A New Approach to Old Cases: Reconsidering Statutes of Limitation

Ehud Guttel
Michael Novick
Ehud Gutel* and Michael T. Novick**

A NEW APPROACH TO
OLD CASES: RECONSIDERING
STATUTES OF LIMITATION†

1 Introduction

To improve the chances of obtaining a correct legal decision, it is desirable to maximize the amount of evidence available to the court. Evidence is maximized to the extent that no existing evidence disappears and all potential evidence is uncovered.1 However, uncovering additional

* Faculty of Law, Hebrew University.
† Much of the research for this article was funded by a Fulbright Grant to Ehud Gutel and occurred at Yale Law School. I am greatly obliged to Professors Guido Calabresi, Dan Kahan, and Alvin Klevorick of Yale Law School. My thinking profited greatly from their comments and suggestions. I am also grateful for the comments of the participants in the law school’s graduate seminar and for those of Hanina Ben-Menachem, Amichai Cohen, Robert Hockett, Israel Gilead, Zohar Goshen, Eli Greenbaum, Alon Harel, Amnon Lehavi, Amit Levy, Daphna Lewinsohn-Zamir, Juhbin Meraji, and Alex Stein.

Since evidence collection is not costless, there is arguably a point at which the marginal cost of obtaining additional evidence exceeds the marginal benefit of improved accuracy. The issue is, of course, empirical. It is important to observe, however, that evidence maximization can in certain ways reduce the costs of the litigation process.

The assumption [that the costs of the trial are a positive function of the amount of evidence] fails because uncovering more evidence may preterm trial processes that would be more costly than the pretrial search for additional evidence. To put this more simply, in many cases, the more that is known about what happened, the greater the chance of a plea bargain or civil settlement. Since those stages of the trial that involve the presentation of and argument about evidence can be far more expensive than the evidence gathering stages, the expected cost of a trial will be a diminishing rather than a positive function of the amount of evidence gathered whenever the expected savings from the diminished likelihood of having to present the evidence to a court exceeds the expected costs of searching for additional evidence. Hence, although [the above] assumption is likely to pertain at the point a trial becomes inevitable, it is wrong to think of total trial costs as increasing always with increasing evidence. Moreover, if appeals and retrials are considered part of the expected trial cost, which seems consistent with the economic perspective, additional evidence can also reduce trial costs by reducing the likelihood that an appeal will be taken or, if there is an appeal, that a new trial will be ordered because of trial court errors.

Richard Lempert, ‘The Economic Analysis of Evidence: Common Sense on Stills’ (2001) 87 Va.L.Rev. 1619 at 1641. In the absence of empirical evidence to the contrary, we assume for the purposes of this article that the marginal benefit from additional evidence in terms of judicial accuracy and lower litigation costs always exceeds the marginal cost of acquiring such evidence.

(2004), 54 UNIVERSITY OF TORONTO LAW JOURNAL
evidence is usually not the principal barrier to evidence maximization: either litigant may typically request a stay of the proceedings in order to pursue new evidentiary leads. Instead, evidence maximization is principally a matter of preserving existing evidence. More particularly, it is essential to protect against the loss of existing exculpatory evidence. Since the plaintiff initiates litigation, we may presume that she will preserve all incriminating evidence. Exculpatory evidence, however, may disappear in the period between the harm and the plaintiff’s initiation of the suit, since the eventual defendant, unaware during this period that she will become the target of a lawsuit, lacks the incentive to maintain such evidence.

Since evidentiary loss is sometimes inevitable, it is also necessary to minimize the consequences of the judicial errors that such losses may generate. When incriminating evidence disappears, the risk of error rests, as an initial matter, entirely upon the plaintiff. The loss of exculpatory evidence, in contrast, exposes the defendant to the risk of error. The legal system should provide a mechanism for fairly and efficiently allocating the risk of error among the parties.

The basic argument of this article is that, in the context of civil litigation, a properly designed statute of limitation can effectively address these two evidentiary imperatives. The current regime of temporal limitation neither maximizes evidence nor adequately responds to the error risks associated with the absence of evidence. This article proposes an alternative regime capable of achieving both goals.

2 The law grants wide discretion to the courts in deciding the actual timing of the trial. Generally, the courts’ ability to require the losing party to compensate the opponent for any loss incurred as a result of stays requested by the loser allows courts to apply liberal stay rules. See notes 16 and 17 infra.

3 The potential role of statutes of limitation in improving the accuracy of judicial decision making has been, to this point, almost entirely overlooked by the legal academy. More generally, statutes of limitation have not, except in connection with adverse possession, been the subject of close academic attention. In his classic analysis of the common law, Oliver Wendell Holmes observed that statutes of limitation ‘never have been explained or theorized about in any adequate way.’ O.W. Holmes, Jr., ‘The Path of the Law’ (1897) 10 Harv.L.Rev. 457 at 476. William Blackstone ventured only that ‘the use of these statutes of limitation is to preserve the peace of the kingdom, and to prevent those innumerable perjuries which might ensue, if a man were allowed to bring an action for an injury committed at any distance of time.’ W. Blackstone, Commentaries on the Laws of England, Book III (1768) at 307. Even in the law and economics literature, where one might expect to find a systematic analysis of alternative possible models, statutes of limitation have generated little interest. The foundational authors address the topic very briefly. See, e.g., R.A. Posner, Economic Analysis of Law, 5th ed. (New York: Aspen Law & Business, 1998) at 90 (noting only that ‘[t]hey reduce the error costs that are caused by using stale evidence to decide a dispute’); W.M. Landes & R.A. Posner, The Economic Structure of Tort Law (New York: Aspen Law & Business, 1987) at 307 (describing briefly the impact of limitation periods on the costs of production); R.
Statutes of limitation currently appear in two forms. The first, and traditional, form defines a fixed period within which the plaintiff may file her claim and bars any claim filed after this period. The second form, the discovery rule, softens the traditional statute’s bar when the plaintiff is reasonably unaware, for some time after the harm occurs, of some of the facts essential to her claim. Under the discovery rule, the statutory clock begins to run only at the time when the plaintiff discovers (or should reasonably have discovered) all the necessary facts. Like the traditional statute of limitation, however, the discovery rule bars the plaintiff from filing after the statutory period has ended.

Because evidentiary loss and factual uncertainty are not operative factors in the current models of temporal limitation, these models do not respond systematically to the two evidentiary concerns identified above. Under the traditional statute of limitation, the only relevant concern is that the plaintiff come to court before the end of the statutory period. There is no reason for the plaintiff to file earlier rather than later within this period, although, given time’s corrosive effect upon the defendant’s exculpatory evidence, it is preferable, from an evidentiary perspective, that the plaintiff file as early as possible. The discovery rule, too, simply requires the plaintiff to bring suit at some point within a fixed period of time beginning at the moment when a reasonably diligent plaintiff could file. The plaintiff in a discovery rule regime may therefore wait to file until the end of the statutory period in order to inflict evidentiary harm upon the defendant. Thus neither the traditional statute of limitation nor the discovery rule creates adequate incentives for a filing timed to avoid evidentiary decay. This shortcoming is more than a theoretical


4 Under the common law, which did not entitle the plaintiff to prejudgment interest, some incentive for early filing existed. Under current law, however, which deems prejudgment interest to be an ‘element of complete compensation,’ to which the plaintiff is entitled (see Loftor v. Frank, 486 U.S. 549 at 558 (1987)), this incentive no longer exists. See also M. S. Knoll, ‘A Primer on Prejudgment Interest’ (1996) 75
flaw. It has been shown that plaintiffs in states with ten-year limitation periods for medical malpractice are 15 per cent more likely to succeed than plaintiffs in states with two-year limitation periods.5 Plaintiffs appear to take advantage of longer limitation periods to delay litigation and allow time to erode their opponents’ evidence.

Nor does the existing system of temporal limitation adequately respond to evidentiary decay that has already occurred. The late discovery scenario offers a clear illustration of this shortcoming. Here the plaintiff learns only after the moment of injury about the harm itself or about some other fact essential to her claim. Under a traditional statute of limitation, the plaintiff may file if the gap between the moment of injury and the moment of filing is not longer than the statutory period. Under the discovery rule, the statutory clock begins to run only from the time of discovery, so long as the plaintiff’s ignorance was not unreasonable. In either case, whenever the plaintiff is permitted to sue, she may take full advantage of any evidentiary loss by the defendant occasioned by the plaintiff’s delay. Neither rule protects the defendant against the consequences of such loss.

The two current models not only deny protection when evidentiary damage has occurred but also extend protection when no evidentiary damage has been suffered. If the plaintiff delays beyond the prescribed period, she is barred from filing, even though her delay may have caused the defendant no evidentiary harm. Here both the traditional rule and the discovery rule, because of their insensitivity to evidentiary loss, inequitably bar the plaintiff’s claim.

This article proposes a new model of temporal limitation that avoids the evidentiary shortcomings of the two current models. This new model transforms temporal limitation from a sanction rule to a price rule.6 The traditional rule and the discovery rule divide time into two zones and impose a sanction – the bar of the claim – when the plaintiff passes from the first into the second. The rule we advance never entirely bars the claim. Instead, it extracts a price that compensates the defendant for his evidentiary loss. This price consists of the total damages claimed by the plaintiff, discounted by the probabilistic value of the lost evidence. Thus,

Tex.L.Rev. 293 at 294 (noting that ‘the laws of most U.S. jurisdictions provide for interest on legal judgments from the time the claim arose until the date of judgment’). For an overview see D. Laycock, Modern American Remedies, 2d ed. (New York: Aspen Law & Business, 1994) at 210–3.


