Proximate Cause and Patent Law

Amy L. Landers

Available at: https://works.bepress.com/amy_landers1/16/
Proximate Cause and Patent Law

Amy L. Landers*

ABSTRACT

The U.S. Supreme Court’s WesternGeco LLC v. ION Geophysical Corp. decision provides the opportunity to open a policy space for proximate cause in patent litigation. As the country’s highest Court has pushed patent law, as well as other statutory torts, in the same direction as other civil causes of action, the full adoption of proximate cause in all monetary forms of recovery appears to be inevitable. At this time, a full adoption of this complex, multidimensional requirement will allow the patent system to resolve several troubling existing problems, including the appropriate extraterritorial reach of the system. Properly implemented, proximate cause in patent law can resolve other difficulties that are explored in this article. These include accounting for the social benefits of infringing implementations, the multiple serial plaintiff problem, and the problem of patent infringement actions brought against end-user consumers as some examples. As the patent system joins other civil causes of action in implementing proximate cause, it is likely that other problems can be implemented and resolved so long as the standard is modified to account for the unique policy problems of the patent system.

TABLE OF CONTENTS

I. Introduction ........................................................................................................................................2
II. WesternGeco and Its International Implications ...........................................................................5
   A. The WesternGeco Decision and its Early Progeny .....................................................................5
   B. The International Impact of the WesternGeco Decision .................................................................9
III. Cause and the Patent System .........................................................................................................11
   A. Connection to the Harm ...............................................................................................................11
   B. Causation and the Reasonable Royalty Remedy ......................................................................13
IV. Rationales for Proximate Cause ......................................................................................................19

* Professor of Law and Director of the Intellectual Property Concentration, Drexel University Kline School of Law. The author wishes to thank participants of the 2018 Intellectual Property Scholars Conference (Berkeley School of Law) and the 7th Mid-Atlantic Patent Works-In-Progress (MAPWIP) Colloquium (American University - Washington College of Law) for thoughtful comments and feedback.
A. An Overview .......................................................... 19
B. Maximizing Socially Beneficial Activity ............................... 22
V. Proximate Cause: Forms of Analysis ...................................... 26
   A. Starting Points .................................................. 27
   B. Harms within the Scope of the Risk .................................. 32
   C. Intervening, Superseding and Efficient Causes ....................... 34
   D. Modifying Proximate Cause for Particular Causes of Action ......... 36
   E. The Appearance of Redundancy .................................... 37
VI. Proximate Cause and Statutory Causes of Action ...................... 39
   A. Overview .................................................................. 39
   B. Examples: Application of Proximate Cause to Statutory Torts ....... 42
VII. Proximate Cause and the Patent System ................................. 46
    A. Background ......................................................... 46
    B. Doctrinal Considerations ........................................... 47
VIII. Resolving Particular Problems in Patent Law ....................... 49
    A. International Considerations and the Scope of the Statute ........ 49
    B. Social Value: Internalizing the Benefits of Innovation ............. 50
    C. Apportionment and the Multiple Plaintiff Problem .................. 53
    D. Consumer Lawsuits ................................................ 56
IX. Conclusion .................................................................. 58

I. Introduction

    Although much has been written about patent infringement and patent damages, the causative relation between them represents comparatively unexplored territory. Patent infringement is a considered a form of tort that originates from the Patent Act statute. Generally, tort law requires that a plaintiff to show of both but-for and proximate cause before an element of recovery will be awarded. In the negligence realm, the mere fact that an accident occurred and that someone suffered an injury is not sufficient to support recovery. Rather, the law requires factual and legal cause between the defendant’s conduct and the plaintiff’s legally cognizable harm. As part of a line of cases analyzing statutory torts, the Supreme Court has held that...

---

proximate cause is presumed for statutory causes of action. More recently, the Supreme Court’s 2018 WesternGeco opinion refers to proximate cause as an undecided issue but provided no guidance into its implementation or structure. It is evident that a full analysis of proximate cause for the patent system is overdue. This article picks up that task and offers the analysis that has been lacking in patent law for incorporation of the doctrine going forward for all forms of monetary damages.

Proximate cause asks whether imposing liability in a particular case is consonant with the law’s purpose. Significantly, this label has two components. First, the phrase “proximate cause” refers to “cause,” which has been considered a question of fact that examines the relation between the tortious act and the plaintiff’s harm. Second, proximate cause is a policy based judgment that limits the legal consequences of a defendant’s actions even where those actions are illegal, tortious or wrongful. As one court states, “the phrase ‘proximate cause’ is shorthand for the policy-based judgment that not all factual causes contributing to an injury should be legally cognizable causes.” As described in Andrews’ Palsgraf dissent, this element has been put into place “because of convenience, of public policy, of a rough sense of justice, the law arbitrarily declines to trace a series of events beyond a certain point.”

Proximate cause is a common law doctrine was developed in the state courts within tort law, and primarily within the tort of negligence. It is not a model of clarity. Juries evidence confusion with the standard. There is no single clear definition, although some principles emerge. As one treatise describes “[t]here is perhaps nothing in the entire field of law which has called forth more disagreement, or upon which the opinions are in such a welter of confusion.” This has been attributed to the fact that the doctrine’s policies evolve and vary across jurisdiction, and factual context.

---

2 Lexmark Intern., Inc. v. Static Control Components, Inc., 134 S.Ct. 1377, 1390 (2014) (“...we generally presume that statutory cause of action is limited to plaintiffs whose injuries are proximately caused by violations of the statute.”).
8 Peter C. Haley, Paradigms of Proximate Cause, 36 Tort & Ins. L.J. 147, 149 (2000) (“tort actions have been the subject of most of the judicial and scholarly attention devoted to proximate cause”).
9 See Steele & Thornburg, Jury Instructions: A Persistent Failure to Communicate, 67 N.C.L.Rev. 77, 91(1988) (85% of actual and potential jurors were unable to understand a pattern proximate cause instruction); Robert P. Charrow and Veda R. Charrow, Making Legal Language Understandable: A Psycholinguistic Study of Jury Instructions, 79 Colum. L. Rev. 1306, 1353 (1979) (nearly one quarter of subjects misunderstood proximate cause to mean “approximate cause” or “estimated cause”).
11 Id. at 263-264; James Fleming Jr. and Roger F. Perry, Legal Cause, 60 Yale L. J. 761, 784 (1951) (“It should be noted at this point that many courts and legal writers have stressed the fact that policy considerations underlie the doctrine of proximate cause. Of course they do, but the policies actually involved often fail to get explicit treatment.”) (footnoted omitted).
Since its inception, proximate cause has been extended beyond common law torts to certain statutory causes of action. These include several with civil enforcement provisions, including antitrust, false advertising and securities fraud causes of action. Recently, a limited number of appellate patent cases considered proximate cause as an appropriate limitation on the types of damages awarded in patent cases concerning lost profits.12 There are few court cases with any rich analysis. In the reasonable royalty context, causative issues do not appear to be considered in awarding monetary damages except in limited circumstances. Although proximate cause has centuries of precedent in other areas of law, it is an under-theorized and under-developed area that holds significant promise for resolving some troubling phenomenon that has been observed over the past few decades. This work seeks to remedy that by examining the doctrine and how it might be applied to facilitate patent law's purposes in both the lost profits and reasonable royalty contexts. In addition, this Article provides several specific examples of instances where the doctrine of proximate cause can be applied to a patent infringement reasonable royalty analysis. Fundamentally, this preview of the application of proximate cause to the reasonable royalty context demonstrates the manner in which this doctrine can open up a policy space that will avoid some forms of injustice that are currently plaguing the system.

This article proceeds in several steps. First, the context of the recent cases in this area are examined. Second, the significant principles of proximate cause are outlined. Third, the Supreme Court’s application of proximate cause to statutory civil actions is examined. Fourth, this work examines several examples where proximate cause can be applied to patent liability determinations. Among this includes the extraterritorial reach of U.S. patent rights. Further, the doctrine can be used to fine-tune liability where the plaintiff’s contribution to the defendant’s implementation is minimal and the social benefits of the defendant’s infringement is significant and predominates. As a further example, proximate cause must be used to analyze damages where apportionment becomes complex where multiple plaintiffs seek recovery against the same infringing implementation to better adjust liability to accord with the purposes of the patent system. Finally, the doctrine can be used to limit liability where patents are asserted against numerous parties that include manufacturers, sellers and consumer users.

II. WesternGeco and Its International Implications

A. The *WesternGeco* Decision and its Early Progeny

Since the earliest part of the nineteenth century, it has been held that domestic patent law does not apply to conduct that occurs outside the country. As a consequence, federal statutes are presumed to apply to conduct within the U.S. This presumption against extraterritoriality is based on comity, a principle which recognizes that all nations are empowered to make the laws that govern which their territories. These decisions are driven by the concern that applying domestic law outside the nation’s borders imposes “the sovereign will of the United States onto conduct occurring within the territorial jurisdiction of another sovereign.” Unless there is an exception, federal statutes are presumed to have no territorial power to operationalize “the respect sovereign nations afford each other by limiting the reach of their laws.” A number of recent U.S. Supreme Court decisions confirm these long-held principles.

As the world has become more interconnected, considerations beyond comity have emerged. Treaties, protocols, and other forms of international cooperation require countries to create laws that are influenced by their international obligations. Under some circumstances, Congress has enacted laws that reach across our borders to reach conduct overseas. To determine whether any particular statute has extraterritorial effect, and thereby rebutting the presumption against extraterritoriality, the courts examine a two-part test developed in the Supreme Court’s *RJR Nabisco v. European Community.* First, the court examines whether the statute gives a clear, affirmative indication that it applies extraterritorially. If the answer is yes, the court may apply the statute to extraterritorial conduct. If the answer is no, then the court must examine the statute’s focus to determine whether the case involves domestic application of the statute. If the conduct that is within the statute’s focus took place in the U.S., then the statute will be deemed applicable even if some relevant conduct also occurred abroad. However, if this occurred in a foreign country, then this is an impermissible extraterritorial application of U.S. law and relief will be denied.

---

13 See, *e.g.*, The Schooner Exch. v. McFaddon, 11 U.S. 116, 147 (1812) (Marshall, C.J.) (“[T]he Exchange, being a public armed ship, in the service of a foreign sovereign . . . should be exempt from the jurisdiction of the country.”).
14 *Kiobel*, 133 S. Ct. at 1667; *See also* F. Hoffmann-La Roche Ltd. v. Empagran S.A., 542 U.S. 155, 167 (2004) (observing that “several foreign nations” have advised the Court that “to apply [U.S.] remedies would unjustifiably permit their citizens to bypass their own less generous remedial schemes, thereby upsetting a balance of competing considerations that their own domestic antitrust laws embody”).
15 *Hartford Fire Ins. Co. v. California*, 509 U.S. 764, 817 (1993); *see also* Am. Banana Co. v. United Fruit Co., 213 U.S. 347, 356 (1909) (“But the general and almost universal rule is that the character of an act as lawful or unlawful must be determined wholly by the law of the country where the act is done” and that any other result would be “contrary to the comity of nations”).
In June 2018, the U.S. Supreme Court decided a patent case which implicated some of these principles, WesternGeco LLC v. ION Geophysical Corp.\(^\text{18}\) The patentee Western Geco asserted four patents claiming technologies used to search for oil and gas under the ocean floor. The claims described improvements to a system that relied on long streamers equipped with sensors that were suspended underneath ships. These sensors were used to map the ocean floor, and the WesternGeco’s improvements contributed to the accuracy of their readings. The resulting maps were more accurate, and therefore more useful for locating offshore oil and gas reserves. Both WesternGeco and the defendant ION are competitors as each perform oil exploration services for their clients. At trial, a jury found that ION had infringed the claims of WesternGeco’s patent and awarded both lost profits and a reasonable royalty as damages. The lost profits damages were attributed to WesternGeco’s loss of ten sales contracts that the patentee lost to ION for services using the patented technology that were performed outside the U.S.

The legal foundation for the case is a series of statutes that define infringement and the extraterritorial reach of U.S. patent law. As a general matter, section 271(a) is the primary section that enumerates the operative acts that constitute infringement—specifically, making, using, offering to sell, selling or importing a device incorporating the invention. The Supreme Court has established that the predicate acts for infringement under this subsection are sufficient only if those acts take place occur in the United States.\(^\text{19}\) This was established in the U.S. Supreme Court’s 1972 decision Deepsouth Packing Co. v. Laitram Corp.\(^\text{20}\). There, the claimed invention was a machine used to devein shrimp. The defendant argued that manufacturing individual machine parts in the U.S. and then shipping them to customers in “three separate boxes, each containing only parts of the machine” for re-assembly by the end users overseas was not infringement under section 271(a).\(^\text{21}\) The Supreme Court agreed. As the Court explained, making an infringing device requires manufacture of “the operable assembly as a whole” and not just its parts in the U.S. Because assembly here was done abroad, there was infringement. Absent an applicable exception, extraterritorial conduct does not infringe section 271(a) even if that conduct would infringe if it occurred in the U.S. This principle has been reconfirmed in a number of cases over the years.\(^\text{22}\)

Deepsouth is consistent with the “basic premise of our legal system that, in general, United States law governs domestically but does not rule the world.”\(^\text{23}\) Subject to some exceptions, a U.S. patent’s power is territorial. Generally, this rule is based on sovereignty, which permits all nations to create the laws that govern within their borders.\(^\text{24}\) Subject to international treaties and obligations, each country can shape its own patent policy and “the laws of others have no claim on it.”\(^\text{25}\) To permit U.S. law to delineate the activity in foreign

\(^{19}\) Deepsouth Packing Co. v. Laitram Corp., 406 U.S. 518 (1972).
\(^{20}\) Id.
\(^{21}\) Id. at 524.
\(^{24}\) Id. at 198.
\(^{25}\) Id.
jurisdictions and infringing, or not, threatens these nation’s ability to craft their own policy choices, issue patents based on their individual standards, and consider which activities should be either covered by—or exempt from—private patent rights. This circumstance avoids the possibility for double patent recovery where foreign activity may infringe both U.S. and foreign patent rights. Further, sovereignty gives other nations the policy space needed to determine their own damages rules that reflect their individual policy decisions based on local conditions.

In addition to the geographic scope of the subsection, section 271(a) subsection requires strict identity—that is, all claim elements must appear in the accused device to support a finding of infringement. In this case, ION did not engage in a predicate act with respect to a device that included all elements of WesternGeco’s asserted claims. Instead, ION made some component parts in the U.S. which were assembled abroad. Critically, ION’s component parts that were made in the U.S. did not infringe WesternGeco’s claims. Rather, these parts must be combined with the entire assembly to meet the strict identity test. Because ION’s combination occurred overseas, section 271(a) was not an appropriate basis to find infringement due to the extraterritorial restrictions on the reach of U.S. patent law.

In *WesternGeco*, the jury found infringement under section 271(f)(2). This section is a limited exception to patent’s prohibition against the extraterritorial reach of a domestic patent. Under this subsection, infringement is established against one who supplies components specially designed to infringe and intends them to be combined overseas into an infringing system. Although some might consider that this subsection overrules *Deepsouth*, that is not an accurate reading. Rather, *Deepsouth*’s territorial limitation continues to govern section 271(a). It is more accurate to say that section 271(f) provides an alternative for establishing infringement on facts analogous to those at issue in *Deepsouth*—that is, shipping parts overseas for assembly there. By fixing the infringing act as the manufacture of the component parts in the U.S., section 271(f) sidesteps the sovereignty problem by focusing the operative act of infringement on U.S.-based activity.

---


28 Microsoft Corp. v. AT & T Corp., 550 U.S. 437 (2007); Dowagiac Mfg. Co. v. Minnesota Moline Plow Co., 235 U.S. 641, 650 (1915) (“The right conferred by a patent under our law is confined to the United States and its territories, and infringement of this right cannot be predicated of acts wholly done in a foreign country.”) (citation omitted).

29 35 U.S.C. § 271(f)(2) states as follows:

> Whoever without authority supplies or causes to be supplied in or from the United States any component of a patented invention that is especially made or especially adapted for use in the invention and not a staple article or commodity of commerce suitable for substantial noninfringing use, where such component is uncombined in whole or in part, knowing that such component is so made or adapted and intending that such component will be combined outside of the United States in a manner that would infringe the patent if such combination occurred within the United States, shall be liable as an infringer.

30 *WesternGeco* LLC v. ION Geophysical Corp., 138 S. Ct. 2129, 2138 (2018) (“The conduct that §271(f)(2) regulates—i.e., its focus—is the domestic act of “supply[ing] in or from the United States.”).
The jury found that ION’s manufacture of certain parts in the U.S. and subsequent shipment of those parts overseas for assembly into an infringing combination satisfied section 271(f)(2). In addition to a reasonable royalty award for supplying the parts overseas, the jury awarded WesternGeco lost profits from contracts that ION had performed abroad because “[a]ccording to WesternGeco, ION's customers would not have been able to win the contracts if they did not have access” to the claimed technology.\(^{31}\) Reviewing this portion of the award, the Federal Circuit held that the lost profit award for these foreign contracts was improper, reasoning that the overseas performance of contracts was extraterritorial and therefore beyond the scope of the U.S. patent right.\(^{32}\) This court explained “[j]ust as the United States seller or exporter of a final product cannot be liable for use abroad, so too the United States exporter of the component parts cannot be liable for use of the infringing article abroad.”\(^{33}\)

The Supreme Court reversed. The Court’s bypassed the first prong of the *RJR Nabisco v. European Community* test and instead examined the second, finding that the focus of section 271(f)(2) is on the predicate act of supplying the component from the U.S.\(^{34}\) Here, the Court determined that ION’s shipment of components of the claim for overseas assembly was within this scope and that therefore section 271(f)(2) has extraterritorial effect. In doing so, the *WesternGeco* Court read 271(f)(2) in connection with the statute governing patent damages, section 283, which reads in relevant part “the court shall award the claimant damages adequate to compensate for the infringement.”\(^{35}\) Finding that lost profits for the overseas contracts was permissible, the Court explained that the lost contracts were “merely incidental to the infringement.”\(^{36}\) Essentially, the Supreme Court conceptualized the foreign sales as connected (that is, “incidental”) to the initial act of infringement of the domestic manufacture of parts, which was the first infringing act under section 271(f). In reaching this determination, the Court expressly refrained from considering whether the doctrine of proximate cause might cut off liability for the lost overseas contracts.\(^{37}\)

Since the *WesternGeco* opinion was issued, a district court applied this holding to damages awarded under section 271(a) in *Power Integrations, Inc. v. Fairchild Semiconductor International*.\(^{38}\) As previously discussed, this subsection has been held by the Supreme Court, in prior cases, to have domestic effect only and was issued years after the Federal Circuit had rejected the patentee Power Integrations’ request for a mix of lost profits and reasonable royalty

---


\(^{32}\) Id. at 1350; see also Power Integrations, Inc. v. Fairchild Semiconductor Int’l, Inc., 711 F. 3d 1348 (CA Fed. 2013).


\(^{34}\) The *WesternGeco* Court explained that although “it will usually be preferable to begin with step one, courts have the discretion to begin at step two in appropriate cases.” WesternGeco LLC v. ION Geophysical Corp., 138 S. Ct. 2129, 2136 (2018).


