very small farm, a one-room slaughterhouse, an optical shop, a winery, a tannery, a semiconductor plant, a steel mill, and so on. Absent the incongruous references to newer technologies, it all sounds like Little House on the Prairie. In the American West, homesteaders did suffer the diseconomies of very small scale in order to claim property rights in the land, but theirs was a life of considerable hardship.

This is not to say that it is impossible for some production activity to exhibit decreasing returns at all levels of activity, including zero, but it does seem unintuitive. Nevertheless, Boldrin and Levine's assertion of everywhere-decreasing returns is superficially plausible because it is familiar. It is familiar because, as noted above, it is convenient and innocuous is the usual context, in which we are concerned with the other end of the cost curve. Specifically, the U-shaped cost curves that we teach in intermediate microeconomics, become curvy backwards Ls because we are concerned with the outputs of individual producers and their aggregation into supply functions.

All that said, one may legitimately raise this question: Is the assumption of initially-diminishing returns inappropriate for a theory of invention or creation? It is essential to Boldrin's and Levine's model, because a would-be inventor—make

that “discoverer”—does not need to suffer some initial cost to have something worthwhile; valuable things—discoveries—show up at the outset. In their model, per unit of input, the most important things occur at the initiation of the process of discovery, at the very beginning of exploration.

B&L do proffer some motivating examples to support their point of departure. “The key intuition is that even with diminishing returns and perfect divisibility, the first few shards of new knowledge—the unfinished notes, the dead ends that have been encountered, the computer program with many bugs—have enormous value in the process of knowledge creation.

But do they? Many of us will recognize the experience that a successful research initiative might trace back to a moment when a particular problem was solved, or a particular insight was gained, or a particular experiment was conducted. Stories of invention often focus on these key moments. But the problem with these stories for Boldrin’s and Levine’s controlling assumption is that these important little moments typically don’t come first. Before they occur, there are many errant trials. Of course, one could assert, as Boldrin and Levine do, that these “encountered dead ends” are extremely valuable. But are they all? If many dead ends are required before we develop key insights, if we have to eliminate many possible solutions before we settle on the one that works, that sure starts to look like initially increasing returns; a certain level of accumulated activity is necessary before answers (outputs) begin to appear.

And what of the efforts that are initiated and abandoned? There are the shards of new knowledge, accumulated dead ends, and buggy software that are never brought to any useful end. This weighs against the Boldrin’s and Levine’s claim that first few shards of knowledge being extremely valuable, a least in aggregate. (Boldrin and Levine provide only a social aggregate production function, so abandonment isn’t really a possibility in their model, but of course it does happen.

Further, the assumption of everywhere-decreasing returns is in conflict with the most ordinary experience. Whether at the individual, institutional or national level, initiating creative or inventive work in new area almost always requires a substantial investment just to get to the frontier. Most scholars have had the jolting experience of beginning a project by finding out that a great deal of good work has already been done on the great new idea we had just "discovered." Companies and research institutions find that they must make a substantial investment to enter a new area before any new results can be produced. The world offers many examples. In spite of the existence of a well developed global automobile industry in the 1960s, the early Japanese cars were terrible. The Japanese were not breaking new ground—say, discovering—until at least a decade into that industry's development. The Korean auto industry has gone through something similar.

In short, B&L challenge the common-sense understanding of invention or creation[4] in which creative works or inventions require a significant investment, while subsequent production of copies are cheap. To do so, they conjure up a model in which some invention comes very easily, and while copies are cheap, they are not so cheap as to choke off "discovery" too soon. To pull this off they posit a technology of knowledge creation that cannot survive serious examination.

AGGREGATION, INDIVIDUALS, and INSTITUTIONS

It is important for Boldrin's and Levine's model that units of "knowledge" are perfectly divisible and useable in their infinitely divided form. "Knowledge is encapsulated in perfectly divisible blueprints."[5] In their model, all these "blueprints" aggregate uniformly. Discovery occurs in a continuous and orderly fashion as these blueprint "atoms" accumulate. One might assume that the

[4] Again, they don't really associate their model with inventing or creating, favoring "discovery" instead.
[5] P. 337, second column, last paragraph.

blueprints are each different, reflecting as they do the accumulation of knowledge, but that doesn't square with their model. To support their assumption, they note how the process of writing and revising leads to a polished product, then offer this non-sequitur: "So it is far from clear that two copies of an incomplete discovery—the notes containing the intuition of the first few experiments, and so forth—would be worth less to consumers than one copy of the complete one."[6] This seems completely contrary to how we write, how we referee, and how we grade our students. What grade do we give a student who has written a bad term paper (or let's just say an "incomplete" one) and then hits the "print" button twice, staples it all together and turns it in? The peculiar slight of hand, "it is far from clear," is not just a passing comment. It appears in the justification of their model and is essential to both the assumption of ever-diminishing returns and the aggregation imposed in their formal model. The alternative understanding, that there is some indivisibility and that a well developed idea might be worth more than two copies of a poorly developed idea, is dismissed as "an extreme view."[7]