6 The terms are borrowed from R. Cooter, ‘Prices and Sanctions’ (1984) 84 Colum.L.Rev. 1925.
for example, if decay in the exculpatory evidence doubles the plaintiff’s chance of winning the case, from a baseline of 30 per cent to 60 per cent, the proposed model halves her potential damages. Because the proposed model attaches a price rather than a sanction to the plaintiff’s delay, we shall refer to it below as the ‘price model.’

The shift from sanction to price is associated with a shift of focus from the plaintiff to the defendant. Under the current models, the important issues are how long the plaintiff delays and, in the case of the discovery rule, whether her delay is reasonable. The price model shifts the focus to the defendant and to the damage that he suffers from the plaintiff’s delay. It no longer matters, under the proposed model, why or how long the plaintiff delays. The plaintiff may file her claim at any time. But whenever the passage of time erodes exculpatory evidence and thus increases the chance of an erroneous decision against the defendant, the plaintiff must compensate him accordingly, whether or not the plaintiff’s delay was reasonable. The proposed model, in short, allows the plaintiff to decide when the litigation will occur, but forces her to take into consideration the effect of this decision upon the defendant’s evidentiary situation.

The new model of temporal limitation addresses both of the evidentiary concerns identified above. First, because any erosion in the exculpatory evidence that derives from the passage of time also symmetrically erodes the plaintiff’s potential reward, she has no incentive to delay filing. By denying the plaintiff any gain from strategic delay, the price model advances the goal of evidence maximization. The proposed model is also sensitive to distortions that arise in cases such as late discovery where, because the plaintiff cannot respond to the model’s incentives, loss of evidence occurs. Where evidentiary decay attributable to the plaintiff’s delay creates the risk of an erroneous decision in favour of the plaintiff, the model protects the defendant by reducing the plaintiff’s reward. Furthermore, the model is not over-sensitive to the threat of evidentiary loss. While the current models bar dilatory plaintiffs even when the defendant has suffered no evidentiary damage, the price model does not; nor, when defendants are unharmed by the delay, does it penalize them.

The structure of the analysis is as follows. Part II elaborates upon the two current models of temporal limitation and describes the operation of the new model. In Part III, we describe the advantages of the price model by comparing it to the current models in each of the two scenarios from which litigation may arise. In the first scenario, the plaintiff is aware of her injury from the moment that it occurs. We show how the two existing forms of temporal limitation provide a windfall for the plaintiff and undermine the goal of evidence maximization by allowing her to delay filing without penalty until the end of the mandated period. The pro-
posed model, in contrast, denies the plaintiff the opportunity for strategi-
gic delay. In the second scenario, the plaintiff discovers her injury only
some time after the injury occurs, leading typically, but not always, to
evidentiary loss. The current models, because they are not key to the
defendant’s evidentiary loss, generate both inequitable and inefficient
results in such situations. They fail to act when the defendant has
suffered evidentiary loss, yet act when he has suffered none. The price
model, in contrast, responds properly in late discovery situations. When
evidentiary loss has occurred, it distributes the error costs associated with
such loss in a manner that is both fair and economically desirable. When
no evidentiary loss has occurred, the model permits the litigation to
proceed unaffected. Part III concludes with some empirical data that
support the theoretical analysis of the late discovery scenario. A broad
survey of outcomes in late discovery cases shows that defendants often
suffer evidentiary losses from which the current models fail to protect
them. The British experience with temporal limitation demonstrates the
way in which the current models can overprotect the defendant, barring
the plaintiff from suit even when no evidentiary loss has occurred.

Although most of the discussion to this point is normative, the analysis
of the three models of temporal limitation also has significant descriptive
power. Part IV shows how the shortcomings of the current rules of
temporal limitation to which the price model responds also explain
inconsistencies in the application of existing statutes of limitation. The
proposed move away from the temporal limitation regime of today is, we
will show, prefigured in courts’ ambivalent attitude toward that regime.

Part V places the argument within a broader context. It explains how
the suggested shift from a sanction to a price model aligns temporal
limitation with a larger legal trend characterized by the substitution of
price-like liability rules for sanction rules. Indeed, we will show that the
proposed model can be formulated as the synthesis of two such recently
evolved price doctrines.

The principal assumption underlying the argument for the price
model is that temporal limitation should fulfill primarily an evidentiary
function. Part VI defends this assumption. We show how non-evidentiary
concerns often associated with statutes of limitation are in fact ade-
quately addressed by other temporal rules. These other rules, however,
do not effectively promote evidence maximization or protect against
evidence decay. This functional analysis offers persuasive grounds for the
conclusion that temporal limitation rules should be dedicated chiefly to
evidence-centred ends. The price model, which best achieves these ends,
is thus the most attractive and efficient form of temporal limitation. In
Part VII we defend the practicability of the price model by showing, first,
that courts routinely make the sort of counterfactual determinations that
it requires and, second, that the price model eliminates as many litigation costs as it generates.

II Three models for statutes of limitation

Current statutes of limitation fall along a spectrum. At one end lies the traditional model of temporal limitation (the 'first' model). Like all statutes of limitation, the first model has two basic elements: the time period during which the plaintiff must initiate her lawsuit and the consequence of not initiating the suit within that period. In the first model, the first element – the time period – is completely rigid. Neither special circumstances associated with the damage nor special attributes of the plaintiff can change it. The second element – the consequence of delay – is likewise inflexible in the first model. After the allotted time passes, the plaintiff’s claim is completely barred, and she loses any right to compensation.

At the other end of the spectrum lies a second, more flexible paradigm (the 'second' model). The first element – the period granted the plaintiff to initiate her claim – is not as rigidly defined in terms of specific units of time. Instead, the second model adjusts the period according to the facts of the particular case. These include, on the one hand, the characteristics of the plaintiff, such as mental capacity, age, or marital status; and, on the other hand, the circumstances of the injury, such as the latency of the damage, the type of harm, and the scientific knowledge available at the time of the alleged harm. The model’s flexibility is, however, confined to its first element. The second element – the outcome of delay – is, as in the first model, rigidly absolute. Once the statutory period has passed, the plaintiff is barred from adjudicating her claim. Thus, while, under the second model, the period of time in which the plaintiff can bring her claim is variable, failure to bring it within the given period carries the same consequences as in the first model. Moreover, the apportionment of the risk of error associated with evidentiary changes remains the same. Up to a certain point, the defendant bears all the risk of error associated with the destruction of exculpatory evidence. After that point, the balance shifts radically: the plaintiff’s claim expires and the defendant is shielded from all risk.

A common instance of the first model is the statute of repose. Statutes of repose establish periods of time, determined ex ante by legislation, within which a plaintiff must initiate her suit. Failure to comply with this requirement completely bars the plaintiff’s claim. These statutes do not permit any equitable pleas for extension. The doctrine of laches exemplifies the second model. Under this doctrine, a court does not invoke a predetermined period but, instead, asks whether the plaintiff exercised
due diligence in bringing her claim. In answering this question, the court may consider any relevant factor.\(^7\) Between statutes of repose and laches fall a variety of temporal limitation statutes that are less rigid than the former but more rigid than the latter. These statutes prescribe fixed periods of time but toll the running of the statutory period in response to certain conditions.\(^8\) The more conditions a statute recognizes, the less it resembles statutes of repose and the more it approximates the laches model.

In the area of civil litigation, although statutes of limitation characterized by double rigidity occasionally occur, the second model dominates in the form of the discovery rule.\(^9\) Rather than dictate a fixed time

---


\(^8\) Most commonly, the statutory period is tolled because the plaintiff was incarcerated, disabled, or a minor when the harm occurred or because the defendant fraudulently concealed the harm. For an overview of the relevant factors see C.W. Corman, Limitation of Actions, vol. 2 (Boston: Little, Brown, 1991) at 1-99.

\(^9\) The discovery rule evolved in the second half of the nineteenth century as part of a general movement towards viewing fault as a ground of liability. While it was initially applied only selectively, it quickly became the dominant form of temporal limitation. For the prevalence of the discovery rule in Canada see K. Roach, 'Reforming Statutes of Limitations' (2001) 50 U.N.B.L.J. 25 (reviewing both legislation and court decisions in Canada applying the 'discovery rule'). For the similar expansion of the rule in the United States see, for example, Corman, Limitation of Actions, supra note 8 at 133-70 ('The acceptance of the discovery rule was rather fast and by 1975 over one-half of the states had adopted some form of a discovery rule, either by statute or by judicial interpretation. Moreover, the discovery rule was sometimes adopted even in the face of opposing legislation'); Griffin v. Correia-Collahan, 74 F. 3d 36 at 38 (1995) (indicating that more than forty states favor the discovery rule); S.D. Gimcher, 'Statutes of Limitations and the Discovery Rule in Latent Injury Claims: An Exception or the Law?' (1982) 43 U.Pitt.L.Rev. 501 (demonstrating the discovery rule's development from a narrow exception into the standard form of temporal limitation); 'Accrual of Statutes of Limitations: California's Discovery Exceptions Swallow the Rule' Note (1980) 68 Cal.L.Rev. 106 at 108 (showing how the discovery rule has become the dominant rule in California); 'Stiles v. United States: FTCA Expanding the Discovery Rule in Occupational Disease Cases' Note (1981) 14 John Marshall L. Rev. 873 at 881-2 (observing the increasing popularity of the discovery rule in the occupational disease context). Even after the heyday of the fault-centred notions on which the discovery rule is based, courts and legislators continue to apply it. In the first edition of the Restatement of Torts, the commentary on the section dealing with statutes of limitation stated briefly, 'Traditionally, [k]nowledge by the injured person of the existence of the tort was immaterial, and it is still true in many of the states.' Restatement (First) of Torts §899 comment. (e) (1934). When Dean Prosser revised the Restatement forty years later, he added that 'in the wave of recent decisions [it was held] that the statute must be construed as not intended to start to run until the plaintiff has in fact discovered the fact that he was suffered injury or by the exercise of reasonable diligence should have discovered.' Restatement (Second) of Torts §899 comment (e) (1974). A decade later, in the fourth edition of Prosser's well-known treatise on torts, the discovery rule was presented as an 'infectious' rule which 'has been spreading from doctors to dentists, accountants
period, the discovery rule establishes the temporal boundary on the basis of fault and reasonableness. Though somewhat narrower than laches, the discovery rule, too, asks courts to determine whether the plaintiff acted reasonably. So long as she did not know and could not have known the material facts of the case prior to the time when she files suit, and therefore was without fault in delaying, her claim is deemed timely. The prevailing mode of temporal limitation thus pairs a flexible time period during which the litigation must begin with a rigid consequence for not initiating the suit within that period.