\(^{36}\) Id. at 2138 (citations omitted).

\(^{37}\) Id. at 2139 n. 3.

damages based on worldwide sales.\textsuperscript{39} Essentially, patentee Power Integrations had argued that lost worldwide sales were the foreseeable result of domestic infringement and “once a patentee demonstrates an underlying act of domestic infringement, the patentee is entitled to receive full compensation for ‘any damages’ suffered as a result of the infringement.”\textsuperscript{40} In this case, the patentee argued that in this industry the nature of the sales and manufacturing solution meant that “a worldwide solution was essential” and therefore “but for” Fairchild's U.S. infringement, Power Integrations would have maintained 100% of [the] worldwide business.”\textsuperscript{41} In 2013, the Federal Circuit rejected this argument, finding that worldwide sales were not recoverable under the principle that the patent laws did not apply to foreign sales.\textsuperscript{42} After the Supreme Court’s \textit{WesternGeco} opinion was issued, the district court in Power Integrations determined that the Federal Circuit’s opinion was effectively overruled.\textsuperscript{43} This district court opinion explained, “Fairchild has identified no persuasive reason to conclude that the interpretation of § 284 should differ here from what was available in \textit{WesternGeco} just because the type of infringing conduct alleged is different. Instead, as Power Integrations puts it, “Section 271(a) ‘vindicates domestic interests’ no less than Section 271(f).”\textsuperscript{44}

\textbf{B. The International Impact of the \textit{WesternGeco} Decision}

The implications of the award of worldwide patent damages that these cases allow represents a significant shift from prior law. The \textit{WesternGeco} Court’s reasoning was based on infringement under section 271(f), which contemplates some extraterritorial activity to the more expansive reliance on section 271(a) by the Power Integrations court warrants attention, particularly in light of the long-held precedent that section 271(a) has been found to have purely domestic applicability under \textit{Deepsouth} and its progeny. For patent owners, these decisions allow the collection of higher damages awards for conduct that has a U.S. basis. More broadly, from a public policy perspective, permitting the award of worldwide damages based on U.S. conduct implicates the sovereignty of other nations in formulating patent policy. By allowing private parties to obtain patents under U.S. standards, and permit them to obtain damages under the standards formulated under U.S. law, the patent system erases each nation’s ability to fine-tune the patent system within their borders to incentivize—or disincentive—particular activity. In other words, if a nation finds that broad experimental use fosters desirable activity, that policy judgment cannot be cleanly implemented in multinational companies must pay U.S.-based judgments that render if more expensive to do in that country.

Among other things, the concept of worldwide damages raises troubling policy concerns incorporated into the Agreement on Trade-Related Aspects of Intellectual Property Rights (“TRIPS Agreement”). This Agreement mandated that each of the over one hundred and sixty

\textsuperscript{39} Power Integrations, Inc. v. Fairchild Semiconductor Int'l, Inc., 711 F.3d 1348, 1371 (Fed. Cir. 2013) (“It is axiomatic that U.S. patent law does not operate extraterritorially to prohibit infringement abroad.”).
\textsuperscript{40} Id. at 1370.
\textsuperscript{42} Power Integrations, Inc. v. Fairchild Semiconductor Int'l, Inc., 711 F.3d 1348, 1371 (Fed. Cir. 2013).
\textsuperscript{44} Id.
members institute a patent system to facilitate worldwide trade. It has an open-ended framework and incorporates a series of flexibilities that allows each member to institute a patent system that fits local priorities.\textsuperscript{45} As stated previously:

\ldots the TRIPS Agreement was conceived to overcome the territoriality of domestic patent systems. At the same time, the TRIPS Agreement was intended to allow other nations flexibility in implementing the specifics of their patent systems to fit their local cultures, economies, and legal systems. By authorizing other sovereigns to enact individualized patentability standards, the TRIPS Agreement permits all members to formulate their own answers to the policy questions that arise in patent cases.\textsuperscript{46}

Imposing remedies on those researching and innovating in foreign jurisdictions circumvents these nations’ ability to shape their own patentability standards to suit their own economies. Under the TRIPS Agreement, such flexibilities are expressly permitted and point toward the optimal path toward maximizing the aggregate global level of invention. Further, this circumstance invites other nations to reciprocate and increase damages for U.S.-based activity to the extent that it implicates a foreign patent. The applicability of the \textit{WesternGeco} to other forms of relief, including the reasonable royalty, are unclear.

More broadly, the \textit{WesternGeco} decision adds another layer to the debate about the appropriate contours of an infringer’s liability. Within this context, it is critical that the element of cause be considered as critically as other elements and as thoughtfully it is in the tort context. This emerging line of case law presents an opportunity to open up a policy space to meaningfully examine whether damages should be imposed for particular types of conduct. As previously mentioned, the \textit{WesternGeco} Court provides at least one limitation, recognizing that “[i]n reaching this holding, we do not address the extent to which other doctrines, such as proximate cause, could limit or preclude damages in particular cases.”\textsuperscript{47} \textit{WesternGeco}’s statement calls for a closer examination of cause in the patent litigation and is consistent with other cases that have begun to recognize cause as an important aspect of patent law.

As one example, an infringement finding based on active inducement must be based on a finding that the defendant’s conduct caused the patent infringement.\textsuperscript{48} In assessing the secondary considerations for obviousness, there must be a nexus between the commercial success of the product incorporating the invention and the claim at issue.\textsuperscript{49} To obtain an injunction, a patentee must demonstrate a causal nexus between the infringement and the products that are the subject matter of the court’s order.\textsuperscript{50} As another, a reasonable royalty must be based on a causal

\textsuperscript{46} Landers, \textit{supra} note 27 at 36.
\textsuperscript{47} \textit{WesternGeco}, 138 U.S. at 2139 n. 3.
\textsuperscript{49} Ormco Corp. v. Align Tech., Inc., 463 F.3d 1299, 1312 (Fed. Cir. 2006).
\textsuperscript{50} Macom Tech. Sols. Holdings, Inc. v. Infineon Techs. AG, 881 F.3d 1323, 1330 (Fed. Cir. 2018) (“to satisfy the irreparable harm requirement, there must be a causal nexus between conduct shown to be wrongful and the alleged harm. This requirement “ensures that an injunction is only entered against a defendant on account of a harm
connection between the award and the infringing features of a product or process.\textsuperscript{51} For the exception to the rule to apply to allow the patentee to recover for a royalty based on the entire product, the patented feature must have a causal connection to the demand for the entire device.\textsuperscript{52} Underlying these decisions is the concept that, just as in tort law, liability does not exist in a vacuum. Rather, liability for patent infringement requires compensation for a particular kind of harm. That analysis that supports rule includes the elements and proximate cause, which provide a number of opportunities to better align liability findings with the purposes of the law.

III. Cause and the Patent System

Proximate cause is an umbrella term that has two components—first, cause in fact and second, (and somewhat confusingly) proximate cause.\textsuperscript{53} Cause (sometimes called “cause-in-fact” or “factual cause”), which is typically a jury question, asks whether the injury would have occurred in the absence of defendant’s tortious conduct.\textsuperscript{54} Although there are exceptions for cases where this test is not an appropriate fit, the general rule queries whether the plaintiff would have been harmed “but for” the defendant’s conduct. Cause-in-fact requires the plaintiff to show that a defendant’s conduct make a material or substantial contribution to the plaintiff’s injury. This showing can be made even if other innocent causes or another’s tortious behavior are contributing factors.\textsuperscript{55}

A. Connection to the Harm

Proximate cause is presumed to apply to civil statutory causes of action.\textsuperscript{56} The limitation has been applied to a number of federal claims.\textsuperscript{57} As the Supreme Court has explained, “[t]hat venerable principle reflects the reality that the judicial remedy cannot encompass every

\begin{footnotesize}
\begin{enumerate}
\item Power Integrations, Inc. v. Fairchild Semiconductor Int'l, Inc., 904 F.3d 965, 977 (Fed. Cir. 2018) ("A patentee is only entitled to a reasonable royalty attributable to the infringing features.")
\item LaserDynamics, Inc. v. Quanta Computer, Inc., 694 F.3d 51, 67 (Fed. Cir. 2012).
\item In some negligence contexts, the term “proximate cause” has been referred to as simply “scope of the risk.” As risk is the foundation of a negligence cause of action, that terminology has not been used in this piece.
\item Keeton et al., supra note 10 at 264. Doctrinally, cause-in-fact is sometimes placed under the umbrella of proximate cause. \textit{Id.} This work excludes cause-in-fact from the proximate cause analysis.
\item Leon Green, Proximate Cause in Texas Negligence Law, 28 TEX. L. REV. 471, 475 (1950) (“There may be and usually are several factors contributing to a plaintiff's hurt, but it is not required that a defendant's negligence be the only cause of the hurt.”).
\item Lexmark Intern., Inc. v. Static Control Components, Inc., 134 S.Ct. 1377, 1390 (2014) (“...we generally presume that statutory cause of action is limited to plaintiffs whose injuries are proximately caused by violations of the statute.”).
\end{enumerate}
\end{footnotesize}
conceivable harm that can be traced to alleged wrongdoing.” 58 This, it is inherent in any compensatory system that a plaintiff demonstrates a causal connection between the defendant’s wrongful conduct and the plaintiff’s harm. As a consequence, as a legal system “we are willing to exonerate a suspected person whenever we decide that his conduct ‘had nothing to do with’ the event in which we are interested.” 59 Essentially, a wrongdoer is not responsible to pay damages unless there is some level of certainty that the plaintiff’s harm would not have occurred but for the defendant’s conduct. To effectuate the law’s goals, causation serves the goals of compensation as well as determining that the appropriate party to pay damages is the one who caused the plaintiff’s injury. 60 In other words, a plaintiff’s compensation depends on identifying a defendant’s particularized conduct that led to the injury. A jury’s understanding of cause in fact controls and that body’s knowledge of cause and effect may be critical to their determinations. 61

Patent law incorporates this compensatory purpose to the liability framework for patent infringement. The patent statute governing monetary relief states that a patentee is entitled to “damages adequate to compensate for the infringement, but in no event less than a reasonable royalty for the use made of the invention by the infringer.” 62 The phrase “damages adequate to compensate for the infringement” makes clear that the purpose of monetary relief for patent infringement is compensatory. 63 As in the common law, the concept behind these measures of compensation is that monetary relief is available for an infringement that leads to the patentee’s legally cognizable harm. This necessarily requires a causal relation between the injury and the infringement. Compensation, which places the patentee in the same position that it would have been absent the defendant’s infringement, depends on ensuring that the defendant’s injury contributed to the harm. 64

This standard has been traditionally applied where lost profits are sought. These awards allow patentees recovery for money that would have been earned if the defendant had not entered the market with a competing infringing product, including lost sales, price erosion and convoyed sales that might have been made absent competition from the infringer. It is well-established that to recover lost profits, the patentee must demonstrate a causative link between the monetary loss

61 William E. Nelson, From Fairness to Efficiency: The Transformation of Tort Law in New York, 1920-1980, 47 BUFF. L. REV. 117, 131–32 (1999) (observing a shifting historic trend that “[a]s juries and others adopted the new paradigm holding that victims of injury should receive compensation from some source, they simultaneously rejected the older, nineteenth century world view that injury, death and other sudden calamities were inevitable, random and frequent events attributable to cosmic rather than human agency”).
63 See Aro Mfg. Co. v. Convertible Top Replacement Co., 377 U.S. 476, 507 (1964) (the statutory measure of damages is “the difference between [the patent owner’s] pecuniary condition after the infringement, and what his condition would have been if the infringement had not occurred.”)
64 See Leon Green, The Causal Relation Issue in Negligence Law, 60 MICH. L. REV. 543, 548 (“The inquiry is limited to the fact of defendant’s contribution to the injury.”).
and the infringement. Yet there is no reason to limit its applicability to lost profits, given that reasonable royalty awards are also compensatory in their purpose.

B. Causation and the Reasonable Royalty Remedy

A second form of damages for patent infringement is the reasonable royalty, which was developed in the legal courts prior to their merger with the courts of equity. During the nineteenth century, a patentee was entitled to recover damages in the amount of an established license rate, if one existed. Over time, the courts began to recognize that—in the absence of evidence establishing lost profits or a customary licensing rate—a patentee could be awarded a reasonable royalty that reflected the value of the patented invention used by the infringer. This remedy relies on the principle that a legally cognizable harm can include the mere act of infringement of the patent right, even if the patentee had suffered no pecuniary loss in the market. This allows patentees who do not sell products to obtain monetary recovery to compensate for the use of the claimed invention. The mere violation of the patent right is the identified harm, unlike lost profits which redresses monetary loss where a competitor infringer sells products and the patentee therefore loses sales that it would have made absent the infringement. As one court described, “the reasonable royalty theory of damages . . . seeks to compensate the patentee not for lost sales caused by the infringement, but for its lost opportunity to obtain a reasonable royalty that the infringer would have been willing to pay if it had been barred from infringing.”

Since the merger of law and equity courts, any patentee may elect to recover either lost profits or a reasonable royalty including those who do—and those who do not—sell product. With a proper evidentiary showing, a patentee is free to demonstrate entitlement to recover monetary damages under both theories. Reasonable royalty awards been the most widely sought-after form of monetary damages, and is available in all patent cases. As a practical matter, patent owners favor reasonable royalty recovery because the standards for proof are less exacting

65 Yale Lock Mfg. Co. v. Sargent, 117 U.S. 536, 552–53 (1886); see also Rite-Hite Corp. v. Kelley Co., 56 F.3d 1538, 1545 (Fed. Cir. 1995)(“the patentee must show a reasonable probability that, ‘but for’ the infringement, it would have made the sales that were made by the infringer.”); Minco, Inc. v. Combustion Eng’g, Inc., 95 F.3d 1109, 1118 (Fed. Cir. 1996); King Instruments Corp. v. Perego, 65 F.3d 941, 952 (Fed.Cir.1995), cert. denied, 517 U.S. 1188 (1996).
67 Dowagiac Mfg. Co. v. Minnesota Moline Plow Co. 235 U.S. 641, 649 (1915) (allowing patentee recovery where “the only measure of damages was such sum as, under all the circumstances, would have been a reasonable royalty for the defendant to have paid”).
68 AstraZeneca AB v. Apotex Corp., 782 F.3d 1324, 1334 (Fed. Cir. 2015).
69 See Zelin Yang, Damaging Royalties: An Overview of Reasonable Royalty Damages, 4 Berk. Tech. L.J. 647, 648 (2014) (“Reasonable royalties are the most common form of damages, accounting for eighty-one percent of the damages awards over the last six years”); See also Christopher B. Seaman, Reconsidering the Georgia-Pacific Standard for Reasonable Royalty Patent Damages, 5 BRIGHAM YOUNG UNIV. L. REV. 1661, 1667 (2010)(“a reasonable royalty recently has become the most commonly employed method for calculating patent infringement damages, supplementing or replacing the traditional remedy of the patentee’s lost profits.”).
and the overall recovery tends to be higher. Yet, from a policy perspective, this form of recovery has been criticized as moving too far from the statute’s compensatory purposes. This may be because, beyond problems calculating the reasonable royalty, causal links between the infringement and plaintiff’s harm remain virtually unexamined.

Although this requirement can be done routinely without changing the result in all cases, in others its application can be outcome determinative. In Ericsson, Inc. v. TCL Communications Technology Holdings (“TCL”), the district court incorporated a causal requirement for lump-sum reasonable royalty award. There, Ericsson demonstrated that TCL’s use of the Android operating system infringed a number of its patent claims. Ericsson’s damages verdict included $75 million as a lump sum royalty, which was based on expert testimony that included a component for the time period that ran from the date of trial until the patents expired. The Ericsson court explained that the award was erroneous, in part because it was based on projected losses and that “projected future loss must be factually and proximately caused by a wrong (such as a tort) that has already occurred.” Finding “there is no exception to these principles in patent law,” the court explained:

Although a lump sum royalty includes compensation for projected future infringement, the patent holder must establish that but-for infringement that has already occurred with respect to a certain product, the infringer would have paid a lump sum royalty for a license to keep selling that certain product through the life of the patent. In other words, what the patent holder actually lost, because of a party’s past infringement, was a lump sum royalty. The lump sum may be based on developments that have occurred after the date of the hypothetical negotiation, including realistic projections.

---

70 See generally PWC, Patent Litigation Study 2018, 6 (May 2018) available at https://www.pwc.com/us/en/services/forensics/library/patent-litigation-study.html; Mark A. Lemley, Distinguishing Lost Profits From Reasonable Royalties, 51 WM. & MARY L. REV. 655, 667 (2009)(“the situation has gotten so bad that some patentees who can prove lost profits elect instead to seek a ‘reasonable’ royalty that is far in excess both of what the parties would have negotiated and of the actual losses the patentee suffered.”).

71 See Brian Love, The Misuse of Reasonable Royalty Damages as a Patent Infringement Deterrent, 74 MO. L. REV. 909, 923-924 (2009); see also Seaman, supra note 69 at 1666 (“there is a growing body of evidence that Georgia-Pacific has resulted in the systematic overcompensation of patent owners in certain industries.”).

72 Certainly, the reasonable royalty analysis has woven some causation concepts into the application of the Georgia-Pacific test. To determine the royalty base for a multicomponent product, the general rule is that the royalty is limited to the value of the infringing feature of the entire product and nothing more. Ericsson, Inc. v. D-Link Sys., Inc., 773 F.3d 1201, 1226 (Fed. Cir. 2014). This requirement is intended to ensure that the compensation award is connected to the defendant’s infringing use and not on the defendant’s sale of products more generally. Lemley, supra note 70 at 662 (recognizing that reasonable royalty awards are available to all patentees, including those who produce and sell products). In a similar vein, admissible licenses to demonstrate the amount of compensation must have a relation to the patented technology. LaserDynamics, Inc. v. Quanta Computer, Inc., 694 F.3d 51, 80 (Fed. Cir. 2012).


74 Id. at 13.
of future sales, but the lump sum must be based on a product accused in the lawsuit and found by the jury to infringe.\textsuperscript{75} Here, the court found that the award was not supported by evidence because it was based on products that “do not yet exist or that have never been found to infringe.”\textsuperscript{76} Thus, any causal connection between the plaintiff’s asserted harm and the monetary damage was lacking. Perhaps it is not surprising that the jury reached this erroneous conclusion, given that the law does not clearly call out the need for this filter before juries decide damages awards during trial.

In another district court opinion, causation was used to reduce an inflated reasonable royalty set forth in an expert report as unsustainably high.\textsuperscript{77} There, the court determined that royalty damages could not be set based on draft negotiation documents for a deal that had never been consummated. As the court explained, the amounts discussed in those documents were not sufficiently reliable to demonstrate a royalty figure that the plaintiff would have obtained but for the infringement. In other words, those negotiation documents were not reflective of any amount necessary to place the patentee “in the pecuniary condition he would have enjoyed but for the defendant's infringement,” which is a classic causation analysis.\textsuperscript{78} It is evident from these examples that courts are capable of applying causative concepts to liability and harm concepts, either whole or in part, to the reasonable royalty award. Yet without a focus on causation throughout the patent system, fine-tuning the relation between liability and damages, and ensuring just results, has been essentially foreclosed.