My burden here is to show that Boldrin's and Levine's characterization of the nature of discovery is peculiar and not likely to fit much of what we think of as creative or inventive activity. Let me shift some of that burden. Here are Boldrin and Levine in a 2008 paper: "The standard case of perfect competition assumes perfect divisibility—that the initial unit may be produced in arbitrarily small quantities. In the theory of innovation, we have dropped that assumption—recognizing that two halves of a book are a poor substitute for the whole."[8]

In their model, the state of knowledge is advanced using the existing knowledge (i.e. the aggregation of those perfectly divisible blueprints) and labor to create more blueprints. They provide a utility function and a budget constraint and the equations of motion for the state of knowledge. From there, they characterize the

[6] P. 338, first column, first new paragraph.
[7] P. 338 Same place.
[8] Boldrin and Levine, "Perfectly Competitive Innovation," Journal *of Monetary Economics*, 55 (2008) 435-53. p. 443

socially optimal dynamic paths for knowledge creation, consumption, and copying. This social optimum appears under the heading "Competition."

Actual competition, however, is nowhere to be found. Boldrin and Levine don't say anything to support this characterization of "competition." There are no firms. There are no individuals at all, just units of labor, perfectly homogeneous and divisible. Treating a calculated social optimum as identical to the competitive outcome may be appropriate in some theoretical contexts, where we know that competition brings about an optimum, but there is no reason to assume that it is obtained here, given the inherent appropriation problems affecting knowledge.

Boldrin and Levine don't quite tell us what they mean by competition, other than it yields the social optimum. In so far as there is anything like competition, it seems to involve unconstrained copying of the blueprints. Production of new knowledge, however, seems to involve a single producer, since the social aggregates enter into a single decreasing-returns production function[9]. Anything like actual competition is nowhere to be found.

In what B&L's call competition, there is no copyright or patent protection to prevent copying, once any consumers have any of the perfectly divisible blueprints. So at the limit, copies are available at the cost of making a copy. This may be what B&L regard as the outcome in the absence of institutions like patent and copyright, but it leaves something out.[10] Even without patent or copyright, people can keep secrets.

This brings us to a contrast between much legal scholarship on patent and the perspective of most economists. Legal scholars are apt to present patent as a social contract in which inventors are given a period of exclusivity in exchange

[9] An alternative interpretation is that the model is of aggregate activity of many actors. For that, however, the entire stock of knowledge would have to be available to all laborers, as in a knowledge commons, but that seems inconsistent with Boldrin's and Levine's presentation.

[10] In other writings, Boldrin and Levine do recognize the mechanisms that creators might use to capture some value from their creations. (cite)

for disclosing their ideas. Economists seem to find this justification to be beside the point, and instead understand patent an incentive system for invention or, almost equivalently, a means by which inventors might hope to recover their costs of invention. Posner and Landes, however, emphasize the importance of the disclosure requirement[11], which they argue addresses problems that would arise as a consequence of efforts of inventors to keep their inventions secret in the absence of protection.

In "A Model of Discovery, however, B&L take no notice of the problem that in the absence of patents, "discoverers" (inventors?) could still keep secrets. Inventors could provide services directly rather than sell products (copies?) that incorporate their inventions. They could take extensive steps to frustrate reverse engineering. They could deliver sealed "black boxes" only to parties that commit contractually not to tamper. Failing those measures, inventors could direct their efforts away from inventions that cannot be profitably exploited without being disclosed. Similarly, absent copyright protection, producers of creative work could inhibit copying by providing only performances—readings, concerts, plays—rather than selling their creative works in tangible form. All of these impose costs on those who would produce or consume inventions and creative works. So it isn't clear that Boldrin's and Levine's "competition," in which copying becomes freely available at some point, is an available outcome. Their solution assumes away one of the central problems that the institutions of intellectual property address.

PATENTABILITY AGAIN: Every Man an Edison?

Although B&L never use the terms invent, invention or inventor, they do specifically mention patents in their introduction and they conclude that under the technology that they have conjured up, "introducing a patent is damaging to welfare…" There is no question that the paper is an attack on patent.

[11] Landes, William M. and Richard Posner, The Economic Structure of Intellectual Property Law. P. 294.

But suppose we take seriously their assumptions about the technology of discovery. Even though their technological assumptions do seem peculiar in the context of actual invention, they might well apply to some forms of knowledge accumulation. The problem is that where they do apply, it is unlikely that the fruit of such accumulations would be patentable, for the process that B&L assume seems to be at the heart of obviousness.

Under B&L's assumptions, the first learning comes easily; infinitely easily at the exact start, by these authors' own account. Where that is the case, efficient arrangements have us back in Mills's hypothetical world where everyone would do a little of everything. In that world, we should all devote at least a bit of our time to discovery. We shouldn't mind, it takes almost no time, and we would need little or no reward for our efforts. Infinity has a big upside.