... manufacturers of defective products, and a miscellany of negligence and other tort actions.' Prosser & Keeton on Torts, 4th ed. (St. Paul, MN: West Publishing, 1984) at 166-7 [hereinafter Prosser & Keeton].

10 As a matter of drafting, the discovery rule occurs as a qualification of the basic limitation rule. Most current statutes of limitation prescribe a basic limitation period, typically three or four years, and then introduce the discovery rule as an exception. See, e.g., Lochinvar Corp. v. Meyers, 930 S.W. 2d 182 at 188 (Tex. 1996) ("The discovery rule is an exception to the general rule that a cause of action accrues when facts come into existence authorizing a claimant to seek a judicial remedy"). While these statutes thus appear to define the temporal period rigidly, the exception in fact renders it flexible. The only relevance of the defined period is for burden of proof. If the plaintiff wishes to initiate her suit after this period, she must convince the court that some of the elements of her claim were inaccessible up to the date of filing. See, e.g., Patimi Transfer Systems, Inc. v. Patimi Handling Solutions, Inc., No. Civ.A. 97-1568, 2001 WL 956641, at *16 (E.D.Pa. Aug. 16, 2001) ("Because the statute of limitations is an affirmative defense, the initial burden of establishing its applicability to a particular claim rests with the defendant. However, under Pennsylvania law, the burden shifts to the plaintiff if he or she asserts that the statute of limitations should be tolled by the discovery rule. If the plaintiff satisfies this burden, the discovery rule delays the accrual of a cause of action until the plaintiff was aware or should have been aware that an injury occurred") (citations omitted).

11 While the discovery rule accepts as a reasonable excuse for the plaintiff’s delay only unawareness with respect to the ingredients of her claim, laches takes cognizance even of practical impediments unrelated to the facts of the litigation. Thus, for example, the victim of an injury who knows all the facts necessary to state a claim but delays filing suit because she has a personal relationship to the injured would find her claim barred by the discovery rule; the doctrine of laches, however, might permit it. See, e.g., Samu v. Williams, 970 F. 2d 1043 at 1048 (2d Cir. 1992) ("Although plaintiff’s failure to pursue the matter aggressively until her adoptive father indicated his support if she desired to do so might be excusable for purposes of laches, it is irrelevant to the question of whether she had notice of her claim for statute of limitations purposes"). Today, for the most part, the terms are used largely interchangeably. See, e.g., Eberbard v. Harper-Grace Hospitals, 445 N.W. 2d 469 at 475 (Mich. App. 1989) ("[A] consideration of the question of laches will in all probability be academic on remand for the reason that the discovery rule applicable to the statute of limitations also incorporates this concept of due diligence. In other words, a finding that plaintiff was not lacking in due diligence would preclude a finding of the "compelling equities" necessary to invoke the doctrine of laches in an action at law"); Beumgort v. Koene Bldg. Products Corp., 666 A. 2d 238 (Pa. 1995) (describing the discovery rule and laches as parallel doctrines).
If statutes of limitation incorporate two elements – the period of time within which a claim must be brought and the consequence of delay – and either element may be rigid or flexible, then our two models do not exhaust the permutations of temporal limitation. The first model is rigid with respect to both elements. The second model is flexible with respect to the first element and rigid with respect to the second. The price model argued for in this article is a third possibility, a statute of limitation in which the first element is rigid and the second flexible. In this model, the time period within which the defendant bears all the risk of error from evidentiary loss is fixed. In the particular version of the price model that is the focus of this article, this fixed period is infinitesimal. In other words, the plaintiff is expected to bring her claim immediately after she suffers harm. The second element of the model – the consequence of delay – is, in contrast to the first two models, flexible: delaying past the statutory limit does not automatically bar the plaintiff’s claim.\textsuperscript{12}

The third model is simply the mirror image of the second. While the second model describes a flexible period and a rigid consequence, the third combines a rigid period with a flexible consequence. The reverse relationship also manifests itself in the two models’ contrasting foci. While the second model looks at the plaintiff, evaluating the reasonableness of her delay, the third model responds to the defendant and, in particular, to the evidentiary harm that he suffers.\textsuperscript{13} It compensates the defendant for the chance of an erroneous decision against him by reducing the plaintiff’s collection right in proportion to the evidentiary damage that her delay has caused the defendant.

To determine the extent to which evidentiary decay creates the risk of an erroneous decision, we must identify a baseline against which to measure error. The baseline situation is one in which all evidence is accessible and none has disappeared. Different courts in the baseline situation, though looking at the same evidence, may, given the latitude typically permitted by the law’s decisional standards, legitimately come to differ-

\textsuperscript{12} The two bipolar factors – first element and second element, flexibility and rigidity – permit, of course, a fourth model, an inversion of the first, in which both elements are flexible. We will, however, devote little time to an independent examination of this model, since, as will become apparent, it offers no advantage over the third.

\textsuperscript{13} Uncertainty arising from the destruction of the plaintiff’s evidence in the tort context is thoroughly treated in a recent study. See A. Porat & A. Stein, Tort Liability Under Uncertainty (Oxford: Oxford University Press, 2000) [hereinafter Tort Liability]. Porat and Stein demonstrate that a probabilistic approach analogous to the suggested price model can increase efficiency and achieve more equitable results. The book is an elaboration of arguments advanced earlier in A. Porat & A. Stein, 'Liability for Uncertainty: Making Evidential Damage Actionable' (1997) 18 Cardozo L. Rev. 1891. The present study follows their work in arguing that the law can and ought to redress evidentiary damage.
ent conclusions. One will give more weight to a particular unit of evidence and decide for the plaintiff; another will find a particular witness more credible and find for the defendant. Thus the baseline evidentiary situation in itself generates no verdict. It generates, rather, a probability distribution. If, for a given case, six out of ten courts would decide in favour of the defendant in the baseline situation, then the correct probability distribution is 60 per cent to 40 per cent in favour of the defendant. 14

When evidentiary loss causes the probability distribution to diverge from the baseline—that is, the distribution when all the evidence is available—the price model functionally restores the correct distribution by adjusting the plaintiff’s collection right. An example shows how the model operates. Suppose that a plaintiff claims $100 in damages. Suppose, too, that if she files immediately after the alleged harm, when all of the evidence relevant to the case is available to the court, she has a 40 per cent chance of winning. The plaintiff thus holds, in essence, forty (100 \times 0.40) statistical dollars in her pocket. Suppose, finally, that the plaintiff waits ten years to file suit and that her chance of winning thereby rises to 50 per cent, since the defendant loses evidence over the course of the decade. At the end of the ten years, the plaintiff has fifty (100 \times 0.50) statistical dollars in her pocket. By delaying, she has taken ten statistical dollars from the defendant. The plaintiff’s advantage arises solely from the decay of the defendant’s evidence, and thus solely from the greater likelihood that the court will rule erroneously in her favour. To prevent the plaintiff from receiving this windfall, her collection right should be reduced in inverse proportion to her improved chance of winning, so that her statistical sum remains the same. Thus, in the above example, the plaintiff, should she win, ought to collect only $80. Since her chance of success has risen to five-fourths of the baseline chance (from 40 per cent to 50 per cent), her collection right must be reduced to four-fifths of its original value (from $100 to $80). The plaintiff remains, in the end, with forty (80 \times 0.50) statistical dollars and has taken no statistical income from the defendant. The plaintiff thus gains nothing from her delay, and the defendant loses nothing. 15

In the proposed model, in contrast to the current models, there is no moment after which the plaintiff cannot bring her claim. In other words,

14 It is of course possible that in the baseline situation, the evidence so clearly favours one party that every court would decide the same way. In this case, the probability distribution is 100% to 0%, and we may speak not only of a baseline probability distribution but also, functionally, of a baseline verdict.

15 For the sake of clarity, the model ignores the time value of money and assumes that the litigants are risk neutral. Thus the plaintiff here is taken to be indifferent as to the choice between a 40 per cent chance of earning $100 now and a 50 per cent chance of earning $80 in the future. Relaxing these assumptions does not change the analysis but only introduces more variables to the proposed model.
the proposed model operates on the basis not of sanction but of price. Under the current models, a claim is either timely, and may be heard, or untimely, and barred. Rather than setting out two binary alternatives, the proposed model attaches a price to the plaintiff’s delay. The injured party can postpone her suit as long as she wishes, provided that she is willing to accept a reduction in the size of her potential compensation.