Some might argue that patent infringement is akin to the tort of trespass, which permits presumed damages for violation of the right of exclusive possession.\textsuperscript{79} Such an analogy, to the extent that it is apt, might rest on the assumption that the mere violation of the patent right is sufficient to demonstrate entitlement to a reasonable royalty without consideration of any causative connection to a pecuniary loss. Although contrary to all of tort law, the argument might continue that, because the violation of the right is the harm, then the link to the reasonable royalty is established merely by a violation of the right.

This argument ignores the statutory structure of the Patent Act. Section 284, which defines the entitlement to a reasonable royalty, is a code section which sets forth various measures of available damages. Section 284 does not define civil liability or the patentee’s harm. Rather, section 281, which defines civil liability, does not contain any language that would rebut the presumption for a causative requirement as a matter of determining liability.\textsuperscript{80}

\textsuperscript{75} Id. at 14 (citations omitted); Allergan Sales, LLC v. UCB, Inc., No. 215CV01001JRGSP, 2016 WL 8222619, at *2 (E.D. Tex. Nov. 7, 2016) (stating rule, “Both a running royalty and a lump sum royalty, however, compensate the patent holder for loss resulting from infringement that has already happened. Although a lump sum royalty includes compensation for projected future infringement, the patent holder must establish that but-for infringement that has already occurred, the infringer would have paid a lump sum royalty.”).

\textsuperscript{76} Id. at 14.


\textsuperscript{78} Id. at *2.

\textsuperscript{79} See Adam Mossoff, The Trespass Fallacy in Patent Law, 65 Florida L. Rev. 1687, 1693 (2013) (outlining the historical origins of the analogy while criticizing that it is applicable).

\textsuperscript{80} 35 U.S.C. § 271 (“A patentee shall have remedy by civil action for infringement of his patent.”).
Supreme Court’s direction, the Federal Circuit has acknowledged that, as a general matter, “express language is not required” because “statutes speak in general terms rather than specifically expressing every detail.”\textsuperscript{81} This section was adopted against a backdrop of common law liability, which overwhelmingly applies the principles of cause in fact and proximate cause as a prerequisite for monetary recovery.

Further, to the extent the trespass analogy might be application, common law trespass actions recognize that the mere violation of the right of exclusive possession do not lead to a full panoply of remedies. For mere violations of the right, a property owner obtains a nominal figure.\textsuperscript{82} In order to recover for any actual pecuniary loss, the landowner must demonstrate a causative connection to the amount of the actual damages suffered.\textsuperscript{83} Thus, a reasonable royalty figure for the mere violation of the patent right does not, absent a showing a causation to the harm, render entitlement to a full reasonable royalty per se available. Moreover, in the tort context, some recent decisions have adopted a rule that de minimis invasions do not warrant recovery of nominal damages at all.\textsuperscript{84}

It is remarkable that patent law has, as a whole, ceded this policy space in a way that is contrary to the manner in which other civil cases are decided. As a practical matter, a patentee should not be able to recover a reasonable royalty for technology that the defendant added that are outside the claim scope as a matter of causation law. Therefore, to the extent that the value of the technology has been enhanced, the patentee cannot demonstrate entitlement for recovery for those improvements. Although one might argue that, at least for cause in fact, a jury question does not permit consideration of policy questions and therefore nothing has been lost for the failure to include it. Yet as Professor Wes Malone recognized, judges exercise considerable control that allows policy to develop in the cause in fact inquiry.\textsuperscript{85} Through pre- and post-trial motion practice and evidentiary rulings, courts can shape the information that the jury hears and provides some level of scrutiny to their decisions. Today, challenges to expert testimony on causality issues can be used to prevent irrelevant or unsupported theories from reaching the jury in tort cases. Appellate courts act as another check on the jury’s finding. Although the majority of cases do not require judicial oversight of cause in fact, tools exist for doing so.

\textsuperscript{81} Rite-Hite Corp. v. Kelley Co., Inc., 56 F.3d 1538, 1547 (1995) (applying a proximate cause analysis to lost profits).
\textsuperscript{83} See Allstate Fire Ins. Co. v. Singler, 14 Ohio St. 2d 27, 29 (1968); Woodbury v. Whitmire, 246 Ga. 349, 350 (1980).
As Malone further explains, policy operates within cause in fact in under-appreciated ways.\footnote{Id. at 60-61.} Some of this work is performed in assessing the counter-factual question of what might have happened in the absence of the defendant’s wrong. This inquiry invites speculation, yet it is a necessary component of the inquiry. Malone points that, facts are presented to the courts but the court must determine their legal sufficiency. The law provides a wide range of acceptable options. As one example, a court accepted that an employer was liable for harm caused to a plaintiff who fell on a coffeepot at the workplace during an epileptic seizure because its presence increased the risk of the possibility of harm.\footnote{Id. at 60-61.} One dissent to this decision argued that “[t]he coffee pot was part of the resulting injury, not its cause.”\footnote{Id. at 977 (Hodges, J. dissenting).} A separate dissent asserted, “[n]o connection is urged (and none was shown) to indicate that the presence of the coffee maker in the perimeter of claimant's fall increased the severity of harm from her fall.”\footnote{Id. at 977 (Opala, J. dissenting).} This range of opinions on the same factual proof illustrates Malone’s point that the legal standard can be controlled to implement liability and compensation policies.

He argues that the failure to comply with a law requiring fire escapes raises a host of causation issues that are decided based on policy perspectives.\footnote{Malone, supra note 85 at 79.} The issue becomes more complex if the rule does not specify the location. As Malone explains, a defendant may show other locations that would have complied with the law but were too far from the victim to make any difference to the outcome.\footnote{Id. at 78.} Other factors, such as a blocked door or too much smoke, might have been overwhelming causes compared to the landlord’s failure to provide a fire escape.\footnote{Id. at 82.} In the tort context, Malone observed that courts have over time shifted their willingness to send cases to a jury where the defendant’s failure to have a life preserver is alleged to have led to the plaintiff’s injury. In other circumstances, the courts have developed causation rules, including burden-shifting and presumptions, that assist in resolving difficult questions of cause in fact. Perhaps one of the most well-known is Summers v. Tice, which shifts the burden of proof to joint tortfeasors where the injury was caused by concurrent joint tortfeasors.\footnote{Summers v. Tice, 33 Cal. 2d 80, 88 (1948).} There is room for policy to operate in the cause in fact space. As detailed in later sections of this article, proximate cause provides an even more compelling case for judicial intervention.

The existence of policy levers for causation in patent law already exist in the lost profit context. There, a patentee may demonstrate entitlement to lost profits through the testimony of customers who testify that they would have purchased the patentee’s products but for the defendant’s infringement. In the absence of such evidence, patent law has devised a causation test that relies on inferences based on a series of factual circumstances, called the Panduit test, to show but for causation to the patentee’s lost sales.\footnote{Panduit Corp. v. Stahlin Brothers Fibre Works, Inc., 575 F.2d 1152, 1156 (6th Cir.1978) (“To obtain as damages the profits on sales he would have made absent the infringement, i. e., the sales made by the infringer, a patent owner must prove: (1) demand for the patented product, (2) absence of acceptable noninfringing substitutes, (3) his

\begin{footnotesize}\footnotesize
86 Id. at 60-61.
87 Flanner v. Tulsa Pub. Sch., 41 P.3d 972, 976, as corrected (Feb. 13, 2002).
88 Id. at 977 (Hodges, J. dissenting).
89 Id. (Opala, J. dissenting).
90 Malone, supra note 85 at 79.
91 Id. at 78.
92 Id. at 82.
93 Summers v. Tice, 33 Cal. 2d 80, 88 (1948).
94 Panduit Corp. v. Stahlin Brothers Fibre Works, Inc., 575 F.2d 1152, 1156 (6th Cir.1978) (“To obtain as damages the profits on sales he would have made absent the infringement, i. e., the sales made by the infringer, a patent owner must prove: (1) demand for the patented product, (2) absence of acceptable noninfringing substitutes, (3) his
\end{footnotesize}
the *Panduit* test for multi-competitor markets. The courts have created another variation that vitiates the patentee’s ability to collect lost profits where the defendant had access to create a non-infringing alternative. In other words, in the lost profits context the courts have formulated rules that exercise policy control over the but-for causation requirement. As proximate cause has remained relatively unexamined for reasonable royalties, comparable work in that area has never been explored.

This is problematic for several reasons. For example, it is questionable whether overseas infringement has a causative relation to any legally cognizable harm. A U.S. patent’s legal effect ends at the country’s borders. As Arthur Ripstein explains “[i]t is a commonplace of legal analysis that not all harms are legal wrongs, and not all legal wrongs are harms.” Presuming entitlement to compensation skips a causative step as well as the opportunity to consider whether the patentee has suffered a legally cognizable harm. As a further example, where there are multiple actors there may be incidental infringements by some that are create such a de minimis harm to the plaintiff that they cannot be said to be a substantial factor to an injury. In other contexts, it is widely that “when the plaintiff has suffered nothing more than *de minimis* or trifling damages at the inception of the defendant's acts, the doctrine of *de minimis non curat lex* bars recovery.” Treating itself as exceptional from other torts, patent infringement leads to reasonable royalty recovery in every case without pausing to consider these circumstances.

All of the judicially-created rules for lost profits constitutes an appropriate use of common law rule-making that fills the gap between the portions of the patent act and section 284, which defines the foundations of monetary recovery. The presumption that all patent infringement, no matter how nominal, extraterritorial or contrary to patent policy, supports a finding of cause-in-fact to the reasonable royalty is noteworthy. As a practical matter, when this occurs in the tort context “[i]t should be quite obvious that, once events are set in motion, there is, in terms of causation alone, no place to stop.” In patent law, this represents missed opportunities to ensure that the purposes of patent law are met in the process of determining liability.

---

95 BIC Leisure Prod., Inc. v. Windsurfing Int'l, Inc., 1 F.3d 1214, 1218 (Fed. Cir. 1993).
101 Keeton et al., *supra* note 10 at 266.
IV. Rationales for Proximate Cause

A. An Overview

In the majority of cases, the application of factual cause does little to prevent limitless liability. Just as the flap of a butterfly’s wing in Brazil has been posited to set off a tornado in Texas, a slight wrong has the potential to impose disproportional liability on a single actor. As Professor Smith recognized, “[t]he fact that the damage would not have happened, ‘but for’ the commission of the defendant's tort, does not invariably justify the conclusion that the tort was the cause, in the legal sense.”102 The courts have developed proximate cause as a policy mechanism to cut off a defendant’s liability either entirely or for particular elements of damages.103 As one state court decision explained in 1865:

In strict logic it may be said that he who is the cause of loss, should be answerable for all the losses which flow from his causation. But in the practical workings of society, the law finds, in this as in a great variety of other matters, that the rule of logic is impracticable and unjust. . . . To visit upon them all the consequences of failure would set society upon edge, and fill the courts with useless and injurious litigation.104

Early legal scholarship recognized that a defendant’s contribution may be remote in space or time and, therefore, not sufficient to warrant the imposition of liability.105 An early example of this principle is the 1793 Ashley v. Harrison, which considered an action brought by the owner of a theater against a defendant who had slandered a singer who was scheduled to perform there.106 After the slander, the performer cancelled her appearance because she feared “being hissed and ill-treated.”107 The court rejected the theater owner’s claim against the slanderer. Lord Kenyon’s opinion observed that, although the performer might have a valid cause of action against the defendant, the theater owner’s harm was too remote. He explained that the singer’s fears might be groundless, and that the defendant’s actions were too removed to support relief. The Ashley case demonstrates that, although a but for cause of the lost theater revenue, the slanderer should not be held liable absent stronger proof of a direct connection between the risks created by making false statements and the theater’s lost revenue in this case.

102 Jeremiah Smith, Legal Cause in Actions of Tort, 25 HARV. LAW REV. 103, 109 (Dec., 1911).
103 See James Fleming Jr. and Roger F. Perry, Legal Cause, 60 Yale L. J. 761, 784 (1951) (noting that proximate cause is driven by policy, “but the policies actually involved often fail to get explicit treatment.”). (“One consideration which is common to all cases under any system is the practical need to draw the line somewhere so that liability will not crush those on whom it is put.”); John G. Fleming, The Passing of Polemis, 39 The Canadian Bar Review 489, 490 (1961).
105 Smith, supra note 102 at 109 (“The defendant's tort must be distinctly traceable as one of the substantial efficient antecedents; as having had a substantial share in subjecting plaintiff to the damage”).
107 Id.
This concept is echoed in the term “proximate,” which suggests either physical or temporal closeness, echoes the ancient concept of in jure non remota causa sed proxima spectator. One theory of proximate cause is exemplified by Justice Holmes’ 1918 advice prescribing that “the general tendency of the law, in regard to damages at least, is not to go beyond the first step.” This principle was explored in depth in the 1919 Joseph H. Beale’s Proximate Consequences of an Act. In this piece, Beale relied loosely on Francis Bacon’s principle of “in jure non remota causa sed proxima spectator” to recognize that various cases considered the proximate cause to be the most recent active force prior to the injury. Beale’s articulation of proximate cause rested primarily on the theory that courts have limited resources, and that efforts to trace all plausible causes to their ends would overwhelm them. As he describes, “the court can give to the tracing of the consequences of any particular act only its fair share of all the available time, considering the other acts which are waiting its attention.”

Over time, this requirement has begun to play a more complex policy role. As Justice Andrews explained in his famous dissent in Palsgraf v. Long Island Railroad, “[w]hat we do mean by the word ‘proximate’ is that, because of convenience, of public policy, of a rough sense of justice, the law arbitrarily declines to trace a series of events beyond a certain point.” Similarly, Leon Green asserts “[t]he doctrine denying recovery in such cases is really based on a so-called public policy, a balancing of interests, with the conclusion that it is better to deny protection to the interest involved under such circumstances than it is to undertake to give compensation under all the difficulties of the case.” Through a number of tests and sub-rules, the law has used proximate cause to do that work by fine-tuning legal responsibility to align legal responsibility with the relevant law.

108 Smith, supra note 102 at 106. This roughly translates to “[i]n law the near cause is looked to, not the remote one.”
111 Beale describes “We may therefore begin our investigation with the assumption that the immediate result of an active force is primarily the proximate result; and that if the principle of proximity is discoverable, it must be by some method of relating the defendant's act to the final active force.” Id. at 641.
112 Omissions and passive activity were included within Beale’s definition of proximate cause as follows:

The whole problem may therefore be stated thus: when is one responsible for the operation (a) of an active force which he has created; (b) of an active force which acts upon a passive force which he created, or upon a passive force which he was legally bound to change.

113 Id. at 641.
Historically, there were a number of asserted justifications for the doctrine. Among these are the difficulty in providing harms from remote causes. A nineteenth century treatise recognizes that the law considers a human actor’s ability to make a moral choice, rather than inanimate albeit dangerous conditions such as water or ice, as the critical factor in fixing liability on a defendant who is capable of reason. Another perspective recognizes the necessity for avoiding crushing liability. As this policy suggests, negligence protects plaintiffs against unreasonable risks. The law tolerates reasonable risks, and indeed such activity is viewed as socially valuable. Proximate cause can be applied to ensure that socially valuable activity is not chilled by onerous liability standards.

In a similar vein, in the 1980’s, Steven Shavell wrote that the causation inquiry should exclude liability for accidents that do not have a sufficiently high potential for incentivizing the defendant to meet the standard of care. He explains that this approach has some advantages, including lower administrative costs for litigating claims that are unlikely to have a deterrent effect. Second, Shavell points out that fine-tuning the analysis will avoid the possibility for crushing liability, which he defines as dis-incentivizing actors from engaging in socially useful conduct because of a fear of overwhelming compensatory costs. He argues that, at least in theory, a properly implemented causation system should exclude certain types of activity from liability so long as it does not impact a defendant’s motivation to exercise due care. Significantly, however, he recognizes that the principles that work in theory may fall down in practice because “injurers who do in fact act with due care or attempt to do so will sometimes be found negligent and have to pay damages.” This is so for a number of reasons, including the fact that errors systematically occur due to human nature and that courts may make mistakes in the over-broad application of due care standards. In this sense, he explains that negligence liability can take on characteristics of strict liability. Shavell argues that, as a consequence, proximate cause can and should correct these problems to ensure that the law operates to maximize social benefits of risk-creating conduct.

117 Marble v. City of Worcester, 70 Mass. 395, 398 (1855) (“The law however looks to a practical rule, adapted to the rights and duties of all persons in society, in the common and ordinary concerns of actual and real life; and on account of the difficulty in unravelling a combination of causes, and of tracing each result, as a matter of fact, to its true, real and efficient cause, the law has adopted the rule before stated, of regarding the proximate, and not the remote cause of the occurrence which is the subject of inquiry.”).  
118 F. WHARTON, A TREATISE ON THE LAW OF NEGLIGENCE 73-84 (2nd ed. 1878) (distinguishing conditions and causes, with the latter warranting a finding of proximate cause).  
122 Id. at 465.  
123 Id.  
124 Id. at 488 (“Other things equal, it is best to restrict the scope of liability in such a way that the incentive to take due care is reduced as little as possible”).  
125 Id. at 489.
B. Maximizing Socially Beneficial Activity

Some recent recognizes that negligence doctrine as applied in the courts is replete with harsh outcomes and errors. Broad duty standards coupled with a breach assessment that does not consider the social utility of the defendant’s conduct, when added to the open-ended cause in fact standard can result in liability that might discourage a defendant from engaging in valuable risk-creating enterprises.

Moreover, the legal standards for negligence law fail to account for social benefits that generated by the defendant’s conduct. In other words, the legal test for assessing reasonableness fails to consider the social good generated by the defendant’s behavior even where the defendant’s conduct is tortious. Absent proximate cause, the railroad might be liable for Mrs. Palsgraf’s injuries despite the attenuated relation between the original tort in assisting the passenger who held the package onto the train, and the scales which fell upon her. To hold the railroad liable for these remote injuries might cause railroads to refuse to carry passengers, or to allow passengers to bring packages onto the train, or to engage in other socially useful activity. This line of scholarship recognizes that the standard elements of a negligence claim can result in the over-deterrence of risk-creating activities and fails to account for their positive impact.

One such work is Mark Grady’s Proximate Cause Decoded. There, he points out that negligence is a harsh standard and impossible to satisfy each and every time that an activity is undertaken. Given the breadth of duty and the requirement to meet the objective standard of care, the fallibility of human nature inevitably leads to slip-ups even among those who typically seek to meet it. According to Grady, proximate cause functions to limit liability for some comparatively blameless activity. As he explains, “unless liability is limited in some way, people will avoid the activities in which their own breach of duty is predictably likely or especially costly, even if these activities are valuable to them and to the community.” As one example, he considers that the doctrine of intervening, superseding cause can be used to save a less blameworthy defendant from liability and thereby preserves the ability for that actor to engage in valuable forms of risk-creating conduct.