Pushing this only a little further, we should all be discovering on all conceivable fronts, or at least on all fronts on which we might derive direct individual benefits. But let's give the case some benefit of the doubt and suppose we only discover in the areas in which we regularly practice, the areas in which we have developed some skills. It is not unreasonable to think that in this context, little discoveries occur all the time: a way to hold a brush, start a screw, multiply by eleven, fold a pair of socks, or mix colors to get green. These things come easily and they are discovered all the time, the same ones, over and over again. We do feel clever when it happens and possibly inclined to share the news with friends. These "discoveries" occur as a byproduct of other activities, the result of little experiments that occur at miniscule cost. Yet they accumulate into the knowledge of a craft. But what we accumulate through all these easy little steps is not invention, it is technique—the ordinary skills of our trades. These things may not be obvious to everyone at every moment, but they do occur to many people on many occasions. They are shared because they came easy and they are understood not to be scarce. They are obvious, even if not patently so.

What they are not, is patentable. To be a patentable, a new idea—an invention—must not be obvious. The conceptual leap required to meet the non-obviousness requirement for a patent would seem to be ruled out by Boldrin's and Levine's construction of discovery. In both their discussion and their formal model they characterize the technology of "discovery" as a process by which knowledge, mixed with labor, accumulates continuously with ever diminishing returns. In fact, they explicitly observe that their characterization of the technology of discovery precludes the "eureka moment" that is sometimes invoked to characterize the required non-obvious advance that is required for profitability. If anything, B&L's model of knowledge creation seems to have more in common with the kind of creation that is protected by copyright, and even some of the details of their model of their model, as I will discuss below, echo the world of copyright.

Patent critics may respond that the patent office has granted patents on some pretty obvious things. While I would note the reasonable retort that some things are obvious only after they have been demonstrated, it is true that obviousness is a difficult matter in practice and the patent office sometimes errs. Indeed, some patents are invalidated in litigation because they are shown to be obvious. There is also a line of criticism that the patent office has set the non-obvious hurdle to low, and some recent court decisions have begun to address that concern[12]. But these are particular flaws in a particular patent system and merely indicate that as an institution, the patent system is likely to be imperfect, not that a patent system is inherently without value.

ERRANT EMPIRICS

B&L claim that their paper is supported by empirical evidence that patents have little impact on innovation. In the opening paragraphs, they note "an emerging

[12] See *KSR Int'l Co. v. Teleflex, Inc.*, 550 U.S. 398 (2007)

body of evidence." (p. 337) It appears to be emerging quite rapidly; by the end of their short paper it has grown to "overwhelming." (p. 342)

Contrary to their claims, the evidence they present is rather scant. They cite[13] a paper by Joel Mokyr as supporting the claim that patents "do not have much impact on innovation." This characterization is a bit of a stretch. For the most part, in the cited paper Mokyr is arguing that the influence of patent on the industrial revolution has been overstated, particularly by Douglas North, who, according to Mokyr, regarded patents as "decisive." Also, Mokyr is considering the effects of the patent system from about 1700 to 1850. A part of his brief is that the patents were very expensive, difficult to obtain, and often ineffective. So if anything, Mokyr is commenting on the effectiveness, or lack of it, of a weak patent system in a technological era that is very different from our own, and his conclusion is essentially that the issue of patent effectiveness during the industrial revolution is unsettled.

Boldrin and Levine also cite their own 2008 book.[14] A central example of that book, the steam engine, has been criticized effectively by George Selgin and John Turner.[15] If the evidence against patents is "overwhelming" it would seem unnecessary to lean on such tenuous support.

CONCLUSION

Many, economists, perhaps most of them, have rejected parts of Milton Friedman's methodological teaching. In particular, they no longer accept an

[13] p. 337
[14] Michele Boldrin and David Levine. 2008. *Against Intellectual Monopoly*. Cambridge: Cambridge University Press.
[15] See George Selgin and John Turner, "James Watt as Intellectual Monopolist: Comment on Bouldrin and Levine," International Economic Review, 47 (4), November 2006. Also see by Selgin and Turner, "Strong Steam, Weak Patents, or, the Myth of Watt's Innovation-Blocking Monopoly, Exploded" 2009 (Unpublished) and "Watt, Again? Boldrin and Levine Still Exaggerate the Adverse Effect of Patents on the Progress of Steam Power." 2009 (Unpublished)

requirement to derive or test the empirical implications of a model. We have absorbed the post-modern skepticism about the decisiveness of any test and have learned the lessons from Kuhn's writings about what scientists actually do. Since tests cannot be decisive, it can not be essential to derive empirical implications test them.

But without saying so, the main stream of economics has retained one part of Friedman's method. For Friedman, the realism of the assumptions of a model was beside the point. Current practice follows seems to follow that disposition: The assumptions of a model are not subject to much scrutiny. We build models. It's what we do.

There is a problem, however, with taking full license regarding our modeling assumptions if we have also freed ourselves from any empiricism. By rejecting part of Friedman—the tests—while retaining full license on assumptions, we have become unmoored from any methodological restraint.

Boldrin and Levine have given us a model. It is a model that appears to have little to do with actual invention, and nothing to do with patentable innovation. Given that, nothing about the patent system or its consequences can serve as an impirical test. It is just a model. To be fair, it is what they promised in the title, "A Model of Discovery."