The price model not only changes the consequences of delay, it also frees us from inquiring into whether or not the plaintiff acted reasonably in delaying. The only relevant issue is the extent to which the plaintiff’s delay caused evidentiary damage to the defendant. It remains important to observe, however, that not every evidentiary advantage arising to the plaintiff from delay necessarily results from evidentiary damage done to the defendant. When, as in the example above, the probability of the plaintiff’s victory increases because the defendant has lost exculpatory evidence, then the plaintiff’s collection right should fall. The plaintiff ought not to receive a windfall that ultimately arises from the destruction of evidence and the consequent increased probability of error in favour of the plaintiff. If, however, the plaintiff’s improved chances are the result of new inculpatory evidence discovered during the time lag, then the plaintiff’s collection right ought not to be reduced. The plaintiff is entitled to the honest reward of her ability to present new evidence before the court.

III The advantages of the price model

The price model improves upon the current sanction rules of temporal limitation in two important ways. First, it eliminates the plaintiff’s incentive to delay filing and thus makes evidence maximization possible. Second, it avoids windfalls to either party. By reducing the plaintiff’s potential reward when her delay has caused evidentiary damage to the defendant, it prevents an injustice to the defendant. By refraining from imposing an artificial bar upon the plaintiff’s suit where the defendant has suffered no evidentiary damage, it ensures that the defendant will not enjoy an unjustified advantage. Part III of this article elaborates upon these two important features of the price model by comparing it to the form of temporal limitation currently dominant, the discovery rule. The comparison is made in the two situations from which litigation may arise. In the first of these, the plaintiff discovers the harm immediately after it occurs. In the second, this discovery comes at some later time. While this section contrasts the price model to the discovery rule in particular, its results may easily be extended to the other limitation rules: laches, statutes of repose, and the intermediate forms that lie between them.

In the analysis that follows, we make two assumptions. First, we assume that each party is capable of predicting the relative effects that the
passage of time will have on both parties’ evidentiary situations. In other words, the discussion supposes that the parties can evaluate whether the passage of time will reduce or improve their chances of victory. Second, for the sake of simplicity, the analysis assumes that both litigants can, after the filing of the case, request to postpone the proceedings if they anticipate that the passage of time will add to their evidence.\(^\text{16}\)

Although generally correct, these two assumptions do not always hold. First, litigants may sometimes be incapable of assessing how the passage of time will affect the evidence. To the extent that this is the case, however, the choice of temporal model is of limited evidentiary importance. If the parties are ignorant of the consequences of the running of time, they cannot be sensitive to incentives. Statutes of limitation can in this case serve little evidentiary purpose. This analysis focuses on situations where statutes of limitation can affect the parties’ behaviour. The second assumption may also not necessarily hold, since courts may refuse to grant continuances. It does, however, considerably simplify the analysis of the parties’ strategic behaviour under each of the models and, to a large extent, approximates actual results.\(^\text{17}\)

The analysis in this section proceeds by way of two illustrative evidentiary scenarios. The conclusion of the analysis is that the price model, in contrast to the discovery rule, can maximize evidence and avoid windfalls to either party. In fact, if the continuance rules are sufficiently liberal, then all the different permutations of evidentiary loss and discovery produce the same outcome: The plaintiff will file her claim as soon as

\(^{16}\) The actual time at which litigation begins is a matter left largely to the court’s discretion. Rule 40 of the Federal Rules of Civil Procedure (Assignment of Cases for Trial) requires that “[t]he district courts shall provide by rule for the placing of action upon the trial calendar (1) without request of the parties or (2) upon request of a party and notice to the other parties or (3) in such other manner as the courts deem expedient.” Similarly, the court is granted wide latitude in deciding pleas for the continuance or stay of the proceedings. See C.A. Wright & A.R. Miller, Federal Practice and Procedure (St. Paul, MN: West Publishing, 1971) at § 2552 (“Motions for a continuance of an action are addressed to the sound discretion of the district court... The grant of a continuance is not likely to be a ground of appeal or otherwise to find its way into a reported decision. Thus the published cases in which continuances were granted are few in number.”).

\(^{17}\) The ability of one party to stay or postpone the proceedings is limited by the harm that the delay would inflict upon the other. See, e.g., Wachovia Bank of Georgia v. Apex Tech of Georgia, 144 B.R. 649 at 652 (S.D.N.Y. 1992) (“[I]n reviewing the denial of a continuance ... the Court must determine whether additional evidence would be valuable enough to warrant the delay in the judicial process and the burden on other parties in the matter”). However, even where postponement of the litigation would harm one litigant, the other may still win a continuance by promising to compensate the damaged litigant. See, e.g., United States v. Three Parcels of Land in City of Fairbanks, State of Alaska, 224 F.Supp. 873 at 874 (D.C. Alaska 1963) (granting a continuance to the defendant in a condemnation case where the defendant agreed to pay reasonable costs that might be occasioned by the United States as a result of a continuance).
possible after the harm occurs, and then either the plaintiff or the
defendant, depending on which party is favoured by the aggregate evidence
to be uncovered after the moment of filing, will ask for a continuance in
order to gather new evidence. The plaintiff’s prompt filing, on the one
hand, protects against the loss of exculpatory evidence, and the continu-
ance, on the other, allows all the evidence accessible after filing to find its
way to court.

A IMMEDIATE DISCOVERY

1 Scenario 1
In this scenario, the evidentiary situation at the time the harm occurs
(T0) gives the plaintiff a 40 per cent chance of winning and the defend-
ant a 60 per cent chance of winning. In other words, were the evidence
available at T0 presented to a court, four out of ten courts would decide
for the plaintiff and six out of ten for the defendant. Between T0 and T1,
new evidence comes to light at a constant rate. By T1, all the remaining
evidence is discovered, some inculpatory and some exculpatory. The
inculpatory evidence improves the plaintiff’s chance by 10 per cent: In
other words, on the basis of the T0 evidence plus the additional inculpa-
tory evidence, five rather than four out of ten courts would find for the
plaintiff. The newly discovered exculpatory evidence improves by 20 per
cent the defendant’s chance of victory, meaning that seven of ten courts
faced with the T0 evidence, the new inculpatory evidence, and the new
exculpatory evidence would decide for the defendant. Between T0 and
the time the plaintiff files suit, however, the defendant, unaware that the
plaintiff will bring suit, does not take care to preserve his evidence. The
exculpatory evidence available at T0 decays at a steady rate until the
plaintiff files suit. By T1, exculpatory evidence with probative value of 20
per cent disappears. Postponement of the litigation from T0 to T1, rather
than improving the defendant’s chance of winning from 60 per cent to
70 per cent (60 - 10 + 20), actually reduces it, from 60 per cent to 50 per
cent (60 - 10 + 20 - 20).

The passage of time works, in this scenario, to the plaintiff’s benefit.
Though more exculpatory than inculpatory evidence emerges, the decay
of existing exculpatory evidence more than offsets this advantage to
the defendant. The plaintiff will therefore file as late as possible. The discov-
ery rule permits the plaintiff to delay, though the extent to which it
permits delay depends on how any particular court construes the rule’s
reasonableness requirement. The more liberal the court’s construction,
the longer the plaintiff may procrastinate. If the court interprets the
discovery rule as requiring that the injured party be aware of all possibly
relevant evidence, the plaintiff can delay her claim all the way until T1,
the time at which all evidence has come to light. If, in contrast, the rule is
construed narrowly, so as to require only awareness of the principal facts of the case, then the plaintiff must file relatively early.

In any case, under all but the most stringent constructions, the discovery rule permits the plaintiff to delay filing until some point after T0. For simplicity of analysis, assume that this point is T1. The outcome that the discovery rule produces is non-optimal in two ways. First, it fails to maximize evidence: Some of the exculpatory evidence available at T0 no longer exists at T1. Second, the discovery rule distorts the parties' respective odds. It is true that at T1, the evidentiary situation is richer than at T0. If we define the probative power of a body of evidence by its ability to improve the odds of the party that it favours, then between T0 and T1, evidence with probative power of 30 per cent (10 per cent for the plaintiff and 20 per cent for the defendant) comes to light, while evidence with probative power of only 20 per cent (all for the defendant) decays. This change in the evidentiary situation, however, actually does more harm than good. Were all the evidence available, the defendant would have a 70 per cent chance of winning and the plaintiff a 30 per cent chance. At T0, the distribution is 60 per cent to 40 per cent in favour of the defendant. The absence of the evidence to be uncovered between T0 and T1 causes a 10 per cent divergence from the baseline. At T1, the combination of discoveries and decay yields an even distribution of 50 per cent to 50 per cent. From T0 to T1, the divergence from the baseline increases from 10 per cent to 20 per cent, even though the court has access to more probative evidence at T1 than at T0. By permitting the plaintiff to delay filing until T1, the discovery rule distorts the probability distribution farther from the baseline, to the undeserved advantage of the plaintiff.

The price model responds to both shortcomings of the discovery rule by eliminating the plaintiff's incentive to manipulate the evidentiary situation. By reducing the plaintiff's potential collection right in proportion to the evidentiary damage inflicted by her delay, the price model forces the plaintiff to internalize the error costs that she generates. The discovery of new inculpatory evidence between T0 and T1 properly increases the plaintiff's odds, but she gains nothing from the decay, during that same period, in the defendant's exculpatory evidence. Therefore, in determining when to file, the plaintiff, under the price model, will simply compare the evidentiary benefit that delay would yield her with the evidentiary benefit that such delay would confer upon the defendant. Since in this case the evidentiary benefit that accrues to her (10 per cent) is smaller than the evidentiary benefit that accrues to her opponent (20 per cent), the plaintiff has nothing to gain from delaying until T1.