---

126 WILLIAM M. LANDES & RICHARD POSNER, THE ECONOMIC STRUCTURE OF TORT LAW, 236 (1987) (finding that in the real world of errors, excessive liability “will have misallocative effects.”); Omri Ben-Shahar, Causation and Foreseeability, in 3 Encyclopedia of Law and Economics 644, 651-652 (Boudewijn Bouckaert & Gerrit De Geest eds., 2000) (“Inasmuch as the application of the negligence rule is plagued with errors and uncertainties ... the unrestricted scope of liability can have the crushing effect”);


129 Id. at 294.

130 Id. at 294.

131 See also Steven Shavell, An Analysis of Causation and the Scope of Liability in the Law of Torts, 9 THE JOURNAL OF LEGAL STUDIES, 463, 500 (1980) (recognizing that holding a defendant liable for unforeseeable harms does not increase that defendant’s incentives to comply with the standard of care).
One example that Grady cites in support of his thesis is *Pittsburg Reduction Co. v. Horton*.\(^{132}\) The first tortfeasor, Pittsburgh Reduction, failed to safely dispose of some dynamite caps, which it openly stored near a schoolyard. A ten-year-old student of the school found them, took them home, and played with them for about a week. In fact, his mother picked them up after he was done playing at the end of the day and that although “she did not know what the shells contained, but she did know that they were shells for some kind of explosives, that her son brought them home, and that he played with them.”\(^{133}\) The student’s father, who also lived in the home, was familiar with blasting caps from his work at Pittsburgh Reduction’s competitor. After the week was over, the student traded the caps to a classmate, who was injured when it exploded when he tried to open one. The court held that the conduct of the student’s parents cut off the causal chain and rendered Pittsburgh Reduction free from liability.\(^{134}\)

In a world without proximate cause, the entity Pittsburgh Reduction would have been considered a concurrent cause and almost certainly be held liable. Storing dynamite caps in an unsecured way in close proximity to the school was a breach and one of the causes of the classmate’s injury. Yet, as Grady explains, the court’s decision excuses Pittsburgh Reduction’s error in light of the parents’ more reckless failure to prevent the injury. He reasons that, as a practical matter, holding both defendants liable would have meant that Pittsburgh Reduction would have paid the judgment given its deeper corporate resources to pay. In effect, the parents would escape liability despite their more predominant contribution to the injury. The use of proximate cause to limit the entities’ liability lightens Pittsburgh Reduction’s cost of doing business and places the full judgment on those who were in better position to appreciate the risk and prevent the injury. Such cases put meat on the bones of the use proximate cause to avoid crushing liability on the less culpable party and implicitly allows the social benefit of their activity to continue. As Grady explains, proximate cause can be used to terminate liability for such a defendant as against a subsequent tortfeasor in a principled, justifiable manner that seeks to carry out the purposes of negligence law—under his view, optimizing the level of socially desirable risk while holding others financially responsible where there are reasons to do so.

Another line of scholarship calls for a reassessment of social cost in assessing tort liability. These works consider tort’s economic purpose of internalizing the negative externalities of conduct if a defendant failed to undertake a reasonable precaution. For example, a car company may fail to optimize the set point to trigger an airbag to minimize potential injuries to passengers. In this example, we might assume that an airbag that deploys at a low rate of speed (for example, while the car is traveling under fifteen miles an hour) will cause more injuries to passengers from the force of deployment compared to a car that has no airbag at all. A company that fails to conduct proper testing to determine the appropriate set-point and fails to warn passengers will be found to be negligent and required to pay for a plaintiff’s injuries caused by these failures. Tort law has been justified by the principle that, unless the car company is forced to internalize the cost of the accidents that occur as a result of its negligence, it lacks sufficient incentives to take these reasonable precautions.\(^{135}\)


\(^{133}\) *Id.* at 649.

\(^{134}\) *Id.* (stating (“It established a new agency, and the possession by [the student] of the caps or shells was thereafter referable to the permission of his parents, and not to the original taking.”)).

Some recognize a distinction between private costs and social costs of tortious activity. To use the above illustration, assume that a car company fails to perform proper testing and creates a set-point for airbag deployment at fifteen or more miles per hour simply by guessing. A plaintiff who is seriously injured while a car is traveling at ten miles an hour might suffer a serious private cost, made worse by the fact that the plaintiff failed to wear a seatbelt. Scholars have observed that awarding this plaintiff full compensation forces the car maker to internalize a higher cost than the social cost of their accident unless the damages calculation accounts for the plaintiff’s failure to mitigate. In other words, the concept that the private cost for recovery is equal to the social cost of the defendant’s conduct, a common assumption underlying discussion about damages, is open to question.

More broadly, these theorists recognize gaps between private and social costs, and that the failure to recognize them can lead to under- and over-deterrence. Consider that the example car company has done a study that concludes that, in the aggregate, fewer people are seriously injured when the airbag deployment set-point is at fifteen miles per hour and up. However, our plaintiff is seriously injured while the car is traveling at ten miles per hour and, due to the peculiar nature of the way that the injury occurred, this plaintiff’s injury would have been avoided if the airbag had deployed at the lower speed (ten miles per hour) than the maker set (at fifteen miles per hour). Finding the car maker negligent under these circumstances internalizes the private cost of the accident but raises the social costs to others who might be more severely injured at collision speeds between ten and fifteen miles per hour if the car maker lowers the airbag deployment set-point in response to the lawsuit. In other instances, this gap that might cause the car maker to absorb higher private costs compared with the socially optimal solution for avoiding injuries across the population. More generally, over-compensation incentivizes injurers to over-invest in precautionary measures or to cease engaging in socially beneficial activities all together.

This principle has led to another insight in the negligence framework. Some conduct that creates risk also creates benefits. Some examples include medical treatments and driving—in one sense, these create risks of harms yet there are benefits to these activities that help society in the aggregate. In other cases, the plaintiff may receive a mix of harms and benefits, such as where a pharmaceutical cure a disease but has an adverse side effect. Another example is a tort that stops the production of goods or provision of services. If consumers obtain those products and services elsewhere, third parties benefit from increased sales beyond those they would have

---

137 See id. at 592-593 (discussing R. Cooter, *The Cost of Coase*, 11 J. of Legal. Stud. 1, 8 (1982), stating “Cooter concurs with the view that the victims' ability to mitigate loss may open up a gap between private loss and externalized social cost, but he argues that this problem can be solved by proper legal rules of loss mitigation.”).
140 Id. at 1531.
141 Id.
made in the absence of a tort. This body of work acknowledges that, although a private individual may be harmed by a defendant’s conduct, there may be positive social benefits that derive from the defendant’s activity that offset the harm, whether fully or partially. Beyond this, it has been recognized that the negligence system induces actors to “engage too little in a socially desirable activity and forgo investments in efficient prevention” where the costs of those preventions is below the private benefit obtained from the increased activity. The benefits and utility of the defendant’s product are explicitly considered in products liability cases for design defect.

Some scholars have explored the concept that the social benefits of a defendant’s negligent conduct should be factored into the negligence liability system. In other words, “The tort system . . . can incorporate positive externalities into its internalization mechanism of liability rules only by offsetting them against negative externalities.” In particular, Israel Gilead and Michael Green suggest that proximate cause’s harm within the risk rule can be used to offset the social benefits of activities in particular circumstances. These include those where the benefits are external in nature, and do not accrue solely to the negligent actor. According to these authors, other conditions where reliance on this mechanism may be appropriate where the benefits outweigh (or at least equal) the harm caused by the negligent conduct. As one example, a plaintiff might undeniably suffer a private loss from an experimental surgery that fails to accomplish the desired result. However, this tortious activity might yield significant public benefits in the form of information that is learned so that the technique is improved and can then be used successfully on all subsequent patients. By using proximate cause, these authors argue, liability can be excluded for conduct that provides net social benefits particularly when the harm from engaging in such conduct is unforeseeable.

Using proximate cause as a policy can be used to ensure that beneficial activity is not deterred by a compensation system that fails to account for a good generated by risk-creating conduct. As will be outlined later in this Article, this principle has valuable implications for patent litigation as well.

146 Gilead and Green, *supra* note 139 at 592.
147 *Id.* at 1535.
148 *Id.* at 1545.
149 *Id.* at 1525.
V. Proximate Cause: Forms of Analysis

Negligence law has provided the largest and richest source of precedent in cases that range back as far as the mid-nineteenth century. There are multiple formulations and theories which have been formulated in the courts. Attempting to locate a unified formulation of the doctrine has been characterized as “fruitless.” Yet there are numerous themes that address “distinct problems, more or less unrelated.” As this phrase suggests, the policies implemented in the proximate cause requirement are varied. In some instances, the courts have used the requirement to prevent crushing liability on a defendant. Historically, the social utility of the defendant’s conduct has been considered as part of the liability calculus. Therefore, proximate cause determinations have been made to preserve certain institutions because of the public benefit that some institutions, including public utilities and other services, provide. It is widely recognized that the courts modify their construction of these terms to facilitate other broad policy goals, such as the imposition of liability to penalize reprehensible behavior and facilitate compensation by imposing liability on a large entity where the primary actor is insolvent.

Proximate cause is capable of eliminating all liability for a negligent defendant. For example, a negligence suit for personal injuries against a restaurant failed for lack of proximate cause because the injury was the result of an unforeseeable act—there, a truck that had been unexpectedly driven through the defendant restaurant’s wall. Alternatively, proximate cause may eliminate recovery for individual elements of a plaintiff’s damages. In other words, a plaintiff may bear the burden of demonstrating proximate cause for some damages but not for others. For example, an accident victim may seek damages for both injuries that occurred from an initial impact and those that developed months later. Depending on the facts presented and the governing standards, a jury may find recovery permissible for the first type but not the second. As a general rule, proximate cause does not cap recovery to a manageable level in all circumstances. For example, a defendant is responsible to pay for the harm caused by his or her actions, even if the extent of the damage caused could not have been foreseen when the injurious act occurred. As one source explains, “If I throw into the ocean a box belonging to A, which I have every reason to suppose empty, but there is hidden in it a purse of gold which is lost, I am liable for the loss.”

151 Dobbs, et al., supra note 114 at p. 273
152 Id. at p. 279.
153 Id.
154 Kelley, supra note 150 at 54.
156 See Richard L. Abel, A Critique of Torts, 37 UCLA L. REV. 785, 811–12 (1990) (“Courts often look for the ‘deep pocket’ defendant (large public or private entities or those likely to be insured) and then construct negligence and causation in order to rationalize the imposition of liability”).
As a practical matter, nearly all injuries are caused by multiple causes consisting of a mix of tortious and non-tortious causes. For liability to attach, it is sufficient that a defendant’s actions be a proximate cause of a plaintiff’s injury; there is no requirement that the defendant be the sole proximate cause.

A. Starting Points

The Supreme Court has acknowledged that proximate cause is “notoriously confusing” and “the lack of consensus on any one definition of ‘proximate cause’ is manifest.” As one opinion explains, “The proximate-cause inquiry is not easy to define, and over the years it has taken various forms; but courts have a great deal of experience applying it, and there is a wealth of precedent for them to draw upon in doing so.” Moreover, although the courts have attempted to infuse proximate cause analysis with the policies of the statute under examination, they appear to have difficulty breaking away from the centuries of precedent the formed the doctrine in the negligence realm.

Nonetheless, a doctrinal outline of proximate cause reveals some structure which varies somewhat among different jurisdictions. Some theories which operate as touchstones for the analysis include the concepts of the directness of the injury and the foreseeability of harm to the defendant at the time of the injurious conduct. Frequently, courts consider a mix of both. Although none offers a complete framework, these operate as starting points for the remainder of the analysis that can operate like a complex web of considerations. Attempting to crystalize this area into a single clear test is not possible, nonetheless a review of some of the primary points of decision yields a valuable foundation.

In some jurisdictions, the inquiry focuses on the directness between the defendant’s conduct and the plaintiff’s injury, which derives from In re Polemis, a British case that is standard in U.S. tort textbooks. Professor Kelley suggests that this standard derives from early tort decisions that considered the distinction between the now-defunct causes of action for trespass and trespass on the case (or “case”). Both were recognized means for a plaintiff to recover in a civil action, but were subject to rigid pleading rules, such that “direct or immediate

---

162 English v. Empire Dist. Elec. Co., 220 S.W.3d 849, 857 (Mo. Ct. App. 2007)(considering both foreseeability and directness); Shelton v. Bd. of Regents of Univ. of Nebraska, 211 Neb. 820, 825-827 (1982)(same); S. Ry. Co. v. Webb, 116 Ga. 152, 42 S.E. 395, 397 (1902)(“The injury must be the direct result of the misconduct charged; but it will not be considered too remote if, according to the usual experience of mankind, the result ought to have been apprehended.”); see generally Jessie Allen, The Persistence of Proximate Cause: How Legal Doctrine Thrives on Skepticism, 90 Den. U. L. Rev. 77, 88 (2012)(“Although the modern cases often talk more about foreseeability, they typically still mix that concept with the idea of unbroken continuity, as courts have been doing for a very long time”).
causation is required for trespass and indirect or consequential causation supports an action in case."\(^\text{165}\) As Professor Kelley explains the facts of an early decision, *Scott v. Shepard*:

Shepherd, a mischievous child, threw a lighted squib, filled with gunpowder, into a partially enclosed market-house. The squib fell on the stand of a seller in the market, who picked it up and threw it across the market-house. It fell on the stand of another seller, who picked it up and threw it, too. On this throw, the squib hit Scott in the face and exploded, putting out his eye.\(^\text{166}\)

To determine whether the plaintiff had properly pled a trespass action, the *Scott* court was required to consider whether Shepard’s actions directly injured Scott, or whether the intervening conduct of the sellers throwing the squib rendered the plaintiff’s harm as indirect and therefore not sufficient to give rise to liability. Those familiar with proximate cause analysis might recognize that the question of directness was translated into a starting point for some forms of proximate cause analysis.\(^\text{167}\) Similarly, Lord Sumner’s decision in the 1920 decision *Weld-Blundell v. Stephens* is emblematic of applying the former directness standard, explaining “the question on whether particular damages are recoverable depends only on the answer to the question whether they are direct consequences of the act.\(^\text{168}\) A common U.S. formulation of the directness standard examines whether there was “an unbroken connection between the wrongful act and the injury, a continuous operation.”\(^\text{169}\) Despite its early origin, the directness standard continues to have viability in a number of courts today in both negligence and statutory causes of action.

A simple case applying the directness standard is the Pennsylvania Supreme Court’s 1921 *Kelley v. Philadelphia, Baltimore & Washington Railroad Co.*\(^\text{170}\) There, the plaintiff was the driver of a car who was injured when driving by a warehouse which had a very large door that had swung out unexpectedly in front of her car. She sued the railroad, which was responsible for removing a removing a post that had been installed by the warehouse to prevent the door from swinging out into the street where the plaintiff had been driving when the accident occurred. Although the railroad was responsible for increasing the danger of the condition, the court found that it was not the proximate cause of the injury. Rather, the court held that the warehouse was legally responsible for the swinging door and was therefore the direct and proximate cause of the

---


\(^\text{166}\) Id. at 58.

\(^\text{167}\) Id. at 75 (According to Kelley, the *Scott* decision became a canon for proximate cause “on the nature of the causal sequence” and “no doubt seemed only natural to judges who looked back into the depths of the common law for substantive proximate cause precedents and, not knowing there were none, dredged up *Scott v. Shepherd*.”).


\(^\text{169}\) Milwaukee & St. P.R. Co. v. Kellogg, 94 U.S. 469, 475, 24 L. Ed. 256 (1876); See also Duphily v. Delaware Elec. Coop., Inc., 662 A.2d 821, 828-29 (Del. 1995) (“[A] proximate cause is one ‘which in natural and continuous sequence, unbroken by any efficient intervening cause, produces the injury and without which the result would not have occurred.’ The mere occurrence of an intervening cause does not automatically break the chain of causation. In order to break the causal chain, the intervening cause must also be a superseding cause, that is, the intervening act or event itself must have been neither anticipated nor reasonably foreseeable by the original tortfeasor.”).

A more complex application of this standard is *Philadelphia v. Beretta*, a 2002 case brought by Philadelphia against fourteen gunmakers for, among other things, negligence and negligent entrustment for the distribution and sale of firearms. This complaint sought the recovery of costs for preventing and responding to gun-related acts within the city’s borders based on the defendants’ marketing and distribution of guns that ultimately fell into the hands of criminals and children. The court applied a six-factor test to determine whether the harm to the city was directly injured by the defendant’s alleged negligence, as follows: (1) the causal connection between the alleged wrongdoing and the plaintiff’s harm; (2) any specific intent to harm the plaintiff; (3) the nature of the plaintiff’s injury within the larger purposes of tort law; (4) whether the damages is highly speculative; (5) the directness or indirectness of the alleged injury; and (6) the aim of keeping the scope of complex trials within judicially manageable limits, i.e., avoiding the risks of duplicative recoveries and the danger of complex apportionment.

Overall, the *Beretta* court affirmed the district court’s finding that the causal connection was too weak to sustain liability. Finding a lack of intent to harm the City, the analysis considered that the City’s harm was derivative of those who had been directly injured by the defendant’s guns. Further, the court considered the City’s harm speculative because the amount of lost resources was difficult to calculate. Holding that any gun’s route from the defendant to the user “long and tortuous,” the court determined that many of the injuries were directly caused by the gun owners. The *Beretta* court concluded that “any effort to compensate plaintiffs would require the expenditure of enormous judicial resources to determine which defendants should bear what percentage of liability.” Significantly, this decision illustrates that the concept of directness can incorporate nuanced policy analysis to consider whether the imposition of liability is consistent with the governing law. By expressly considering the larger purposes of the law, the speculative nature of the plaintiff’s injury and the defendant’s intent, the analysis goes beyond an impulsive label.

Other courts consider whether the plaintiff’s injury would have been foreseeable to the defendant at the time of the injurious conduct. The textbook example of this standard is *Overseas Tankship (U.K.) Ltd. v. Morts Dock & Engineering Co. (The Wagon Mound)*. Under current U.S. law, there is no uniform formulation. Some include the concept that “[a] party's act is the proximate cause of an injury if it is the natural and probable consequence of the act and should have been reasonably foreseen and anticipated in light of the circumstances.”

---

172 Id. at 423.
173 Id. at 425.
174 Fleming and Perry, supra note 11 at 491 (describing *Lynch v Knight* (1861), 9 H. L. C. 577).
Proximate cause is cause which in natural and continuous sequence, unbroken by an efficient intervening cause, produces the injury, and without which the injury would not have occurred, the injury being the natural and probable consequences of the wrongful act.  

Foreseeability is a term of art with different facets. In the negligence realm, foreseeability may use any number of policy levers, including: 1) whether the type of harm is foreseeable due to the nature of the risk created; 2) whether the manner of harm is foreseeable; and 3) whether the plaintiff is part of a foreseeable class of persons. In negligence cases, “[n]ot all courts use all three, but many commonly employ more than one of these limits.” Moreover, in the state courts these limitations may be implemented as part of proximate cause but others might treat some of these considerations as part of the duty inquiry.

One example of the application of the foreseeability standard is Mason v. Nashville. There, the school district was sued by a plaintiff who had been injured by a student who used a razor that had been provided to her as in a cosmetology course. The plaintiff alleged that the school district was negligent for giving these to students despite a zero-tolerance policy against the possession of a weapon. In this case, the Mason court found foreseeability lacking. Consistent with a number of state court decisions which evaluate foreseeability in the proximate cause context in a highly context-specific manner, the court found that the instructor had properly counseled the students in safety measures, there were no prior incidents of razor misuse, and the school had no reason to foresee that the particular student who had used the razor would be capable of being the aggressor in a fight.