Under the price model regime, the plaintiff is in fact indifferent as to the timing of the litigation. If she files the claim at T0, the defendant,
alerted to the suit, will prevent the decay of any existing evidence. Knowing that the passage of time will bring to light more exculpatory than inculpatory evidence, he will request a continuance to T1. When the litigation begins at T1, the plaintiff will be eligible to collect her entire claim but will face a defendant whose evidentiary position is intact. If the plaintiff makes her initial filing only at T1, some of the defendant’s evidence will have decayed, and the plaintiff’s collection right will be proportionally reduced. When litigation begins at T1, the plaintiff will be able to collect only a part of her claim, but her opponent’s reduced evidentiary situation will improve her chances of collecting that part. Since, for the plaintiff, both starting times lead to the same result in statistical dollars, she will as easily choose one as the other.

The plaintiff’s indifference under the price model highlights the fact that the model denies her any windfall from the defendant’s evidentiary loss. The price model thus fulfills the second evidentiary goal of eliminating the error costs associated with evidentiary decay. At the same time, the plaintiff’s indifference means that the price model, unmodified, does not systematically maximize evidence. Plaintiffs will, on average, file at a point halfway between T0 and T1, and exculpatory evidence with probative power of 10 per cent will be lost. Even in this respect, however, the price model may be more favourable than the discovery rule. While plaintiffs under a narrowly construed discovery rule may, on average, file earlier than they would under the price model, a broadly construed discovery rule allows plaintiffs to delay filing all the way until T1.

It is possible, moreover, to modify the price model slightly so that it consistently maximizes evidence. Under the version of the model described above, the plaintiff is indifferent as to when to file her claim, since T0, T1, and all the times in between promise the same number of statistical dollars. If, however, the plaintiff must not merely compensate but slightly overcompensate the defendant for any loss of exculpatory evidence, then it is in her interest to bring her claim before any such loss occurs. Take again the example described above. Assume that the plaintiff is claiming $100 in damages. The correct allocation of probabilities is 70 per cent to 30 per cent in favour of the defendant. The plaintiff, then, should properly have thirty statistical dollars (100 × .30). As a result of the plaintiff’s decision to delay until T1, exculpatory evidence is lost that reduces the defendant’s odds to 50 per cent and improves the plaintiff’s odds to 50 per cent. The unmodified price model, as explained above, reduces the plaintiff’s collection right to $60, so that she still has only thirty statistical dollars (60 × .50) and thus gains nothing from the delay. But the delay also causes the plaintiff no loss. If, however, we add a peppercorn – say, $1 – to the proportional reduction of the collection right, so that it falls to $29, then the plaintiff has an incentive to file her claim before any decay in the exculpatory evidence occurs. Under the
peppercorn price model, then, the plaintiff will file before any of the defendant’s evidence is lost. After the plaintiff files, the defendant will request a continuance, since time is expected to bring to light more exculpatory than inculpatory evidence. By T1, all the evidence on both sides will have been uncovered and made available.

2 Scenario 2

In this scenario, the passage of time, rather than bringing to light more exculpatory than inculpatory evidence, instead favours the plaintiff. More inculpatory than exculpatory evidence is uncovered between T0 and T1. Otherwise, the scenarios are identical. The discovery and price models produce essentially the same results as in the first scenario. Under the discovery rule, the plaintiff will file at the latest point that the rule allows. Exculpatory evidence will decay until that point, and the plaintiff will enjoy the consequent windfall. Under the peppercorn price model, the plaintiff will again file at T0 to avoid the penalty triggered by the defendant’s evidentiary loss. Here, however, since the passage of time produces more evidence for the plaintiff, she rather than the defendant will request a continuance.

B LATE DISCOVERY

1 The under- and overprotectiveness of the discovery rule

The analysis to this point has demonstrated that in cases where the plaintiff becomes aware of her injury immediately after it occurs, the price model, supplemented by the peppercorn penalty, both maximizes evidence and avoids windfalls. This section carries this analysis to the second situation in which litigation may arise, namely, late discovery. In late discovery situations, there is some stretch of time between the moment of injury and the moment at which the plaintiff becomes aware that litigation is possible. From this second moment onward, the analysis in section A is applicable: The current regime of temporal limitation permits the plaintiff to postpone filing until the end of the statutory period, while the peppercorn price model discourages her from doing so. The following discussion focuses on the evidentiary changes that may occur before the plaintiff becomes aware that she may sue.

Our analysis will show that because the discovery rule focuses exclusively on the reasonableness of the plaintiff’s delay, it generates two problems in the late discovery situation. First, the discovery rule frequently fails to protect defendants: When the plaintiff’s delay is reasonable, the rule permits her to proceed with her suit and collect full damages, even though her delay may have wreaked evidentiary harm upon the defendant. Second, the discovery rule also, at times, overprotects defendants. When the plaintiff’s delay is unreasonable, the discovery rule bars her
suit from proceeding, even though her opponent's evidentiary situation may not have been in any way compromised by the delay.

Both the plaintiff and the defendant may suffer evidentiary harm during the period between the moment of injury and the moment at which the plaintiff becomes aware that she can sue. Where faultless tardy plaintiffs lose evidence, the law already adequately protects them. Typically, this protection comes in the form of favourable factual presumptions. If the plaintiff proves that some piece of evidence has become unattainable, courts assume that it would have supported the plaintiff's claim and place upon the defendant the burden of showing otherwise. The more likely it is that the missing evidence would have buttressed the plaintiff's account, the heavier the defendant's burden to demonstrate that it would not have. In other words, where the blameless plaintiff discovers her injury after the erosion of some inculpatory evidence, the law responds by reducing, in a roughly proportional way, the plaintiff's burden of proof.

One reason that courts are sympathetic to plaintiffs' loss of evidence in late discovery cases is that defendants in such cases are often parties such as doctors, manufacturers, and state agencies that are statutorily obliged to document their activities. If, over the course of time, they fail to maintain records, the violation of their statutory duty gives rise to a presumption against them with respect to the missing information.18

Even when there is no explicit legal duty to maintain evidence, courts make factual presumptions against the defendant when they suspect that the faultless plaintiff may have suffered evidentiary damage. Once the plaintiff establishes the defendant's negligence, it becomes the defendant's duty to prove the absence of damages or causation. For example, in Half v. Lone Palm Hotel,19 the plaintiff sued the defendant motel operators for the wrongful deaths of her son and husband, who had drowned in the motel's pool. The defendants had negligently failed to provide lifeguard supervision. The court presumed that the accident had been caused by the defendants' negligence and put the onus on the defendants to prove otherwise. The court explained that 'the absence of such a lifeguard in the instant case ... not only stripped decedents of a significant degree of protection to which they were entitled, but also deprived the present plaintiffs of a means of definitively establishing the facts leading to the drownings.'20

18 See, e.g., Kronich v. United States, 150 F. 3d 112 (2d Cir. 1998) (ruling, in a suit against a CIA agent for harms caused to the plaintiff in an LSD experiment, that the discovery rule could be applied to preserve the claim and that the jury could draw an adverse inference against the agent based on the agent's inability to provide relevant documents).
20 Ibid. at 474-5. For more cases and further analysis, see S.A. Spitz, 'From Res Ipsa Loquitur to Diethylstilbestrol: The Unidentifiable Tortfeasor in California' (1990) 65
In *Half*, the uncertainty arose because the defendants breached a legal duty in failing to provide a lifeguard. Courts, however, are inclined to interpret an uncertainty against the defendant even where the uncertainty is not directly attributable to the defendant’s wrong but is, instead, a product of the passage of time. In the diethylstilbestrol (DES) cases, for example, the plaintiffs, the injured offspring of pregnant mothers who had ingested the drug, failed to file their claims within the limitation period. Because their injuries had become manifest only years after their mothers had used the drug, the courts applied the discovery rule and accepted the plaintiffs’ late filing. The plaintiffs’ late discovery of their injuries had not only pushed them past the limitation period but also rendered it impossible for them to identify which of the defendant drug manufacturers had produced the particular doses that each mother had taken. The courts, acknowledging the evidentiary harm that the passage of time had inflicted upon the plaintiffs, assumed that each of the defendants had manufactured the drug ingested by each mother and required each defendant to show that it had not produced the drug in a given case.21

Finally, the rule of *res ipsa loquitur* pushes this technique one step further, enabling courts to interpret lacunae in favour of the plaintiff not only when the defendant’s wrong is unrelated to the lacunae but even when no wrong on the defendant’s part has been shown. Although typically the plaintiff must prove her claim by concrete evidence, when no such evidence exists, *res ipsa loquitur* allows the court to infer from the general facts of the case to the liability of the defendant.22 In late discov-

Ind.L.J. 591 at 604-5. Another illustrative example is *Bigelow v. RKO Radio Pictures*, 327 U.S. 251 (1946). In this case, the defendants, distributors of feature films, conspired to provide films to some movie theatres on better terms than those offered to the theatres of the plaintiffs. The defendants contested that the plaintiffs had not proven that the defendants’ conspiracy had inflicted actual economic harm. The majority found that the plaintiffs’ inability to prove harm was due to the defendants’ wrongful behaviour. Despite the fact that the defendants had no duty to maintain evidence, the court concluded that, since ‘[t]he most elementary conceptions of justice and public policy require that the wrongdoer shall bear the risk of the uncertainty which his own wrong has created,’ the plaintiffs were not obligated to establish their harms with certainty.

Ibid. at 580.