As another example, in the Texas decision Ryder Integrated Logistics, Inc. v. Fayette County, the court considered whether a police department was liable for injuries to a truck driver who was killed as he drove past the scene of a traffic stop on the side of a highway. The plaintiff’s injury occurred when he was distracted by the police vehicle, which was in the process of turning around so that its headlights and flashing lights faced oncoming traffic. As with Mason, the Ryder analysis was based on a highly fact-specific analysis.

As a tort that considers foreseeability as determinative of liability in other elements, the introduction of foreseeability into the proximate cause inquiry focuses its application to the purposes furthered by the negligence cause of action. In this way, foreseeability enables the

---

179 id.
180 id.
182 id. at 226-227.
184 id. at 928 (finding that “a reasonable peace officer could have foreseen that driving westbound near an eastbound shoulder at night—with headlights and emergency lights illuminated—might confuse drivers, disrupt traffic, and lead to a collision much like the one that ultimately occurred.”).
inquiry of whether the imposition of liability is consonant with the purposes of this particular cause of action. To a significant degree, negligence law is concerned with the liability for risk taken in light of the foreseeable harms.\textsuperscript{185} This principle is evident in negligence duty formulations which consider foreseeability as a question of law for assessing “whether the category of negligent conduct at issue is sufficiently likely to result in the kind of harm experienced that liability may appropriately be imposed on the negligent party.”\textsuperscript{186} Foreseeability is integrated in the standard of care, which inquires whether the defendant’s burden of avoiding an accident would have been reasonable in light of the foreseeable risks to the plaintiff.\textsuperscript{187} As part of the overall negligence inquiry, the proximate cause standard has been defined through the use of such terms as “reasonably foreseen” and “natural” and “probable.” Theorists have considered various reasons that courts apply the foreseeability doctrine despite the fact that this consideration already exists in other elements of the tort. Professor Grady considers that this standard eliminates liability for accidents that are merely fortuitous occurrences.\textsuperscript{188} Professor Kelley posits that the concept of foreseeability may seem to be “an accurate and illuminating way of expressing delicate judgments about the mutual expectations in social systems of reliance that underlie some of our basic judgments about wrongdoing in any particular social context.”\textsuperscript{189} Yet as the proximate cause as a legal requirement evolved, these explanations may be considered incomplete. Moreover, as a malleable concept, foreseeability has been criticized “with its triple role and its accordion-like meaning” as “clearly one of the murky concepts that has led students and scholars to think that negligence law lacks conceptual integrity.”\textsuperscript{190}

An example of a mixed standard is *Pepsi-Cola Bottling Co. v. Von Brady.*\textsuperscript{191} The court’s application demonstrates that policy rules are woven into the analysis to determine whether a particular defendant’s action is the proximate cause of an injury. There, the plaintiff was a driver traveling in the right lane who was cut off by the sudden turn of a third party, who swerved into her from the left lane. To avoid a collision, the plaintiff turned right and hit the defendant Pepsi’s truck which was parked illegally on the side of the road by its employee. She sued the Pepsi as the owner and employer, claiming that but for the illegally parked truck she would have avoided the injury because there would have been sufficient space for her car to navigate. Clearly, Pepsi’s actions were in breach of local parking regulations. The court recognized that the location of the Pepsi truck was the most direct source of the impact, questioning whether “the parking of the truck created a condition prior to the accident cannot be disputed, but was it such a ‘condition’ that liability can be predicated thereon?”\textsuperscript{192} Rejecting this, the court instead held that the third party was the sole and proximate cause of the injury. The court explained that the Pepsi truck had

---


\textsuperscript{186} Ballard v. Uribe, 41 Cal. 3d 564, 573 n. 6 (1986).

\textsuperscript{187} See authorities cited *supra* at notes 181-184.

\textsuperscript{188} Mark F. Grady, *Proximate Cause Decoded*, 50 UCLA L. REV. 293, 300 (2002).

\textsuperscript{189} Kelley, *supra* note 165 at 77.


\textsuperscript{191} Pepsi-Cola Bottling Co. of Tulsa, Okl. v. Von Brady, 386 P.2d 993 (1963).

\textsuperscript{192} *Id.* At 996; see generally, F. WHARTON, A TREATISE ON THE LAW OF NEGLIGENCE 73-84 (2\textsuperscript{nd} ed. 1878) (distinguishing conditions and causes).
“merely furnish[e]d a condition by which the injury was made possible.”

Additionally, the opinion held that the third party’s action could not have been foreseen by Pepsi’s employee at the time that the truck was illegally parked. Although the Pepsi-Cola court might have held that the truck was a direct (and therefore proximate) cause of the plaintiff’s injury given that it provided source of the point of impact, it chose instead to predicate liability solely on the third party’s more active negligence.

In the abstract, the choice between these starting points may not be critical. Each formulation is sufficiently malleable to accommodate a range of policy choices. To trace any of these standards reveals shifts that function as a marker of the larger societal changes and attitudes about recovery, liability and risk. Nonetheless, one must be cognizant of how these terms are used in our current legal and economic culture. Further, as these standards become crystallized into precedents, each begin to reflect outcome-determinative meanings that reflect policy choices. As some examples, certain states have created default rules that events that include a plaintiff’s suicide, excessive drinking or product misuse sufficient to constitute an intervening, superseding cause that terminates the defendant’s liability. These decisions implement policies that concern compensation, responsibility, and risk.

More broadly, proximate cause shapes liability in negligence cases in a manner that is consistent—and conforms—to the purpose of the cause of action. Rather than considering concepts such as directness and foreseeability as immutable, it is critical that the context for their construction is considered. The functions that the requirements serve is translatable to patent law. However, the constructions of those terms, developed in the negligence context, might be less amenable to direct importation into the intellectual property sphere.

B. Harms within the Scope of the Risk

Negligence might be defined as a protection against the unreasonable risk of harms created by others. In the proximate cause analysis, the plaintiff’s injury must derive from the type of harm that renders the defendant’s conduct tortious. Called the “harm within the risk” rule, this is intended to ensure that “liability; whatever risks, in other words, made one negligent should also be the only risks for which one should have to pay if they have materialized in actual

193 Id. at 996.
194 John G. Fleming, The Passing of Polemis, 39 The Canadian Bar Review 489, 492 (1961) (observing with respect to one such change “was to be attributable to a deep shift in social philosophy, associated with the modern welfare state, in which judges have on the whole collaborated with juries to ensure a much wider range of loss distribution than would have been acceptable forty years ago”).
195 White v. Lawrence, 975 S.W.2d 525, 530 (Tenn. 1998) (recognizing that “suicide may constitute an intervening cause if it is a willful, calculated, and deliberate act of one who has the power of choice.”) (citing cases); Marshall v. Clark Equip. Co., 680 N.E.2d 1102, 1108 (Ind. Ct. App.), decision clarified on denial of rehe’g, 683 N.E.2d 1351 (Ind. Ct. App. 1997) (“An intervening cause, such as a misuse of the product, will relieve the manufacturer of liability when the intervening acts could not have been reasonably foreseen by the manufacturer.”); Sapko v. State, 305 Conn. 360, 388 (2012).
harm.” Derived from cases written in the 1850’s, this principle limits the plaintiff’s compensation to the injuries that are the same *types of harm* that render the defendant’s acts tortious.\(^{198}\)

Professor Goldberg has a set of useful hypotheticals to explain how this requirement operates. As he describes, imagine a driver who carelessly throws a half-filled coffee cup out of his window.\(^{199}\) In the first, the coffee splashes on the windshield of an approaching car. As it blocks the vision of the driver of that car, he swerves, strikes, and seriously injures a pedestrian standing on the far sidewalk. In this case, the plaintiff’s injury is arguably within the scope of the risk of the defendant’s activity. Under a proper analysis, the defendant is not required to foresee the precise manner in which the plaintiff was harmed but only that a passerby might be physically harmed by this careless act. Alternatively, consider that the cup falls harmlessly on the street.\(^{200}\) Infuriated by this thoughtlessness, the plaintiff raises his arm in anger, which inadvertently knocks into a hidden beehive located on a nearby tree limb. According to Goldberg, the plaintiff’s injury from the bee stings, although a *cause in fact* of the defendant’s action, is not *proximately* caused by it. As he explains, “the risk of a pedestrian's being injured by a bee sting was not one of the risks that inclines us to label throwing litter out of a moving car to be a careless activity.”\(^{201}\) In other words, bee stings are not within the scope of the risk of littering from a moving car.\(^{202}\)

To properly apply this standard, one must first identify the nature of the tortious conduct and the scope of the relevant risk. Unquestionably, there are assumptions and judgments made during this definitional act. A broader definition of the activity will result in more instances where liability will be imposed—for example, “driving” has a broader range of available risks compared with the narrower activity of “throwing a cup out of a car.” Next, one considers the relevant risks of the activity and determine whether the injury has a sufficient relation to at least one of those risks. For example, it has been said that the risk of handing a loaded gun to a child is that the risk is of injury from the child firing the gun without understanding its dangers and


\(^{198}\) *Id.* at 340.


\(^{200}\) *Id.*

\(^{201}\) *Id.* at 1338.

\(^{202}\) See also DAN B. DOBBS, PAUL T. HAYDEN AND ELLEN M. BUBLICK, *The Law of Torts* § 198 (2d ed. [date]), which provides the following example:

…[S]uppose that the defendant negligently manufactures a vacuum cleaner so that it does not have good suction. After several frustrating days using the cleaner, the purchaser takes it to the repair shop. On her way to the shop, the purchaser is struck by a car and suffers injury. The manufacturer was negligent. The manufacturer's negligence was one of the many factual causes leading to the purchaser's harm—but for the manufacturer's negligence, the purchaser would have stayed home safely cleaning carpets and would not have been struck by a car. Yet here again, legal professionals are likely to agree that the manufacturer's conduct in making the poor vacuum was not a proximate cause of the plaintiff's injury, because the defendant's negligence did not create or increase the risk of injury in a vehicular collision.
consequences. Therefore, if the child drops the gun and injures her foot, that harm is outside the scope of the risk created through the negligent entrustment. This would trigger proximate cause to terminate liability based on the manner in which the accident occurred. That is, rules against giving children guns are not enforced by compensation for an injury that would have occurred if the defendant had handed the child a non-dangerous, but heavy, object such as an iron pot. In other words, the harm within the risk rule is intended to ensure the necessary correlation between the defendant’s tort and the plaintiff’s harm.203

As one source explains, this limitation ensures fairness from a corrective justice standpoint that “the defendant is not held legally responsible for the materialization of a harm that is not within the set of possibilities that supply a reason for exercising due care.”204 This aspect of the proximate cause requirement provides a vital link between a tortious event and the plaintiff’s harm. If the law permitted plaintiff’s recovery for innocent conduct: 1) socially valuable non-tortious activity would be discouraged; 2) the plaintiff would be overcompensated; and 3) negligence would operate in a similar manner to strict liability, in the sense that liability would be imposed on a defendant without a showing of fault, but without the limitations required to bring a strict liability claim.205

Inherent in this rule is the concept that, absent the meaningful application of proximate cause, the other elements of a negligence claims are over inclusive.206 In other words, following the letter of the other elements of the torts results in the imposition of liability for reasonable risks.207 The tort duty rules coupled with the standard breach analysis can deliver results that, particularly when coupled with the expansive cause in fact analysis, delivers results that are inconsistent with the purposes of the law. Significantly, proximate cause cuts off liability despite the fact that a defendant breached a duty to an injured plaintiff. As one source explains, proximate cause “serves as the final and finest sieve” that considers the specific context in which the alleged harm occurred.208 The harm within the risk rule provides a vital link that excludes injuries that result from either socially beneficial risks or those that are merely fortuitous from the liability calculation.

C. Intervening, Superseding and Efficient Causes

Proximate cause doctrine relies on the mechanism of the intervening, superseding cause to cut off liability in particular cases.209 In some jurisdictions, a parallel concept is used to cut off

---

203 Heidi M. Hurd and Michael S. Moore, *Negligence in the Air*, 3 Theoretical Inquiries L. 333, 334 (2002). If, however, the tortious action is defined as handing a small child a very heavy object—in this case, it happens to be a loaded gun—then liability can be imposed.


205 See *id.* at 1521-1522.

206 Steven Shavell, *An Analysis of Causation and the Scope of Liability in the Law of Torts*, 9 THE JOURNAL OF LEGAL STUDIES, 463, 489 (1980) (“In reality a significant element of strict liability is inherent in the negligence rule: injurers who do in fact act with due care or attempt to do so will sometimes be found negligent and have to pay damages.”).


209 Milwaukee & St. P.R. Co. v. Kellogg, 94 U.S. 469, 475, 24 L. Ed. 256 (1876); See also Duphily v. Delaware Elec. Coop., Inc., 662 A.2d 821, 828-29 (Del. 1995) (“In order to break the causal chain, the intervening cause must
liability at its efficient cause (sometimes called the operating efficient cause or efficient intervening cause), which “is usually termed the cause, whose share in the matter is most conspicuous, and is immediately preceding and proximate to the event.”\textsuperscript{210} These concepts have been described as follows:

The issue of an intervening cause occurs when a new and independent force ... so interrupts the chain of events that it becomes the responsible, direct, proximate, and immediate cause of the injury.\textsuperscript{211}

Such acts typically occur between the defendant’s initial injurious act and the time that the plaintiff is injured. In negligence cases, “[i]f the intervening force actually does intervene — producing harm by active operation after the negligent commission or omission of the actor sought to be held liable, then, usually, that actor’s conduct is exonerated provided that the intervening force was not foreseeable.”\textsuperscript{212} Such acts are third parties or independent forces such as an unforeseen storm, physical aspects of a landscape or road, or other significant contributing factors to the plaintiff’s injury that are more proximate in time or space.

There are significant exceptions to the intervening, superseding cause rule. One is that a second tortfeasor engaged in a risk that is objectively foreseeable to the defendant will not break the causal chain. As with foreseeability in other contexts, the term is a legal conclusion that varies based on both the facts and policy judgment. One example is the Texas Supreme Court’s \textit{Biggers v. Continental Bus System, Inc.}\textsuperscript{213} In \textit{Biggers}, a decedent whose car had been hit head-on by one of Continental’s buses that was being driven at an excessive speed. At the time of the accident, the decedent’s car was driving in the wrong lane toward the bus. Continental argued that proximate cause was lacking because “one operating an automobile on a highway is not required, under any circumstances, to foresee or anticipate that an automobile traveling in the opposite direction may enter the wrong traffic lane.”\textsuperscript{214} Recognizing that other jurisdictions were in conflict, the Texas Supreme Court sided with the line of cases holding that such events were foreseeable.\textsuperscript{215} The court explained “[i]t is a matter of common knowledge that a fair proportion of collisions on highways are ‘head-on’ collisions, resulting from one of the automobiles being in the wrong lane of traffic.”\textsuperscript{216} On the basis, the \textit{Biggers} court affirmed a jury finding that the bus’s excessive speed was the proximate cause of the collision.

\textsuperscript{210} See, e.g., Powers v. Standard Oil Co., 98 N.J.L. 730, 734, 119 A. 273, 274 (Sup. Ct.), aff'd, 98 N.J.L. 893, 121 A. 926 (1923); Tapp v. Blackmore Ranch, Inc., 254 Neb. 40, 50, 575 N.W.2d 341, 349 (1998); Smith v. Secrist, 590 S.W.2d 386, 389 (Mo. Ct. App. 1979). This article will refer to these standards collectively as the intervening, superseding cause.


\textsuperscript{212} 3 American Law of Torts § 11:9.


\textsuperscript{214} \textit{Id.} at 357.

\textsuperscript{215} \textit{Id.} at 357-358.

\textsuperscript{216} \textit{Id.} at 358.
**Biggers’** holding underscores that foreseeability operates as a policy judgment, in part because of the unusual way that the plaintiff was injured. The decedent’s car began driving in the wrong lane because it was part of a line of cars that was forced to suddenly slow down. This was because the lead car hit its brakes and slowed from driving 40-45 miles per hour down to 10-15 “to observe the creek to ascertain if it was too muddy for fishing.”\(^{217}\) The decedent, driving behind this lead car, slowed to avoid an accident. However, the car behind the decedent was unable to brake quickly enough and continued forward, pushing the decedent’s car into oncoming traffic right in front of the speeding bus. The defendant bus company argued “that its driver could not reasonably foresee the unusual circumstance” that led the decedent’s car to arrive in the wrong lane.\(^ {218}\) In response, the Biggers court over-ruled precedent that had held “one’s negligence cannot be a proximate cause of an injury unless the precise manner or means of the injury resulting from the negligence can be foreseen.”\(^ {219}\) Instead, the Biggers court applied the rule that proximate cause was established so long as a hypothetical reasonable person driving the bus should have foreseen the risk of excessive speed to any car that entered the wrong lane for any reason. In this way, the court considered both the nature of the risk (excessive speed) coupled with the foreseeability of the plaintiff (any car in the wrong lane) to further a policy that favors compensation to injured plaintiffs. Additionally, by steering the analysis toward the generalized nature of the risk of excessive speed and away from the foreseeability of the particularized manner in which the accident occurred, the Biggers court relied on all three policy guides toward recovery for the decedent.

More broadly, these cases illustrate that in close cases the concept of foreseeability, which tracks other policies woven throughout negligence law, can operate within proximate cause to separately allow or withhold damages to further the law’s goals. These turning points—whether under the direct or foreseeability standard—can be similarly implemented in the patent system to guide liability as a matter of policy. Over time, these precedents can be crystallized to provide certainty and act as an incentive to structure transactions appropriately.

**D. Modifying Proximate Cause for Particular Causes of Action**

Although the roots of proximate cause developed primarily in negligence law, the requirement applies to other common law torts. In such cases, the proximate cause is fine-tuned to fit contextually with the relevant cause of action. For example, some sources hold that those who act willfully for the purpose of causing harm cannot rely on the argument that foreseeability is lacking and that, therefore, recovery should be denied for lack of proximate cause.\(^ {220}\) In such cases, the inquiry collapses to the cause-in-fact question and a defendant cannot rely on the argument that the harm was unforeseeable to terminate liability. For example, those engaged in an intentional tort will be liable for all harm caused, including harms that are not reasonably foreseeable.\(^ {221}\) This is because the law seeks to powerfully discourage intentionally wrongful

---

\(^ {217}\) *Continental Bus*, 157 Tex. at 354.

\(^ {218}\) *Id.* at 358.

\(^ {219}\) *Id.* at 359.

\(^ {220}\) See Restatement (Second) of Torts § 435A; Kimberlin v. DeLong, 637 N.E.2d 121, 127 (1994) (“The rule has therefore emerged that the doctrine of superseding cause is inapplicable to willful torts”).

\(^ {221}\) Restatement (Third) of Torts § 33(a) (“An actor who intentionally causes harm is subject to liability for that harm even if it was unlikely to occur.”).
conduct. In effect, the rule requires the defendant to become liable for consequences that are entirely unexpected, even if the defendant would not have responsible for those harms if he had been merely negligent.

In contrast, a California court narrowed the scope of proximate cause in a case asserting negligence against the state, rendering it more difficult for the plaintiff to recover where governmental decision-making was challenged. There, the plaintiff alleged that a governmental actor had failed to order a sufficient psychological evaluative test for an inmate, Pitre, who was soon to be released from prison. According to the plaintiff, an adequate evaluation, as required by a state statute, would have led to Pitre’s psychiatric commitment. Instead, she was murdered by Pitre soon after his release. The relevant statute directed, among other things, that government’s immunity rested on proximate cause between the government employee’s conduct and the harm. The court recognized that its conclusion about proximate cause was based on public policy.

After reviewing a line of California governmental immunity decisions which found proximate cause lacking in cases brought against governmental decision makers, this Court directed that the plaintiffs’ complaint be dismissed, finding proximate cause lacking as a matter of law. Given that evaluators had discretion in the review process, the court found that “[a]s a matter of policy” the proximate cause inquiry was “problematic, because it trenches closely upon the discretionary functions of the evaluation process” established by statute. In short, to protect the discretion of governmental decision makers, the court held that proximate cause was not met. In such instances, proximate cause’s black letter standard is varied based on the policy reasons that derive from the nature of the claim.

As in tort law, proximate cause can be varied to apply to different instances of infringement in the patent context. Thus, proximate cause in cases of willful infringement might allow the capture of liability compared to cases where the infringer had no notice of the patent. Similarly, proximate cause in cases of attenuated liability, such as where a manufacturer provides a very minor infringing component of a multicomponent device, proximate cause might be applied more stringently.

E. The Appearance of Redundancy

One pervasive characteristic of proximate cause analysis is the doctrine’s tendency to apply standards that are duplicative of other elements of the substantive cause of action. For example, in negligence law the proximate cause relies heavily on foreseeability, which is central
to aspects of negligence law’s duty and breach analysis. In other words, the proximate causal chain encompasses foreseeable types of harm, and the determination over whether an event is an intervening cause rests on whether it is foreseeable. At the same time, one way to establish a legal duty toward another is based on whether or not the plaintiff should have been a foreseeable victim that might be harmed by the defendant’s risk-creating activity. Foreseeability governs the scope of the defendant’s duty, in that traditional breach analyses encompass limitations for the objectively foreseeable likelihood of harms to others.

Professor Greene argues that proximate cause’s “chameleon quality permits it to be substituted for any one of the elements of a negligence case when decision on that element becomes difficult.” Although there may be some truth in his argument, the analysis of redundant elements evidences that courts apply foreseeability differently within proximate cause compared to others. This allows courts to fine-tune liability where the other elements of the claim do not lead to just results that are commensurate with the law’s purpose. In the negligence context, this occurs where there is a categorical duty that has been crystallized through precedent. The legal system tends to create such precedent to provide some level of certainty for actors engaged in risk-creating activity. That is, prospective defendants should know where their legal liability begins and ends so that they can shape their activities to minimize the potential for payment. The negligence system benefits from this stability, because it allows for the creation of legal rules that can incentivize risk creators to aim for an optimal level of risk. This balance is important, because a burdensome negligence standard incentivizes risk creators to undertake wasteful spending on protective measures or curtailing socially valuable activity entirely. Yet, as applied in some cases, negligence law’s categorical rules can deliver harsh results in the absence of doctrines to limit liability. In other words, a negligence system does not impose liability for all risks, but relies on doctrines that include proximate cause to calibrate liability to guide defendants to engage in appropriate levels of risk.

For example, in Pine v. Arruda, a woman was injured at a golf course when she was hit by a member who was driving a golf cart, which suddenly accelerated when the plaintiff’s golf bag fell on the accelerator. The accident occurred during a sudden rain and lightning storm that arose during tournament, when a member tried to pull the cart out of the rain into a covered area so that the plaintiff’s bag would not get wet. The plaintiff sued the golf course alleging negligence for its failure to exercise due care by failing to have more covered storage areas and assistance with golf carts during rain delays. Applying the categorical duty rule for business invitees, the court found the defendant golf course owed the plaintiff a legal duty “to maintain its property in a reasonably safe condition in view of all the circumstances, including the likelihood of injury to others, the seriousness of the injury, and the burden of avoiding the risk.”

Accepting the defendant’s argument that liability under these circumstances would impose “liability for all injuries resulting from the negligence use of their golf carts no matter who was using them or how attenuated the link was between the conduct of the country club and the

---

226 See e.g., Lopez v. McDonald's Corp., 193 Cal. App. 3d 495, 508 (Ct. App. 1987) (“The defendant owes a duty of care to all persons who are foreseeably endangered by his conduct, with respect to all risks which make the conduct unreasonably dangerous.”).
229 Id. at 285.
resulting injury,” the court found that the plaintiff’s claim should be rejected for the lack of proximate cause. Among other things, the court consider the golf cart’s driver as an unforeseeable intervening cause that terminated any potential liability. The Pine court’s entry of judgment in favor of the defendant demonstrates that, more broadly, proximate cause allows for the courts to craft liability rules that allow risk-creating businesses to engage in appropriate levels of risk. Further, proximate cause prevents defendants from incurring liability for injuries over which they have little to no control.

Certainly, a court might have concluded differently than the Pine court did. This is the type of policy judgment that allows courts to set an appropriate level of risk as a matter of policy. A court might determine, for example, that the defendant country club needed to incentives to provide more protective measures than were taken here. The existence of proximate cause gives courts the flexibility and policy space to make that judgment. In the patent context, the existence of a cause-in-fact inquiry in the lost profits inquiry therefore does not preclude the applicability of proximate cause in that area. Nor does it preclude consideration of both cause-in-fact and proximate cause for reasonably royalty relief.

VI. Proximate Cause and Statutory Causes of Action

A. Overview

The proximate cause limitation of negligence law cannot be imported wholesale into all statutory causes of action. Given the disorderly state of the common law across the fifty states, the Supreme Court has recognized the difficulty of stating a precise proximate cause test for federal statutory claims. Many of the leading cases integrate a patchwork of considerations. Each bears the same judicial stamp that considers the appropriate limits on liability. Yet another consistent theme among these authorities is that the policies of proximate cause are malleable to fit the policies inherent in the relevant statute.

In some cases, the Court has applied a modified scope of the risk analysis. For example, in an antitrust action, Blue Shield of Virginia v. McCready, the Supreme Court set forth the relevant test as a mix of the common law standard coupled with limitations that are relevant to the purposes of the antitrust statute. As a further example, in the Supreme Court’s Associated General Contractors of California v. California State Council of Carpenters, the Court’s analysis explicitly excluded consideration any harms caused by conduct that were not redressable

---

230 Id. at 284-285.
232 Blue Shield of Virginia v. McCready, 457 U.S. 465, 478, 102 S. Ct. 2540, 2547–48 (1982) (describing the relevant test as comprising two-factors: (1) to the physical and economic nexus between the alleged violation and the harm to the plaintiff, and (2), more particularly, to the relationship of the injury alleged with those forms of injury about which Congress was likely to have been concerned in making defendant's conduct unlawful and in providing a private remedy under § 4).
by the antitrust law as irrelevant.\textsuperscript{233} This finding evolves the common law harm within the risk rule to a “harm within the statute’s purpose” rule to fine tune the imposition of liability according to policy relevant to the cause of action at issue. In \textit{Lexmark Int’l, Inc. v. Static Control Components, Inc.}, the Court instructed that that the “[p]roximate-cause analysis is controlled by the nature of the statutory cause of action,” and that an important part of the proximate cause inquiry turns on whether “the harm alleged has a sufficiently close connection to the conduct the statute prohibits.”\textsuperscript{234} These decisions underscore that proximate cause must be modified to further the purposes of the statute that is asserted to create the defendant’s liability. For patent law, the relevant inquiry must therefore be viewed in light of the goals of the system which gives rise to new invention and innovation.

Certainly, the leading cases use parts of the structure of proximate cause as that doctrine developed in negligence law. For example, the directness of the plaintiff’s harm is a key consideration for proximate cause in many statutory causes of action; although the foreseeability inquiry controls in others.\textsuperscript{235} In implementing these starting points, those terms remain as malleable as it is in common law tort cases. For example, in \textit{Lexmark Int’l v. Static Control Components}, the Supreme Court considered proximate cause in the context of a federal statutory cause of action for false advertising. There, the Court applied the directness standard.\textsuperscript{236} In doing so, the Court wove in some of the considerations that are similar to those discussed in the \textit{Beretta} case.\textsuperscript{237} The \textit{Lexmark} opinion considered the directness of the plaintiff’s asserted injury, the policy of keeping the scope of trials within judicially manageable limits, and avoiding the risks of duplicative recoveries and complex apportionment.\textsuperscript{238}

In other cases, the Court has considered a number of policy considerations relevant to the asserted cause of action that are akin to those considered in \textit{Beretta}.\textsuperscript{239} These include whether there are victims who are more directly impacted and therefore have sufficient incentives to vindicate the same wrongful conduct.\textsuperscript{240} Also, recognizing that defendants have limited resources, courts examine whether awarding damages to those who are indirectly harmed might

\textsuperscript{233} Associated General Contractors of California, Inc. v. California State Council of Carpenters, 459 U.S. 519, 526-527 (1983) (“ Such deceptive diversion of business to the nonunion portion of a so-called ‘double-breasted’ operation might constitute a breach of contract, an unfair labor practice, or perhaps even a common-law fraud or deceit, but in the context of the bargaining relationship between the parties to this litigation, such activities are plainly not subject to review under the federal antitrust laws.”) Although the language of the \textit{Associated General Contractors} decision does not appear to expressly create a proximate cause requirement for Clayton Act claims, the opinion has been read to do so. See Holmes v. Securities Investor Protection Corp., 503 U.S. 258, 268 (1992) (construing \textit{Associated General Contractors}, stating that in that case “we held that a plaintiff’s right to sue under § 4 [of the Clayton Act] required a showing that the defendant’s violation not only was a ‘but for’ cause of his injury, but was the proximate cause as well.”).


\textsuperscript{236} S. Pac. Co. v. Darnell-Taenzer Lumber Co., 245 U.S. 531, 533, 38 S. Ct. 186 (1918).

\textsuperscript{237} See supra note 170 and accompanying text. See also S. Pac. Co. v. Darnell-Taenzer Lumber Co., 245 U.S. 531, 533, 38 S. Ct. 186 (1918) (Holmes, J) (recognizing “the general tendency of the law, in regard to damages at least, is not to go beyond the first step.”).

\textsuperscript{238} \textit{Lexmark}, supra note 234 at 543-544.

\textsuperscript{239} See note 170 and accompanying text.

\textsuperscript{240} \textit{Anza v. Ideal Steel Supply Corp.}, 547 U.S. 451, 458 (2006).
preclude meaningful recovery others who are more directly injured. Courts seek to avoid allowing multiple recoveries and fashioning complicated rules apportioning damages. This rule has been supported by the concern that multiple claims to recovery that would deplete the defendant’s resources and therefore render it difficult for others to obtain meaningful financial recovery.

Another issue is the potential difficulty ascertaining the amount of the plaintiff’s actual damages compared with other factors that may have led to the plaintiff’s losses. This consideration recognizes that plaintiff’s harms might be caused from multiple sources, including those that do not derive from the defendant’s wrongful conduct. As one opinion explains, “the more indirect the injury, the more difficult it becomes to ascertain the amount of a plaintiff’s damages attributable to [the defendant's wrongdoing], as distinct from other, independent, factors.” Some dicta suggests that the mere potential complexity of apportioning harms is not determinative. Nonetheless, consistent with the common law, where there are a multiplicity of factors that contribute to the plaintiff’s injury, which leads to a conclusion that the defendant’s conduct is not a substantial factor to the plaintiff’s harm, recovery is denied.

Just as in negligence cases, some of these considerations are redundant of other elements of the statutory cause of action. Proximate cause applied in the antitrust context is one example, where both consider the nature of the plaintiff’s injury as central to the resolution of both standing and proximate cause. Just as in negligence law, this injury is analyzed differently for the statutory tort. For example, for purposes of standing in antitrust cases, the plaintiff’s harm is considered to see whether it is within the zone of interests protected by the antitrust law. In contrast, for proximate cause in antitrust cases the question of the plaintiff’s injury “is not one of standing, but of directness.”

Similarly, in Lexmark v. Static Control Components, a false advertising case, the Court considered both the type of injury suffered by the plaintiff for standing purposes and then its directness to determine proximate cause. Although those conversant with patent analysis are accustomed to compartmentalizing issues within particular elements, the trend in proximate law

241 Associated General Contractors, 459 U.S. at 545.
242 See Associated General Contractors, 459 U.S. at 544.
244 Lexmark Intern., Inc. v. Static Control Components, Inc., 134 S.Ct. 1377, 1392 (2014) (stating that the “potential difficulty in ascertaining and apportioning damages is not . . . an independent basis for denying standing where it is adequately alleged that a defendant's conduct has proximately injured an interest of the plaintiff's that the statute protects.”).
245 See RESTATEMENT (SECOND) OF TORTS § 281 (1965)(“Where the harm which in fact results is caused by the intervention of factors or forces which form no part of the recognizable risk involved in the actor's conduct, the actor is ordinarily not liable.”); Perron on behalf of Jackson v. J.P. Morgan Chase Bank, N.A., 845 F.3d 852, 858 (7th Cir. 2017)(plaintiffs' attempt to recover for dissolution of their marriage due to the defendant’s violation of the federal Real Estate Settlement Procedures Act (RESPA) was rejected for lack of proximate cause because “This kind of claimed harm is far too attenuated from the alleged violation to cross the proximate-cause threshold.”).
246 See South Dakota v. Kansas City S. Indus., Inc., 880 F.2d 40, 46 (8th Cir. 1989) (“as the Supreme Court has observed, the determination of whether there is antitrust standing is similar to the determination of whether there is proximate cause.”).
248 Lexmark, 134 S.Ct. at 1392.
analysis expressly allows for redundancy that considers different aspects of policy in different claim elements.

B. Examples: Application of Proximate Cause to Statutory Torts

A review of a number of Supreme Court cases demonstrates the manner in which the court applies the doctrine of proximate cause to federal statutory claims. As early example is Associated General Contractors of California v. California State Council of Carpenters (“Associated General Contractors”), the Supreme Court examined whether a union could bring a claim for treble damages § 4 of the Clayton Act, which authorizes recovery by “any person who shall be injured in his business or property by reason of anything forbidden in the antitrust laws.” 249 The union alleged $25 million in damages based on the defendant efforts to persuade employers to hire non-union workers. 250 As a result, the complaint alleged that contractors and subcontractors who employed union workers lost contracts and the defendant’s conduct interfered with employers’ ability to make “free choices between market alternatives [and] is inherently destructive of competitive conditions.” 251 Although the Court concluded that the law’s statutory history did not discuss proximate cause, the Associated General Contractors Court read the requirement into the statute. Ultimately, the Court determined that the plaintiff union could not bring its claims under the Clayton Act because its allegations did not establish that its injury had been proximately caused by the defendant’s alleged conduct. 252

To decide the question, the Associated General Contractors Court relied on both the type of harm and foreseeable plaintiff inquiries. Specifically, the Court held that these damages were not the type of harm that Congress sought to protect by enacting the antitrust laws. Further, the Court found that the union was not within the class of plaintiffs that Congress sought to protect with these laws, evoking the common law’s “foreseeable plaintiff” concept. 253 In addition, the Court considered the directness of the union’s injuries, evoking Polemis’ legal rule. Yet unlike Polemis which used directness to expand liability, the Associated General Contractors Court read the word “direct” narrowly to hold that the union could not establish proximate cause under the facts presented.

---

250 Id. at 528. The relevant allegations redressable by the Clayton Act are described by the Court are “The complaint alleges that defendants “coerced” two classes of persons: (1) landowners and others who let construction contracts, i.e., the defendants' customers and potential customers; and (2) general contractors, i.e., defendants' competitors and defendants themselves. Coercion against the members of both classes was designed to induce them to give some of their business—but not necessarily all of it—to nonunion firms.” Id. at 527-528.
251 Lexmark, 134 U.S. at 1386 (stating that, in Associated General Contractors, the Court “held that the statute limited the class to plaintiffs whose injuries were proximately caused by a defendant's antitrust violations.”).
252 Associated General Contractors, 459 U.S. at 540. In so doing, the Court drew on antitrust’s standing jurisprudence citing Brunswick Corp. v. Pueblo Bowl-O-Mat, Inc., 429 U.S. 477, 487-489 (1977) (explaining “In each case its alleged injury must be analyzed to determine whether it is of the type that the antitrust statute was intended to forestall. In this case, particularly in light of the longstanding collective bargaining relationship between the parties, the Union's labor-market interests seem to predominate, and the Brunswick test is not satisfied.”).
253 Id. at 540 ("Set against this background, a union, in its capacity as bargaining representative, will frequently not be part of the class the Sherman Act was designed to protect").
The Associated General Contractors Court relied on several policy considerations, including that the harm from the alleged conduct was diffused among others with more direct injuries. In addition to the fact that these plaintiffs had incentives to vindicate their own claims, the Court found that the union’s injuries were, by comparison, “speculative.” The Court observed that the union’s alleged harm might have been caused by other factors unrelated to the defendant’s conduct, and that its claim raised the “risk of duplicate recoveries on the one hand, or the danger of complex apportionment of damages on the other.” Additionally, the Court expressed concern that allowing the union’s complaint would mean that “the District Court would face problems of identifying damages and apportioning them among directly victimized contractors and subcontractors and indirectly affected employees and union entities,” which might render the litigation unmanageable.

In a subsequent decision, Holmes v. Securities Investor Protection Corporation, the Supreme Court applied this standard to a claim brought under the Racketeer Influenced and Corrupt Organizations Act (RICO). There, the Court held that an entity obligated to reimbursed economically injured broker-dealers could not demonstrate proximate cause. Despite legislative history suggesting that RICO should be liberally construed to effectuate its remedial purpose, the court determined that the plaintiff must demonstrate a “direct relation between the injury asserted and the injurious conduct alleged.” Primarily, the Court relied on the Associated General Contractors Court’s direct harm analysis because Congress drafted RICO based on the Clayton Act.

Here, the Holmes Court found that the plaintiff was unable to prove a sufficiently direct injury. In addition to relying on statutory policies, the Court invoked some policies that derive from the common law. As the Holmes Court observed, in the case of those indirectly harmed, it would be difficult to trace whether the plaintiff’s harm was due to other causes, such as “the broker-dealers’ poor business practices or their failures to anticipate developments in the financial markets.” The Court emphasized that those who were more directly injured had sufficient incentives to bring suit to remedy their losses, alleviating the concern that tortious conduct would occur without consequence. Critically, the Court recognized that the defendant had limited resources and that to allow those indirectly injured to recover might deprive the directly injured of any compensation. Further, as the Holmes Court explained, beyond this, the “claims of the indirectly injured would force courts to adopt complicated rules apportioning damages among plaintiffs removed at different levels of injury from the violative acts, to obviate the risk of multiple recoveries.”