22 See Porat & Stein, *Tort Liability*, supra note 13 at 87 (‘Under the [general] principle, naked statistical evidence that merely typifies the case at hand as belonging to a general negligence category will not open the defendant’s pockets to the plaintiff; however, it might do so when the res ipsa loquitur presumption enters the scene’).
ery cases in which the passage of time erodes direct inculpatory evidence, the doctrine protects the plaintiff by functionally shifting some of her burden to the defendant.23

These mechanisms by which the law converts evidentiary harm to the plaintiff into factual presumptions in her favor ensures that the loss of inculpatory evidence will not legally impair the plaintiff. By lightening the plaintiff’s burden of proof, the law provides her with much the same protection that the price rule provides the defendant. Just as the price rule lowers the plaintiff’s potential winnings to offset the loss of exculpatory evidence, so these presumption-shifting rules offset the plaintiff’s loss of evidence by relaxing her evidentiary obligations.24

Defendants in late discovery cases do not enjoy the same sort of protection. Given this imbalance, the discovery rule’s failure to adjust for defendants’ evidentiary losses in late discovery cases exposes them to unfair evidentiary risk. The plaintiff, herself shielded from the consequences of decay in inculpatory evidence, is allowed to take advantage of the defendant’s evidentiary damage. The price model, in contrast, restores the balance between the litigants. By reducing the plaintiff’s potential reward in proportion to the defendant’s evidentiary damage, it balances the advantage that elements of the law outside temporal limitation provide the plaintiff. Thus the price model adequately avoids under-protectiveness of defendants in late discovery situations.

Overprotectiveness results most commonly when, despite the plaintiff’s unreasonably late discovery, the defendant suffers little or no

23 See, e.g., Thompson v. Lista, 420 N.E. 2d 232 (1981) (holding, in a malpractice claim against the defendant doctor for alleged negligent treatment of the plaintiff’s hernia, that the discovery rule rendered the claim timely, and that the jury could find the doctor liable by applying the doctrine of res ipsa loquitur). The extent to which the doctrine presumes against the defendant varies according to the facts of the case and the jurisdiction that applies the doctrine. See Prosser & Keeton, supra note 9 at 258–9 (‘As a general proposition ... the procedural effect of res ipsa may be said to be a matter of the strength of the inference to be drawn, which will vary with the circumstances of the case. [Some jurisdictions] give res ipsa loquitur a greater effect than that of a mere permissible inference for the evidence. They have held that it shifts to the defendant the burden of going forward with the evidence, or creates a presumption which requires a direct verdict for the plaintiff unless the defendant offers sufficient evidence to meet it. [Other jurisdictions] have gone further, and have held that is shifts to the defendant the ultimate burden of proof’).

24 See, e.g., Kronisch, supra note 18 at 127: ‘We must ... attempt to determine whether there is any likelihood that the destroyed evidence would have been of the nature alleged by the party affected by its destruction. This inquiry is part of our attempt to place the innocent party in the same position he would have been in had the evidence not been destroyed by the offending party.’
evidentiary damage. Since the application of the discovery rule turns entirely on the reasonableness of the plaintiff’s delay, the rule bars negligently dilatory plaintiffs even when their negligence causes no evidentiary damage. The defendant in such cases receives a windfall equivalent to the plaintiff’s alleged damages multiplied by the probability that the plaintiff would have won had the litigation occurred.25 The price model, in contrast, avoids this windfall by allowing plaintiffs to file at any point.

To this point, we have contrasted the ways in which the price model and the discovery rule deal with evidentiary erosion with respect to a given plaintiff and defendant. Our implicit criterion has been fairness: the price model is superior to the discovery rule because it gives no unfair advantage to either side. The argument for the price model becomes still more compelling when the comparison is extended to the models’ overall economic consequences. Because the price model prevents either litigant from receiving a windfall, it ensures that activities that cause damage will be neither over-deterred nor under-deterred.

Consider, for example, a manufacturer whose product occasionally causes $60 of damage to its users. Ten consumers, alleging that the product caused them such injury, sue after some time has passed from the moment of the alleged injury. Had the plaintiffs filed immediately, when all the evidence was available, they would each have had a 20 per cent chance of winning, and the manufacturer would, over the ten cases, have had to pay $120. Economic efficiency therefore demands that the manufacturer internalize exactly this amount. As a result of the plaintiffs’ delay, however, the manufacturer loses some evidence, and the plaintiff’s chances rise to 30 per cent. Under the price model, since the plaintiffs’ chances have risen by three halves (3/2), their collection right must fall to two-thirds (2/3), from $60 to $40, so that, though the manufacturer must now pay in three rather than two cases, he will still pay only the desired $120 (3 × $40). Under the discovery rule, in contrast, the manufacturer may end up paying anywhere from $0 to $180 (3 × $60), depending on how many of the ten plaintiffs were reasonable in their delay. If none was reasonable, then the discovery rule bars all of their claims. The manufacturer pays nothing and is therefore under-deterred. If all were reasonable, then all may sue. The manufacturer will lose in three of the ten cases and pay a total of $180. Here he is over-deterred.

25 Cases of overprotectiveness may occur even in the scenario of immediate discovery if the plaintiff, despite her immediate discovery, delays past the end of the statutory period. Such delay is unlikely, however. In any event, our analysis of the overprotectiveness problem in the context of late discovery applies unchanged to the immediate discovery context.
This result may be formalized through an analysis of the aggregate error costs generated by each rule. Error costs occur when an innocent defendant pays or an aggrieved plaintiff fails to receive payment. While neither the price nor the sanction model systematically minimizes error costs, the former distributes error costs equally between the litigants, so that the uncertainty to which evidentiary decay gives rise does not allow either party to externalize the costs of its activity. This proposition is best demonstrated by building upon the above example.

Under the above facts, and assuming that all the plaintiffs were reasonable in delaying, the discovery rule generates $60 in error costs. Three of the ten plaintiffs win, whereas, had there been no evidentiary decay, only two of them would have won. There is thus one incorrect transfer of $60. The price model also permits three plaintiffs to win, but each collects only $40. Two of the three victorious plaintiffs would have won even had there been no evidentiary decay and are therefore each entitled to $60. The shortfall in these two cases thus generates $40 (2 × $20) in error costs. The third victorious plaintiff should not have won. Her victory therefore adds another $40 in error costs. The price model thus generates a total of $80 in error costs, $20 more than the discovery rule.

Under a slightly modified version of the above example, the discovery rule produces larger error costs than the price model. Suppose that the plaintiffs' delay causes the manufacturer more severe evidentiary damage, so that the plaintiffs' chances of winning rise from 20 per cent to 60 per cent. Under the discovery rule, assuming again that the plaintiffs did not delay unreasonably, six rather than two plaintiffs win. The four incorrect decisions generate error costs of $240. The price model responds to the evidentiary decay by proportionally reducing the plaintiffs' collection right from $60 to $20. As under the discovery rule, six of the ten plaintiffs win, but each collects only $20. Two of the six rewards are under-compensatory: the plaintiffs in these cases would have won even in the absence of evidentiary decay and are therefore each entitled to $60. Error costs from these two cases amount to $80 (2 × $40). The other four victorious plaintiffs won only because of the evidentiary decay. Error costs from these cases are, in total, $80 (4 × $20). The price model thus produces, in the end, $160 in error costs, $80 less than the discovery rule.

The key difference between the two examples is the extent of the evidentiary damage caused by the delay. In the first case, the evidentiary

---

decay caused the plaintiffs' chances of winning to rise from 20 per cent to 30 per cent, and the discovery rule produced smaller error costs. In the second case, the plaintiffs' chances rose from 20 per cent to 60 per cent, and the price model proved superior. These two cases are instances of a more general rule: When the plaintiff's chances of victory less than double, the discovery rule generates smaller error costs. When her chances more than double, the discovery rule yields larger error costs. Therefore neither model systematically minimizes error costs.

From the perspective of economic efficiency, however, the question of minimization is less important than the question of allocation. In each of the above illustrations, the discovery rule forces the defendant manufacturer to bear the full burden of the error costs ($60 in the first case, $240 in the second). The manufacturer is therefore over-deterred. The price model, in contrast, always divides the error costs evenly between the litigants ($40 to each in the first case, $80 to each in the second). The error costs on the two sides offset each other, so that, in the end, the effect is the same as though there were no error costs at all, and the manufacturer is neither over-deterred nor under-deterred.

2 Empirical data

Subsection 1 argued that the discovery rule sometimes under-protects and sometimes overprotects defendants in late discovery cases. Subsection 2 buttresses this argument with empirical data from the United States and Britain.

i Discovery rule under-protectiveness

If, in late discovery cases, plaintiffs indeed receive protection for the erosion in their inculpatory evidence while defendants suffer (uncompensated) evidentiary harm, then plaintiffs should enjoy a higher rate of success in late discovery cases than in otherwise identical cases in which the suit begins shortly after the injury. Results in a wide-ranging survey indicate that the discovery rule does indeed generate a plaintiff bias in late discovery cases. The passage of time appears, in this scenario, to weaken defendants' ability to protect themselves.

27 Were the magnitude of error costs the only consideration, the most desirable rule of temporal limitation would be one that, like the price model, focuses on the defendant's evidentiary position but that, unlike that model, assumes the form of a sanction: The plaintiff's claim would be barred when and only when, as a result of the loss of exculpatory evidence, her chances more than doubled. This rule is not, however, a plausible option, since it creates incentives for the plaintiff to behave opportunistically. Before filing, the rational plaintiff would allow the defendant's evidence to decay to the point that her plaintiff's chances doubled.

28 For a formal mathematical model, see Appendix.
Our survey measured the effect of late discovery in the area of product liability. Product liability was chosen because the category is well defined, homogeneous, and suitably large and because it has been the subject of extensive empirical analysis on the issue of plaintiffs’ success rates. We utilized the most comprehensive database of products liability cases, composed of all published federal and state decisions in the area of product liability for the years 1979 to 1999. We divided these cases into two groups. The first group included all product liability cases that commenced within the prescribed statutory period. It thus excluded cases in which the court implemented the discovery rule and cases in which the plaintiff’s claim was time barred. The second group, derived from the same database, included all product liability cases in which late discovery led the courts to implement the discovery rule. During the period between injury and discovery, both the plaintiff and the defendant were susceptible to evidentiary damage. In all other respects than the time of discovery, the two groups of cases did not systematically differ. If late discovery had the same evidentiary consequences for both the plaintiff and the defendant, then the frequencies of plaintiff success in the two groups should have been roughly the same.