---

254 Id. at 542.
255 Id. at 544.
256 Id. at 545.
258 Id. at 268 (Holmes).
259 Id. at 267 (Holmes).
260 Id. at 272.
261 Id. at 269 (citing cases).
These same policy considerations were examined at length in another RICO case, *Anza v. Ideal Steel Supply Corp.* The plaintiff Ideal, a company that supplies equipment for making steel, challenged the actions of its competitor National. Ideal alleged that National violated RICO by selling products tax-free to customers who paid cash, which amounted to a price discount. According to Ideal, National then mis-reported its tax earnings by submitting fraudulent tax returns to the New York State Department of Taxation and Finance to conceal these transactions. Applying the *Holmes* standard, the Court ruled that Ideal could not establish proximate cause for its lost sales due to price competition from National’s tax-free policy for customers who paid cash. The *Anza* Court explained that the directly injured entity was the state of New York and that there were numerous policy reasons for denying relief to Ideal. The Court observed “[o]ne motivating principle is the difficulty that can arise when a court attempts to ascertain the damages caused by some remote action.” In a passage that evokes Andrews’ *Palsgraf* dissent, the *Anza* Court noted that ascertaining damages for those who were indirectly harmed posed almost insurmountable problems. For example, here the Court explained, “Ideal's lost sales could have resulted from factors other than petitioners' alleged acts of fraud. Businesses lose and gain customers for many reasons, and it would require a complex assessment to establish what portion of Ideal's lost sales were the product of National's decreased prices.”

In *Lexmark v. Static Control Systems*, the Court considered whether a manufacturer of parts used to refurbish Lexmark’s ink toner cartridges had standing to bring a federal false advertising claim. Static Control allegations targeted Lexmark’s “Prebate” program, which offered customers a 20% discount so long as the customer agreed to return the cartridge after it was spent. As a supplier of parts to entities that refurbished cartridges, the plaintiff’s business depended on customer’s willingness to purchase refurbished cartridges. Static Control alleged that, through the Prebate program, Lexmark essentially misrepresented to consumers and that Lexmark cartridges purchased through the Prebate program could not be refurbished. Also, Lexmark sent letters to refurbishers stating that it was illegal to sell refurbished toner cartridges that use Static Control’s parts.

The *Lexmark* Court considered whether the plaintiff had alleged a harm that was proximately caused by Lexmark, an inquiry which it explained “is controlled by the nature of the statutory cause of action.” In concluding that Static Control’s allegations met the standard, the Court focused on three areas of inquiry. First, the plaintiff’s harm must be within the scope of the rights protected by the statute. Second, the injuries alleged must be proximate caused by the

---

263 *Id.* at 454.
264 *Id.* at 458.
265 *Id.* at 459.
267 *Id.* at 1384 (Specifically, the Court summarized Static Control’s claims that “it alleged that through its Prebate program Lexmark ‘purposefully misleads end-users’ to believe that they are legally bound by the Prebate terms and are thus required to return the Prebate-labeled cartridge to Lexmark after a single use.”).
268 *Id.* at 1391.
269 *Lexmark*, 134 S.Ct. at 1393. For this element, the Court stated that the plaintiff met the requirement by alleging that it was “a ‘perso[n] engaged in’ ‘commerce within the control of Congress’ whose position in the marketplace
defendant’s conduct—that is, they must have a sufficiently close connection to the conduct prohibited by the statute. In Lanham Act cases, this would be met by a competitor whose injuries included sales that were lost because of the defendant’s misrepresentation. Additionally, the Lexmark Court stated that reputational harms that stemmed from the defendant’s statements were actionable under false advertising law. With respect to this factor, the Court observed “[w]hen a defendant harms a plaintiff’s reputation by casting aspersions on its business, the plaintiff’s injury flows directly from the audience’s belief in the disparaging statements.”270 The Court noted that, according to the plaintiff’s allegations, the defendant’s statements implied that Static Control’s customers were engaged in illegal activity. If, based on that representation, consumers ceased buying refurbished cartridges, then the Court reasoned that this was sufficient direct to impact Static Control’s sale of parts and therefore harm the plaintiff’s income.271 Certainly, the Court could not have done so in a false advertising case, because that cause of action which does not allow consumers standing to bring suit. Nonetheless, the Court widened the ability to demonstrate proximate cause beyond competitors and allowed this parts supplier to demonstrate harm by extending directness to encompass anything that was not “‘too remote’ from the defendant’s unlawful conduct.”272

In contrast, CSX Transportation v. McBride, the Supreme Court relaxed the proximate cause standard for negligence claims filed under the Federal Employers’ Liability Act (FELA).273 The relevant section of this act allows employees to recover from railroads for personal injuries or deaths that result “in whole or in part from [carrier] negligence.”274 The CSX Transportation Court determined that a combination of reasons supported this result, including the statute’s “in whole or in part” language that the Court construed to expand recovery compared to common law proximate cause. The Court relied on Congress’ intent to enact FELA with humanitarian and remedial goals.275

In result, the CSX Transportation found that the jury’s common sense was the primary limit on compensation worker injuries. This relaxed standard allows compensation for a broader class of harms compared to common law negligence cases.276 As the statute limits recovery to federal employees injured on the job, the court found that proximate cause did not add any further limitation on the type of plaintiff who could recover.277 Throughout, the opinion relies on a close reading of the statute to form the iteration of proximate cause in the FELA context. Unlike many other statutory proximate cause cases which draw heavily on common law

270 Id. at 1393.
271 Id. at 1395 (“Static Control’s allegations suggest that if the remanufacturers sold 10,000 fewer refurbished cartridges because of Lexmark’s false advertising, then it would follow more or less automatically that Static Control sold 10,000 fewer microchips for the same reason…”).
272 Lexmark Int’l., 134 S. Ct. at 1390.
274 Id. at 688 (quoting 45 U.S.C. § 51).
275 Id. at 691-692.
276 Id. at 703-704 (“If negligence is proved, however, and is shown to have “played any part, even the slightest, in producing the injury. . .then the carrier is answerable in damages even if “the extent of the [injury] or the manner in which it occurred” was not ‘probable’ or ‘foreseeable.’”) (citation omitted).
277 Id. at 704 (finding that there are “there are no unforeseeable plaintiffs in FELA cases.”).
principles drawn from negligence cases, *CSX Transportation* deeply embeds the legislative structure and language throughout the legal standard in a comprehensive manner.

**VII. Proximate Cause and the Patent System**

**A. Background**

The Supreme Court’s presumption for proximate cause applies to patent cases. The word “compensate,” long used in the common law to include proximate cause, together with the objective of the compensation (“use made of the invention”) suggests a necessary link between the plaintiff’s harm and the monetary award. Some nineteenth century cases use language that suggests that proximate cause was considered in awarding damages. In the Federal Circuit’s *Rite Hite* decision, proximate cause is expressly discussed as a governing principle for lost profit awards. Most recently, the Supreme Court’s 2018 *WesternGeco* opinion refers to proximate cause as an issue that could limit the reach of lost profits damages. As part of a larger trend of the Court’s presumption that proximate cause is to be read into all federal statutory causes of action, express consideration of this element in the patent system is overdue. Under this reasoning, proximate cause is read into the statute to apply to both lost profits and reasonable royalty forms of recovery.

In terms of the form that proximate cause might take, the Federal Circuit’s *Rite-Hite* majority includes a curious mix of terminology. Although the decision characterizes proximate cause as a matter of foreseeability, the opinion proceeds to discuss that “[s]uch labels have been judicial tools used to limit legal responsibility for the consequences of one's conduct that are *too remote* to justify compensation.” The court therefore appears to be relying on both foreseeability and the remote consequences by using verbiage that suggests that the court considers that proximate cause is, as a policy matter, a doctrine that is used to address the

---

278 See Lexmark Intern., Inc. v. Static Control Components, Inc., 134 S.Ct. 1377, 1390 (2014) (“...we generally presume that statutory cause of action is limited to plaintiffs whose injuries are proximately caused by violations of the statute.”).

279 35 U.S.C.A. § 284 (West)(“Upon finding for the claimant the court shall award the claimant damages adequate to compensate for the infringement, but in no event less than a reasonable royalty for the use made of the invention by the infringer . . . .”); see also Lucent Techs., Inc. v. Gateway, Inc., 580 F.3d 1301, 1324 (Fed. Cir. 2009)(“awarding damages through litigation attempts to assess ‘the difference between [the patentee's] pecuniary condition after the infringement, and what his condition would have been if the infringement had not occurred.’”)(citing Yale Lock Mfg. Co. v. Sargent, 117 U.S. 536, 552 (1886)).

280 Carter v. Baker, 5 F. Cas. 195, 202 (C.C.D. Cal. 1871) (“Their profits must be the direct and legitimate fruits of that patent. They may have sustained damages from this source, but they are too remote. It rarely happens, that all the damages, incidental and remote, resulting from a wrongful act, are permitted to be recovered by the law.”); Buerk v. Imhaeuser, 4 F. Cas. 594, 595 (C.C.S.D.N.Y. 1876) (“The damages in such a case must be confined to the direct and immediate consequences of the infringement, and not embrace those which are both remote and conjectural.”).


283 *Rite-Hite*, 56 F.3d at 1546 (emphasis added).
“background question whether the asserted injury is of the type for which the patentee may be compensated.”

This nuanced point is key. As the Rite Hite court recognized, proximate cause is a tool that can be used by the judiciary to properly implement the purposes of patent law, observing that “judicial relief cannot redress every conceivable harm that can be traced to an alleged wrongdoing.”

As the court explained, “[t]he general principles expressed in the common law tell us that the question of legal compensability is one to be determined on the facts of each case upon mixed considerations of logic, common sense, justice, policy and precedent.”

B. **Doctrinal Considerations**

As previously described, as a doctrine proximate cause has two dimensions—factual cause (or cause-in-fact) and proximate cause. Although these doctrines might not be significant in all cases, in some circumstances the doctrine would allow the system to better align damage awards with the purpose of the law.

Perhaps because they are faced with evidentiary standards that require more demanding economic standards of proof in accordance with the Daubert standards, over the past several years patentees have sought ever-widening theories of recovery. Significant here, some rely on incomplete statements of proximate cause, such as the proximate cause concept of harm that follows in a “natural and unbroken” sequence from the original infringement. In other words, some companies have asserted the right to relief from damages that “inexorably flow” from the infringement. This parallels the patentee’s argument in the Power Integrations case, which sought to expand liability to worldwide activity as a foreseeable consequence of domestic infringement. Perhaps the reasons that the courts view these theories with some skepticism is that they rely on incomplete statements of proximate cause. Specifically, these efforts omit all of the limitations that are central to the proximate cause inquiry. By relying on expansive language alone, these efforts stand on flawed theoretical grounds. To fully implement the doctrine, consideration of its limitations must be incorporated.

As CSX Transportation v. McBride suggests, the doctrine operates best when the fundamental structure of proximate cause doctrine is modified to fit the relevant statute. Application of proximate cause as it exists in negligence law would be a poor fit to further the purposes of the patent system, including with concept that intangible information and the

---

284 Id.
285 Id.
286 Id. at 1547 (quotations and citations omitted).
288 See, e.g., Edwards ex rel. Fryover v. Anderson Eng’g, Inc., 45 Kan. App. 2d 735, 740 (2011) (“Proximate cause is cause which in natural and continuous sequence, unbroken by an efficient intervening cause, produces the injury, and without which the injury would not have occurred, the injury being the natural and probable consequences of the wrongful act”) (citations omitted).
290 See footnote 38 and accompanying text.
dynamics of the commercial innovative context is distinct from the principles governing physical injury. Certainly, proximate cause holds significant promise to open a policy space for the patent system to address a number of issues that have remained unresolved. Delivery of this promise depends on the law’s willingness to adapt the law to the patent context.

One question arises as to the proper form on inquiry. If the negligence standard and the Federal Circuit’s *Rite Hite* are any guide, both foreseeability and the directness standard might be sufficiently malleable to accommodate the policy drivers for patent cases.\(^\text{291}\) Notably, with some exceptions, the general trend in Supreme Court cases has relied on the directness standard, or according to *Lexmark* the “not too remote” standard. This standard appears to be better suited for commercial cases, which are distinct from negligence law’s foreseeability framework. Significantly, relying on the foreseeability standard in statutory causes of action alone might pull decision-makers toward using negligence precedent, which has little relevance to the patent context, which is akin to strict liability.\(^\text{292}\) Additionally, the factual circumstances that give rise to liability in the commercial context, which include false advertising, antitrust and patent law, are more alike compared with the wide-ranging and unpredictable possibilities that give rise to negligence claims. Thus, the directness standard may more appropriate for commercially-based torts.\(^\text{293}\)

Applying proximate cause to the patent system’s purposes presents some challenges. As a practical matter, the purpose of the patent system has not been consistently articulated.\(^\text{294}\) Although the constitutional aim is broadly stated “to promote the [p]rogress of [s]cience and useful [a]rts,”\(^\text{295}\) modern sources cite the system’s potential to drive innovation, job growth, and a successful economy.\(^\text{296}\) Alternatives consider the distributive and social justice aspects of the system.\(^\text{297}\) As a practical matter, rights holder use them to licensing, create new forms of revenue,

\(^{291}\) *Rite Hite*, 56 F.3d at 1546 (stating “For example, remote consequences, such as a heart attack of the inventor or loss in value of shares of common stock of a patentee corporation caused indirectly by infringement are not compensable” and “the reasonable limits of liability encompassed by general principles of law can best be viewed in terms of reasonable, objective foreseeability”) (emphasis added).


\(^{293}\) Timothy Richard Holbrook, *Extraterritoriality and Proximate Cause after WesternGeco*, Yale Journal of Law & Technology (Forthcoming); Emory Legal Studies Research Paper (Oct. 2018); available at SSRN: [https://ssrn.com/abstract=3259667](https://ssrn.com/abstract=3259667) (observing that “the blanket ‘foreseeability’ test is prone to abuse. Indeed, it is an ever-changing standard.”).


\(^{296}\) *ECON. & STATISTICS ADMIN. & U.S. PATENT & TRADEMARK OFFICE, INTELLECTUAL PROPERTY AND THE U.S. ECONOMY: INDUSTRIES IN FOCUS i* (Mar. 2012) (“Innovation protected by IP rights is key to creating new jobs and growing exports. Innovation has a positive pervasive effect on the entire economy, and its benefits flow both upstream and downstream to every sector of the U.S. economy.”); See also VANNEVAR BUSH, SCIENCE: THE ENDLESS FRONTIER, A REPORT TO THE PRESIDENT (1945) (discussing the importance of scientific innovation).

\(^{297}\) See MADHAVI SUNDER, FROM GOODS TO A GOOD LIFE: INTELLECTUAL PROPERTY AND GLOBAL JUSTICE 178 (2011); Anupam Chander & Madhavi Sunder, *Foreword: Is Nozick Kicking Rawls’s Ass?*, 40 U.C. DAVIS L. REV. 563, 568–71 (2007); Peter Lee, *Toward a Distributive Commons in Patent Law*, 2009 WIS. L. REV. 917, 928–31 (criticizing the patent system as being a detriment to low-income populations); See also ROBERT P. MERGES, *JUSTIFYING INTELLECTUAL PROPERTY* 121–36 (2011) (discussing how current intellectual property law addresses
securitize assets, create defensive portfolios, and as reputational signals to facilitate business transactions. Further, companies use patents to capture the value of their research and development, which is useful for acquisitions and sales. Resolving the conflicts between these different goals of this system is beyond the scope of this piece. Nonetheless, the potential for one or any of these to drive policy results through the mechanism of proximate cause must be kept in mind.

VIII. Resolving Particular Problems in Patent Law

A. International Considerations and the Scope of the Statute

As an initial matter, given the strong policy against the extraterritoriality of the application of U.S. patent rights, the Supreme Court’s suggestion that proximate cause might limit liability for extraterritorial infringement appears sound. The doctrine provides a useful platform to terminate liability for factual circumstances like those in *WesternGeco*. As a practical matter, foreign sales activities are outside the scope of a domestic patent. As previously argued, there is a question of whether a plaintiff can establish any harm for the use of a U.S. patent for overseas activity. Because a domestic patent has no power overseas, the sales contracts that are the subject of *WesternGeco* are outside the scope of the right and therefore are unrecoverable. In U.S. law, there are two hundred years of precedent upholding the principles that U.S. patent rights are territorial, absent compliance with the statutory exceptions. To the extent that the scope of the statute inquiry in taken seriously in the proximate cause inquiry, as least as seriously as it in negligence law, this inquiry would terminate liability for ION’s overseas activity under the “harm outside the scope of the risk” rule. To the extent that patentees seek recovery for such activity, their options include using patents filed internationally to prevent such conduct.

Moreover, there is a distinction between the act of “making” an infringing machine domestically under section 271(a) and the follow-on international sales from the use of the machines abroad. In *WesternGeco*, ION was liable for the domestic manufacture of the infringing machines and was liable for a reasonable royalty for that activity. The patentee sought additional lost profits damages for the contracts that were international in scope. Here, the patentee received full compensation for all direct harms from the infringement in the form of a reasonable royalty. Applying standard proximate cause analysis, the harm from the performance of overseas contacts is an indirect harm under patent law particularly given the harms to the system which flow from awarding damages for overseas activities through such mechanisms as the intervening, superseding cause. In other words, this act represents a point where proximate cause can cut off the causal chain of claimed damages.

distributive justice); Madhavi Sunder, Novartis v Myriad: The Indian and U.S. Supreme Courts on Patents and Public Health, 35 EUR. INTELL. PROP. REV. 711, 714 (2013) (“[T]he purpose of intellectual property law need not be seen in narrow terms of incentives for innovation alone.”).


Draft - 2/20/2019
Do not distribute or cite without permission
This principle can be applied to the facts of the WesternGeco case. There, the jury awarded WesternGeco lost profits from contracts that ION had performed abroad using components assembled from parts shipped overseas from the U.S. The Supreme Court reasoned that these sales were causally connected to the U.S.-based infringing conduct of manufacturing the parts domestically for assembly abroad, which constitute infringing acts under section 271(f). Yet, considering the presumption against the extraterritorial application of U.S. patent law abroad as detailed in the Microsoft decision, the contracts performed might be struck from the judgment because they constitute an indirect harm. Thus, as conduct crosses over into areas of the world with their own sovereign patent law, such conduct might be viewed as subject to an intervening, superseding act as it crosses outside the U.S. territory. Alternatively, extraterritorial harms might be viewed as outside the scope of the risk of the infringement of a U.S. patent. If decisions that include the Supreme Court’s Microsoft case, which was part of a centuries-long tradition of applying a presumption against the extra territorial reach of U.S. patent law, means what it says, a reasoned application of proximate cause would cut off liability activities which occur off U.S. shores.

Similarly, this mechanism can also be used to address Power Integrations, in which the patentee claimed entitlement to damages based on worldwide sales under section 271(a). To the extent that reliance on this infringement subsection is sustainable, proximate cause can be used to cut off liability for foreign sales. As a practical matter, the patentee’s ability to obtain patents overseas under the TRIPS Agreement provides an ample remedy for offshore activity. As in the above example, reliance on either the lack of a cognizable harm, or the scope of the risk doctrine, are sufficient to terminate liability in that instance.

B. Social Value: Internalizing the Benefits of Innovation

The negligence system’s goal is to impose a just level of compensation that creates incentives to engage in an optimal level of risk. Recovery is not intended to maximize recovery for all plaintiffs regardless of the strength or justness of the claim. In the negligence context, it has been recognized that proximate cause operates as an important doctrine because “unless liability is limited in some way, people will avoid the activities in which their own breach of duty is predictably likely or especially costly, even if these activities are valuable to them and to the community.” Similarly, proximate cause for statutory torts is used to accomplish a number of goals associated with the relevant cause of action.