29 The Product Liability Reporter (PLR) is a case-reporting service published biweekly by the Commerce Clearing House (CCH). For an account of the quality and comprehensiveness of the collection, see J.A. Henderson & T. Eisenberg, ‘The Quiet Revolution in Products Liability: An Empirical Study of Legal Change’ (1990) 37 U.C.L.A. Law Rev. 479 at 500 (noting that over the same period, the PLR identified substantially more cases than did a similar service of Westlaw); T. Eisenberg & J.A. Henderson, ‘Inside the Quiet Revolution in Products Liability’ (1992) 39 U.C.L.A. Law Rev. 731 [hereinafter ‘Inside the Quiet Revolution’]. For the purposes of this analysis, we followed Eisenberg and Henderson in excluding asbestos cases, in which a special statutory regime produces an abnormally high rate of plaintiff success. See Eisenberg & Henderson, ibid. at 734, explaining that asbestos cases form a separate category governed by asbestos-specific statutes.

30 Properly, cases in which the decision was based not upon the facts but upon some matter of law should also be excluded from the first group. However, the number of such cases appears to be sufficiently small that they may be disregarded.

31 Within each group, cases varied with respect to the nature of the alleged defect and the size of the alleged damage, but, apart from the time of discovery, there was no characteristic that was common to the cases of one group but absent from cases of the other group.

32 This conclusion assumes that plaintiffs are able to predict accurately whether courts will accept their requests to implement the discovery rule. If they cannot make this prediction with certainty, then the frequency of plaintiff success will, sperius sperius, be larger in late discovery cases. To see why, suppose that an individual would not file a claim for $y$ unless she believed she had at least an x chance of winning. Consider two identical claims with $y$ values that, according to their material facts, each present an x chance of success. In one case, however, the claimant learns about the harm only after the statutorily prescribed period. If she is certain that the court will not apply the discovery rule, then she will not bring suit. If she is certain that the court will apply the
To ascertain the rate of plaintiff success in the first group, we relied upon past statistical analyses of the cases in the database. Because the cases in the second group had never been examined as a distinct category, and because the decisions in the database often concerned not the final judgment but, rather, the question of whether or not to implement the discovery rule, it was necessary to inquire of the attorneys in these cases about the ultimate outcomes. In a short questionnaire, we asked them to identify the party for whom the court had decided and, if the case had settled, which (if either) party had received the more favourable terms. The results indicate that plaintiffs did much better in cases where courts implemented the discovery rule than in cases initiated within the limitation period. In cases in which the court reached a decision, timely plaintiffs prevailed only 43 per cent of the time, while dilatory plaintiffs discovery rule, she will bring suit. If she is uncertain – if, in other words, the odds that the court will agree that her delay was reasonable, and consequently implement the discovery rule, are smaller than 100 per cent – then she will not file suit, since her total chance of victory is less than x. In other words, if the plaintiff’s success depends on the implementation of the discovery rule and the odds of its implementation are smaller than 100%, her chance of victory must be less than it would otherwise have had to be in order for her to file suit. If, for example, there were only a one in two chance that the court would agree to apply the discovery rule, a plaintiff would file suit for $5 only if her chance of winning were at least 2x. The consequence, in this case, would be that the frequency of plaintiff success in proceedings in which the courts apply the discovery rule (group 2) would be twice as high as in similar timely cases (group 1). The cases in the PLR suggest, however, that this consideration may be ignored, since plaintiffs are, in fact, capable of accurately predicting whether the court will implement the discovery rule. Courts reject requests to apply the discovery rule in only one case out of seven, and even this figure probably understates the extent to which plaintiffs’ requests are successful. Since the statute of limitation is an affirmative defence that the defendant must raise, there are presumably cases in which the plaintiff’s delay was so clearly reasonable that the defendant does not raise a limitation defence at all. In such cases, the plaintiff has, functionally, correctly predicted the court’s willingness to implement the discovery rule. Statistical analysis fails to account for these cases because there is no explicit ruling on the discovery rule.

33 In calculating the plaintiffs’ frequency of success in timely products liability cases for the years 1979 to 1989, we followed the conclusions of Eisenberg and Henderson’s research (which also used the pool of PLR cases) as summarized in ‘Inside the Quiet Revolution,’ supra note 29 at 740. For the years 1990 to 1999, we relied upon a more recent study; see J. McIver, ed., Products Liability: Verdicts, Settlements and Statistical Analysis (Horsham, PA: LRP Publications, 2000) [hereinafter Products].

34 When the two attorneys agreed that the settlement favoured a particular party, we counted the case as a victory for that party. When the attorneys disagreed, or when they agreed that the settlement favoured neither, we counted the case as a draw.

35 Between 1979 and 1989, plaintiffs’ rate of success averaged approximately 43 per cent (dropping from 56 per cent in 1979 to 31 per cent in 1989). See Eisenberg & Henderson, ‘Inside the Quiet Revolution,’ supra note 29. From 1990 to 1999, plaintiffs’ rate of success averaged approximately 42 per cent (rising from 39 per cent in 1990 to
won in 78 per cent of the cases.\textsuperscript{34} In cases involving settlement, too, the discovery rule appears to have favoured plaintiffs. The attorneys' responses indicated that only 10 per cent of settlements in the late discovery group favoured the defendant.\textsuperscript{35} Although data regarding settlements in the first group of cases are unavailable, defendants' willingness to accept unfavourable settlements in late discovery cases does at least support the finding that the litigation prospects for late discovery plaintiffs are especially good.

46 per cent in 1999). See McIlvreen, \textit{Products}, supra note 33. Plaintiffs' average rate of success over the entire period of the survey (1979 to 1999) was approximately 43 per cent. Plaintiffs' low rate of success in product liability cases is supported by many other surveys. See, eg., R.W. Wright, \textit{Once More into the Bramble Bush: Duty, Causal Contribution, and the Extent of Legal Responsibility} (2001) 54 Vand.L.Rev. 1071 at note 55, who observes that '[s]tudies have uniformly found that, while juries find in favor of plaintiffs in around 50% of all tort cases (compared to around 60% of contract cases), they are much more likely to find in favor of defendants in the two most controversial and publicized areas of tort liability: medical malpractice and product liability.' See also D.J. Devine et al., \textit{Jury Decision Making: 45 Years of Empirical Research on Deliberating Groups} (2001) 7 Psychol.Pub.Pol'y & L. 622 at 708 (Almost every major study on the topic has found that plaintiff success rates vary considerably as a function of various factors, particularly case type.... Specifically, plaintiff success rates tend to be highest in automobile negligence and contract-related cases, hovering close to the 60% mark for both case types. On the other hand, plaintiff success rates are somewhat lower in high-stakes cases involving product liability (40%) and medical malpractice (30%).'); T. Eisenberg et al., \textit{Federal Product Liability Litigation Reform: Recent Developments and Statistics Litigation Outcomes in State and Federal Courts: A Statistical Portrait} (1996) 19 Seattle U.L.Rev. 433 at 436 (finding in a 1991-1992 survey that '[i]n both state and federal courts, noticeably lower success rates exist in the two most discussed areas of modern tort law: products liability and medical malpractice. In medical malpractice jury trials, plaintiff success rates in state court are 30 percent and in federal court are 26 percent. In products liability cases, success rates are 40 percent in state court and 37 percent in federal court'); T.A. Eaton & S.M. Talarico, \textit{A Profile of Tort Litigation in Georgia and Reflections on Tort Reform} (1996) 50 Ga.L.Rev. 627 at 663–4 ('The Georgia data on outcomes of jury trials again parallel those contained in recent studies of tort trials in other jurisdictions. The recent BJS study of jury verdicts in large counties examined over 9400 tort jury trials. These data reflect an overall plaintiff success of 49.9%, with plaintiff prevailing in a higher percentage of automobile (60%) than premises liability (49%), products liability (40%), or medical malpractice (30%) jury trials.').

\textsuperscript{34} We identified 122 cases that involved the application of the discovery rule. Of these 122 cases, thirty-two (26 per cent) reached verdicts and thirty (25 per cent) were settled. Of the remaining sixty cases (49 per cent), some are still pending, while in others, the advocates did not provide sufficiently clear responses. Of the thirty-two cases that were decided, twenty-five (78 per cent) were for the plaintiff and seven (22 per cent) for the defendant.

\textsuperscript{35} In fifteen of the thirty settled cases (50 per cent), the litigants’ attorneys agreed that the settlement had been a victory for plaintiff. In three (10 per cent) they agreed that the defendant had won. In twelve (40 per cent), the attorneys either considered the settlement a victory for neither litigant or disagreed about who had won.
The results of this survey are confirmed by a similar study of medical malpractice suits. In an analysis of nationwide data collected by the National Association of Insurance Commissioners and the US General Accounting Office, researchers discovered a significant positive correlation between the maximum delay permitted by the discovery rule and the rate of plaintiff success. Claims brought in states that cap the discovery period at ten years have a 10 per cent higher probability of success than those brought in states that limit the discovery period to two years.38

ii Discovery rule overprotectiveness

The discovery rule is overprotective of defendants when it bars unreasonably late plaintiffs even though, were all evidence for both sides available, the plaintiff would have won. This windfall to the defendant is especially likely to occur in cases where the defendant suffers little or no evidentiary harm from the plaintiff’s unreasonable delay. Empirical data from England suggest that the extent of the windfall is considerable.