Currently, patent law awards damages aimed to award the full value of the infringed technology. By focusing on compensation and deterrence, the system does not account for any net social benefit in the infringement analysis, whether explicitly or implicitly. In patent litigation, liability is imposed under standards akin to strict liability against those who have no

---

301 Id. at 294.
actual notice of infringement. A defendant’s mens rea is not an element, at least for a direct infringement claim. Because of high search costs and fuzzy claim boundaries, even the most careful innovator may incur infringement liability. In other words, even the most careful innovators may find that they are systematically subject to patent infringement liability.

As a practical matter, some infringing conduct has both negative (infringing) impacts as well as positive (non-infringing) components. Particularly for multi-component products, the infringing portion may be minor or an insignificant to the overall operation of the device. In such cases, the heavy transaction costs and damages figures may discourage innovators from including features with positive social impact. In some cases, proximate cause may be implemented toward a closer alignment between the law’s purpose and the availability of monetary relief, particularly in close cases. In particular, this may occur where the plaintiff’s contribution to the defendant’s implementation is minimal and the social benefits of the defendant’s infringement is significant and its social value predominates in the infringing implementation.

The implementation of proximate cause creates a policy space for fact-finders to consider the broader social benefits of the defendant’s conduct. In other words, although an individual plaintiff may be harmed by the defendant’s conduct, this activity may give rise to positive externalities that render full liability unjust. At present, the violation of the patentee’s rights is the beginning and the end of the damages inquiry. The fact that a defendant’s conduct infringes the patent has been considered sufficiently blameworthy to crowd out all other considerations. Yet proximate cause theorists, including Gilead and Green, have recognized the importance of relying on proximate cause standards to allow defendants to internalize the social value of their conduct where the standards allow. That principle can be directly imported into patent law to prevent injustices that threaten to harm the social good. In such circumstances, it might be said that the harm to the plaintiff is too insubstantial to be considered a cause of the harm.

There are some available models to frame the examination of social value in patent law. As one example, to the extent that the defendant has added inventive elements to the infringing design, literature examining the social value of patents provides useful insight into the social benefits of intangibles. In this vein, those whose infringing implementations evidence an overwhelming percentage of inventive activity are in a materially different position from copyists who add little to nothing in the field. According to this literature, the positive externalities of this activity includes such examples as those that reduce the use of hazardous

---

chemicals, provide environmental benefits and assist the economically disadvantaged. Each of these includes implementations that have both a private benefit to the producer (as these products and services are sold for profit), as well as a public benefit that does not inure to the originator but to the public as a whole.

As has been recognized in the negligence context, in patent law the patentee’s individual harm may be at variance with the social value of the claimed invention. Although private individual may be harmed by a defendant’s patent infringement, the patentee may in fact obtain some benefits through the infringement. For example, it may be that the plaintiff is seeking to spread adoption of a particular industry standard that is furthered by the defendant’s sales of infringing product. In that instance, the infringement may deliver a mix of harms and benefits. In addition, there may be broader social benefits that derive from the defendant’s innovative activity. Proximate cause provides a doctrinal platform for the consideration of these factors to ensure that judgments are in accordance with the purposes of patent law. In such instances, it may be appropriate to use proximate cause to terminate liability using one of the many tools of proximate cause. For example, to the extent that the factual scenario mirrors those of the WesternGeco opinion, liability for overseas conduct might be terminated. Where the plaintiff benefits from the infringement more than is harmed, such damages might be considered to be beyond the nature of the statutory cause of action. In other word, a patentee whose business interests are furthered more than harmed by the defendant’s infringement might not establish a patent harm. This becomes an application of the policy that defines the scope of the relevant interest and stake in light of the law’s purposes. Just as some harms are not considered “antitrust harms” where violations of those laws are at issue, in the patent context the courts should reserve the ability to define whether a patentee’s asserted harm fits within the scope and purpose of the statute.

Such results might be particularly appropriate where the plaintiff’s contribution to the infringing device or process is minimal and the infringing implementation represents a significant innovative improvement. This may occur where the patent infringement is limited to a small portion of a multi-feature product. Under the current law, a patentee who seeks monetary damages against a defendant’s multi-feature product is limited to the value of the portion of the infringing device that incorporates the patented features. In other words, it has long been the law that the patentee:

…must in every case give evidence tending to separate or apportion the defendant's profits and the patentee's damages between the patented feature and the unpatented features, and such evidence must be reliable and tangible, and not conjectural or speculative; or he must show, by equally reliable and satisfactory evidence, that the profits and damages are to be calculated on the

---

310 Virnetx, Inc. v. Cisco Systems, Inc., 767 F.3d 1308, 1326 (2014) (“No matter what the form of the royalty, a patentee must take care to seek only those damages attributable to the infringing features.”).
whole machine, for the reason that the entire value of the whole machine, as a marketable article, is properly and legally attributable to the patented feature.\footnote{Garretson v. Clark, 111 U.S. 120, 121, 4 S.Ct. 291 (1884).}

Although a simple product might incorporate only a few features, some incorporate thousands.\footnote{See RPX Corp., Registration Statement (Form S-1) 55 (Sept. 2, 2011), available at http://www.sec.gov/Archives/edgar/data/1509432/000119312511012087/ds1.htm (“Based on our research, we believe that there are more than 250,000 active patents relevant to today’s smartphones . . .”); Amy L. Landers, \textit{Let the Games Begin: Incentives to Innovation in the New Economy of Intellectual Property Law}, 46 \textit{Santa Clara L. Rev.} 307, 341 (2006) (“[S]oftware and computers are examples of ‘system’ products – they comprise thousands, even hundreds of thousands, of individually functioning components and features all assembled in a package for a customer.”).}

For example, the average cell phone includes multiple hardware and software components that perform numerous complex tasks. A patentee’s right may extend to a remarkably small component within a much larger feature set. At the same time, the infringing implementation might evidence significant social benefit that, under the current state of the law, is not considered. This state of affairs might be an impediment for innovative companies to create products, even though doing so creates significant social benefits. In such cases, a patentee’s claim that it would have garnered a significant royalty might be trimmed, perhaps to nominal figure, given the purpose of the patent laws.

Proximate cause can be implemented as a policy lever where the infringing product or process generates a substantial social benefit, particularly there is a minimal technological contribution by the patentee. Treating non-willful infringers that create implementations which generate social goods, including public goods, should be considered in light of the goals of the patent system. Integrating proximate cause into the liability analysis can be used to direct the outcome of individual cases. In cases at the margins, there is room for policy to operate to more tightly align monetary damages with the purpose of the law. Just as the legal system forces defendants to internalize the costs of their conduct, proximate cause might allow some defendants to internalize the benefits as well. Moreover, this aspect of proximate cause can push patent law toward accomplishing social good, as the system was ultimately intended to do.

\section*{C. Apportionment and the Multiple Plaintiff Problem}

As previously described, a patentee is entitled to recover damages for the use of the claimed invention. Yet some devices incorporate multiple features, which raises the possibility that multiple patentees will claim infringement for different functions or portions of the same accused device or process. In theory, each plaintiff would seek and obtain their individual share of recovery. Yet there is no central coordinating oversight mechanism that assures a fair result under those conditions. As a practical matter, each patentee will time their lawsuit for their own convenience and financial advantage. There is no requirement that all patentees claiming the right to recovery to the same apparatus or process file suit such that they are before the court at the same time. This circumstance raises the problem that a defendant must pay overlapping amounts toward multiple plaintiffs through rounds of liability trials.
Further, patent royalty damages are not capped by the courts. Any single jury is entitled to determine the value of the invention as implemented in the infringing implementation without regard for any earlier or later judgement awarded to other patentees who filed suit with patents of their own. Assuming that the defendant has finite resources, patentees who sue earlier are more likely to recover the amount of the judgment compared to those who sue later after the defendant’s funds are exhausted.

Conditions created by the admissibility standards suggest that each of these awards may lead to overcompensation on the plaintiffs’ side. As a practical matter, the courts do not demand precision is estimating damages. Although in theory, the plaintiff’s recovery would be limited to the value of the technology, as a practical matter the courts admit that they “have long acknowledged that any reasonable royalty analysis necessarily involves an element of approximation and uncertainty.” The patentee’s proffered royalty is approximate at best and skewed to maximize recovery. The courts permit a broad range of evidence to be considered. For example, in one case a patentee was permitted to submit comparable licenses relevant to value of the patented technology as evidence of the appropriate royalty base despite the fact that at least one post-dated the relevant time period by two years, did not involve the patents-in-suit and failed to disaggregate non-patent related services from the technology at issue. More generally, the controlling test for determining the reasonable royalty amount has been criticized as delivering inaccurate results.

Juries, when presented with these estimates, are not well-equipped to fix damages figures with exactness. For infringement allegations that focus on infringing portions of multicomponent products, patentees are not required to isolate the feature that incorporates the patent with precision; rather this estimate may be approximate. As the Federal Circuit has explained “we have never required absolute precision in this task; on the contrary, it is well-understood that this process may involve some degree of approximation and uncertainty.” Each patentee, acting rationally to maximize their recovery in a lawsuit, would err on the side of over-inclusivity. In other words, if ten patentees seek recovery for infringement of ten different

315 ActiveVideo Networks, Inc. v. Verizon Commc’ns, Inc., 694 F.3d 1312, 1333 (Fed. Cir. 2012). Beyond the broad admissibility standard, there are myriad reasons to consider patent licenses as unreliable sources of information for patent damages. See Erik Hovenkamp & Jonathan Masur, How Patent Damages Skew Licensing Markets, 36 THE REVIEW OF LITIGATION 379, 384 (2017) (stating “licensing terms are actually less reliable as a proxy for harm than they would be if the licensing-based damages standard did not exist.”).
316 Daralyn J. Durie and Mark A. Lemley, A Structured Approach to Calculating Reasonable Royalties, 14 Lewis & Clark Law Review 627, 628 (2010) (“Simply handing the question of reasonable royalty to the jury, without more, is not a recipe for precision in damages analysis. But the fifteen-factor test may actually be worse because it overloads the jury with factors to consider that may be irrelevant, overlapping, or even contradictory.”).
317 See Bernard Chao and Roderick O’Dorisio, Saliency, Anchors & Frames: A Multicomponent Damages Experiment, Draft Paper at 4 (2019) (“Juries still issue what appears to be disproportionately high patent damage awards despite both Supreme Court and Federal Circuit decisions aimed at constraining these awards to the value of the patented feature.”) (draft on file with author).
318 Virnetx, Inc. v. Cisco Sys., Inc., 767 F.3d 1308, 1329 (Fed. Cir. 2014) (“a patentee must be reasonable (though may be approximate) when seeking to identify a patent-practicing unit, tangible or intangible, with a close relation to the patented feature.”).
319 Id. at 1328.
components of a single multi-function product, in the aggregate these patentees are quite likely to recover more than the total value of the infringed technology in the aggregate.

To the extent that overcompensation is occurring in these circumstances, attempting to award damages on a seriatim basis resembles a pie where every slice is slightly larger than it should be. As one source recognized for multicomponent products, “[i]n many situations, patentees are overcompensated – and socially valuable, but potentially infringing, commercialization is over deterred – as a result.”320 Although in theory, this issue could be avoided with better apportionment rules that solution has not been forthcoming. As Shavell’s work establishes in the tort context, correction at the proximate cause stage is needed to properly balance the purposes of patent law and to foster new invention. By imposing a heavy burden on those who have infringed, particularly where such infringement is not willful, patent law’s purpose of fostering new innovation cannot be met. A policy space that incorporates proximate cause is needed. Just as the Pittsburgh Reduction court placed liability on the most culpable actor, proximate cause can be used in a similar manner to maximize incentives to avoid patent infringement.

Beyond all of this, a plaintiff may establish a higher royalty based on sales of the full value of the infringing product under the entire market value rule.321 This plaintiff’s royalty is based on a percentage of the price of the entire product.322 When this rule is applied, a patentee may recover based on the value of an entire apparatus or system that contains several features or components “where the patented feature creates the basis for customer demand or substantially creates the value of the component parts.”323 Essentially, this rule operates as an exception to the rule that the patentee’s recovery is limited to the contribution of the patented component in a multi-component implementation. As one court explains, “[t]he entire market value rule indeed permits damages on technology beyond the scope of the claimed invention.”324 Logically, “if a patentee wins an entire market value rule case, no other patentee should be able to recover any damages based on the sale of the same product.”325 Yet there are currently no mechanisms in patent law which prevents this circumstance from occurring.326

For multi-component products, the totality of these circumstances can create the potential for overlapping liability that will weigh heavily on a single defendant. This might exhaust the defendant’s resources or, at a minimum, result in an outsized total liability. These possibilities raise some of the same apportionment and line-drawing problems identified by the Associated General Contractors Court as significant reasons to find that proximate cause is not met. As that

321 LaserDynamics, Inc. v. Quanta Computer, Inc., 694 F.3d 51, 67–68 (Fed. Cir. 2012) (discussing the application of the entire market value rule, stating “in any case involving multi-component products, patentees may not calculate damages based on sales of the entire product, as opposed to the smallest salable patent-practicing unit, without showing that the demand for the entire product is attributable to the patented feature.”).
325 Lemley, supra note 70 at 664.
326 Id.
court identified, proximate cause can be used to alleviate “the danger of complex apportionment of damages on the other.”

As this quote suggests, proximate cause may assist to resolve the multiple plaintiff problem. Given the potential for overlap in recovery, particularly if each award is inflated, an infringer is subject to multiple rounds of liability pays more than the conduct warrants. The current admissibility standards tolerate approximation and estimation. If one feature warrants application of the entire market value rule, all earlier awards on individual component parts are not reduced by that finding. A defendant’s total liability should be subject to more careful apportionment principles, particularly if the defendant’s implementation evidences the creation of social benefits. Just as negligence law does not ask for the elimination of all risk but rather optimal risk, the patent system should not create circumstances that unduly burden innovators, particularly those that infringe without notice of the patent in suit. Yet complex calculations that guard against the possibility of future judgements are, as a practical matter, extremely difficult—if not impossible—to make. These concerns highlight a special role for proximate cause where liability for multifunctional products are the subject of liability actions brought by multiple parties who are not before the court at the same time.

As a practical matter, proximate cause can be used to terminate liability in particular kinds of cases to minimize these concerns. For example, in some instances a defendant’s conduct with respect to a particular infringement feature may be indirect compared to others who might bear direct responsibility for the infringement. In others, the infringement may concern a de minimis feature which imposes onerous liability considering the nature of the infringing implementation. For this component, other policy concerns of patent law can be integrated into the analysis. In appropriate cases, this can integrate consideration of the social value of the infringing implementation.

D. Consumer Lawsuits

Sources have documented the practice that patentees deploy to sue all parties who might be engaging in infringement. This practice is possible because the patent system defines infringement as one who “makes, uses, offers to sell, or sells any patented invention, within the United States or imports into the United States any patented invention.” In addition, there are provisions which authorizes relief for other forms of activity, including for the indirect infringement of the claim. In a typical supply chain, one entity is responsible for “making” an infringing device, “selling” and “offering to sell” is done by a retailer, and consumers who purchase the product “use” the patented invention in the products or services that they purchase. These are treated as separate acts of infringement. On that basis, some patentees assert claims against anyone who engages in any of this activity, despite the fact that all lawsuits relate to the making, selling and use of the same device.

---

329 Id. section 271(b), (c), (f) and (g).
Patentees who seek to maximize the revenue obtainable from patent lawsuits and associated settlements have used the rules to file lawsuits against small companies, retailers and consumers who use the patented technology in addition to, or instead of, deeper-pocketed manufacturers. These consumer suits maximize the patent holder’s revenue through the ability to collect nuisance settlement amounts from more resource constrained parties who would rather pay quickly than pay for the full cost of defense. From the patentee’s perspective “serial nuisance filings against resellers or users quickly become more profitable than litigating on the merits against the original manufacturer.” These settlements lead to revenue without the risk of invalidity challenges that manufacturers have the incentives to raise in court challenges that proceed through summary judgment and trial.

Suits against consumers raise a plethora of thorny damages issues, including the potential overlapping and duplicative recovery. As described in the prior section, there is no assurance that the patentee’s recovery will necessarily reflect an accurate value of the patented technology. This is even more extreme for consumer suits because one of the primary reasons that patentees pursue small entities and customers is to obtain fast, early settlements that reflect a discount on the cost of litigation rather than the value of the technology. Even if some cases go to trial, as Brian Love and James Yoon point out, these suits are subject to distortive effects that skew monetary recovery upward.

Ideally, the patent system would develop rules to grapple with the total number of damage awards that a patentee obtains to ensure that there is not an aggregate level of overcompensation and to minimize double and overlapping recovery. At present, the patent system’s “current case law fails to achieve a socially optimal balance between patentees’ rights to enforce their patents and society’s interest in policing and properly valuing patented inventions.” Instead, the system has not attempted to so in favor of allowing patentees to engage in strategic behavior to maximize more than the value of their patent while minimizing the possibility of invalidation.

Using proximate cause’s directness standard and the doctrine of intervening cause, the doctrine can do some doctrinal work to minimize these distortions while continuing to compensate the plaintiff for legitimate harms. Many of the policies discussed by the Supreme Court’s Lexmark opinion support denying liability determinations to end-user infringers where the more culpable parties are available, including the directness of the plaintiff’s asserted injury, the policy of keeping the scope of trials within judicially manageable limits, and avoiding the risks of duplicative recoveries and complex apportionment. A plaintiff is capable of obtaining full compensation from manufacturing and sellers may be found to have obtain sufficient compensation for all harm suffered by the infringement. In this circumstance, the making and selling of the infringing good can be determined to be the proximate cause of the injury and the

---

331 Id. at 1605.
332 Love and Yoon, supra note 320 at 1634 (noting that seeking recovery against “larger device allows patentees to ask for larger damages amounts without appearing unreasonable”).
333 Id. at 1635.
consumer’s uses more remote (and therefore not compensable) causes of the patentee’s harms. As a policy matter, this allows the court to avoid duplicative and overlapping liability. Although such a determination is highly fact-dependent, as all proximate cause matters are, proximate cause can provide the policy levers necessary to allow patentees to protect their patent rights while policing society’s interest in ensuring proper valuation of inventions.

IX. Conclusion

The Supreme Court’s *WesternGeco* opinion serves as a useful prompt to consider the nature of the causation inquiry for both aspects of proximate cause—that is, cause in fact and proximate cause. Although much of the patent system has not considered these as viable elements of a claim for patent infringement, it is evident that both can undertake important roles to ensure that the system operates in a just manner. At present, there are instances where applying the elements of patent law to particular cases can result in judgments that threaten to chill valuable innovative activity, as well as activity that is beyond the scope of United States law. This result should not be surprising, given that this requirement serves a vital function across other areas of the law that are designed for civil recovery.

---

335 See Dmitry Karsh ted, *Causal Responsibility and Patent Infringement*, 70 Vand. L. Rev. 565, 625 (2016) (recognizing that manufacturers may be better positioned to avoid infringement, and therefore cheaper cost avoiders, compared to consumers).