As in the United States, statutes of limitation for personal injury cases in England include a general discovery rule that tolls the statutory period for reasonably tardy plaintiffs.39 The English legislature, however, also grants courts broad discretionary power to accept any untimely claim, even when the plaintiff’s delay was negligent.40 The English statute instructs judges that they should not accept an untimely claim if, *inter alia*, the delay caused evidentiary harm to the defendant.41 The implication is that, in cases in which English courts apply their discretionary power, the

38 Sloan et al., ‘Effects of Tort Reform,’ supra note 5.
39 The Limitation Act 1980 (U.K.), 1980, c. 58, s. 11(4), states that in personal injury cases, the ‘period applicable is three years from: (a) the date in which the cause of action accrued; or (b) the date of knowledge (if later) of the person injured.’ A similar rule is stated in s. 11A for ‘defective products’ cases.
40 Ibid. at s. 33 (1): ‘If it appears to the court that it would be equitable to allow an action to proceed having regard to the degree to which – (a) the provisions of section 11 or 11A or 12 of this Act [the statutes of limitation for personal liability] prejudice the plaintiff or any person who he represents, ... the court may direct that those provisions shall not apply to the action.’ Similarly, the Defamation Act 1996 (U.K.), 1996, s. 5(4), grants the English courts discretionary power to disapply statutes of limitation in claims of defamation or malicious falsehood. Three other common law jurisdictions also grant courts discretion in deciding whether to implement statutes of limitations: Western Australia, Manitoba, and Nova Scotia. For a brief comparison between the English limitation statute and the limitation statutes in each of these jurisdictions, see Law Commission for England & Wales, Limitation of Actions, Consultation Paper No. 151 at 321 (1998).
41 Ibid. at s. 33(3): ‘In acting under this section the court shall have regard to all the circumstances of the case and in particular to ... (b) the extent to which, having regard to the delay, the evidence adduced or likely to be adduced by the plaintiff or the defendant is or is likely to be less cogent than if the action had been brought within the time allowed by section 11, by section 11A or (as the case may be) by section 12.’
discovery rule, standing alone, would have barred the plaintiff, despite the fact that her delay caused no evidentiary harm to the defendant. Empirical analyses of English case law have identified a significant number of cases in which an unreasonably late plaintiff was permitted to file and proceeded to win. In each such case, the discovery rule, without the English modification, would have granted a windfall to the defendant.

IV Manipulation of the discovery rule

Evidence for the inadequacy of the discovery rule, and for the claim that its inadequacy lies principally in its failure to account for the defendant’s evidentiary position, arises also from an examination of the complicated jurisprudence that has accreted to the discovery rule in American courts. In response to the discovery rule’s pro-plaintiff bias in late discovery situations, some courts have developed a set of doctrines designed to protect defendants by limiting the application of the rule. Part IV, section A, details these doctrines in order to demonstrate the inadequacy of the


The recommendations of the Law Commission that examined the English statutes of limitation provide further evidence of the discovery rule’s overprotectiveness. In its initial report, the Commission advised abolishing the discretionary power of the courts to disapply the limitation period. In a revised proposal, however, the Commission, responding to vociferous objection from consultees, reversed its position and advocated the retention of section 33. While section 33 discretion produces uncertainty, the Commission concluded that this cost was outweighed by the section’s flexibility, which makes it possible for courts to avoid the injustice involved in barring dilatory plaintiffs whose delays, while unreasonable, cause no harm to the defendant. See Law Commission, Limitation of Actions (report no. 270) at 90–91 (2001), online: Law Commission < http://www.lawcom.gov.uk/library/lc270/lc270.pdf> (explaining the reasons for the change in the Committee’s recommendation).
discovery rule. This section is built upon the four basic elements that the plaintiff must establish to file a tort claim: breach of duty, the identity of the injurer, harm, and causation. In each attempt to narrow the discovery rule, the courts have proposed to exclude a different element from the basic information necessary to trigger the statutory clock. As we will show, none of these exceptions is a sound application of the fault principles underlying the discovery rule and none has, in practice, won wide or lasting acceptance. The logical poverty and minimal success of these exceptions suggest that they are not natural boundaries of the discovery rule but, rather, impositions upon it designed to correct its flaws.

The courts that have applied these exceptions have not, in general, directly associated them with the discovery rule's failure to account for the effect of the plaintiff's delay upon the defendant's evidentiary position. Section B argues for such a link by providing other instances of manipulation of the discovery rule that clearly respond to the issue of evidentiary loss.

A EXCEPTIONS TO THE DISCOVERY RULE

1 Breach of duty

In *United States v. Kubrick*, 44 a veteran brought suit under the Federal Tort Claims Act (FTCA) 44 against a Veterans' Administration hospital to recover for a hearing loss allegedly caused by the hospital's negligent treatment. Kubrick, who had been treated for an infection in his leg, was aware of his ear injury and of the factual nexus between this injury and the hospital's conduct. However, his doctors had failed to recognize the hospital's negligence prior to the end of the limitation period. In light of his continuous consultation with his doctors and the persistent refusal of the Veterans' Administration to admit its responsibility, 45 Kubrick's late discovery of the hospital's negligence was determined to have been blameless. 46 However, the Supreme Court, although previously willing to apply the discovery rule in FTCA cases, here found Kubrick's claim time barred. For purposes of the statute of limitation, explained the Court, one must distinguish between ignorance regarding the liability of the defendant and ignorance with respect to the injury or its cause. The statutory period would be tolled only if the plaintiff had been reasonably unaware of the harm or of the factual connection between the defen-

45 Ibid. at 114.
46 Ibid. at 127-8. The district and circuit court judgments in favour of Kubrick had followed prior decisions in holding that a malpractice plaintiff under the FTCA must know the legal implications of the facts, as well as the facts themselves, before the limitation period begins to run. See ibid. at 121, note 8.
dant's behaviour and the injury. The discovery rule, however, would not postpone the limitation period if the victim had been reasonably unaware of the injurer's negligence.\footnote{7} In that case, the limitation period would run from the time of injury.\footnote{8}

In cases factually similar to Kubrick, later courts called upon to determine how state statutes of limitation should be implemented generally rejected the distinction drawn by the Supreme Court. Faithful to the idea that blameless plaintiffs should not be time barred, they refused to differentiate between reasonable ignorance with regard to causation or injury and ignorance of the defendant's breach of duty.\footnote{9} These courts, recognizing the theoretical untenability of a position that had claims accruing somewhere between the moment of injury and the moment of full knowledge, instead ruled that the discovery rule tolled the passage of time as long as the victim could not reasonably have known that the defendant had acted wrongfully. Some also pointed to the practical diffi-

\footnote{7} The majority supported its decision with two arguments, both unconvincing. First, the Court noted briefly that the FTCA waived the traditional immunity of the United States and that courts should not take it upon themselves to extend the waiver beyond the point that Congress intended (ibid. at 117–8). This argument loses force in light of the fact that the Court accepted the discovery rule in Urie v. Thompson, supra note 9 at 169–70, and afterward implemented it widely in different federal suits, despite the absence of any explicit Congressional enactment. The Court's second argument advanced a substantive distinction between different types of unawareness. The fact of injury, suggested the Court, 'might be unknown or unknowable until the injury manifests itself; and the facts about causation may be in the control of the putative defendant, unavailable to the plaintiff or at least very difficult to obtain. The prospect is not so bleak for a plaintiff in possession of the critical facts that he has been hurt and who has inflicted the injury. He is no longer at the mercy of the latter. There are others who can tell him if he has been wronged, and he need only ask.' Kubrick, supra note 45 at 122. This argument, however, allows only that ignorance of the fact that a particular act is the cause of the injury is unreasonable, while such ignorance of causation is unreasonable, it should not toll the running of the statutory period.

\footnote{8} It appears that Kubrick offered a convenient opportunity for the Court to restrict the discovery rule and to address the rule's pro-plaintiff's bias. Though the Veterans' Administration had initially refused to compensate Kubrick for the hearing injury, it eventually accepted his claim. The decision against Kubrick was therefore of little importance in the particular case and was likely designed primarily to serve as a precedent for future cases. See Kubrick, supra note 45 at 117, note 5.

\footnote{9} See, e.g., Buschmann v. Georgetown College, 518 A.2d 425 (D.C. App. 1986) (holding that in a suit for dental malpractice and breach of warranty, the cause of action would not accrue until the patient knew or by exercise of reasonable diligence should have known of the injury, of its cause in fact, and of some evidence of wrongdoing). The court in Buschmann listed twelve other states (Arizona, Colorado, Hawaii, Maryland, Michigan, New Hampshire, New Jersey, North Dakota, Oregon, Pennsylvania, Utah, and Washington) that in practice reject Kubrick and continue to toll the limitation period as long as the plaintiff is reasonably ignorant of the injurer's breach of duty. None of these states has, in the two decades since Buschmann, adopted the Kubrick rule. See also Corman, Limitation of Actions, supra note 8 at 147–50.
culty raised by *Kubrick*. Holding that awareness of the injury and of its cause suffices to start the clock would, as one court put it, 'almost force' plaintiffs to file suit against any possible defendants without knowledge of their wrongdoing, simply to avoid the statute of limitation from commencing. This overcaution would congest the courts with potentially frivolous claims.\(^{50}\)

Even some federal courts rejected *Kubrick*. Plaintiffs who knew about the injury and were aware of the factual connection between the defendants and the harm, yet reasonably knew nothing of the defendants' negligence, were permitted to take advantage of the discovery rule. While paying lip service to the Supreme Court precedent in *Kubrick*, these courts refused, in practice, to distinguish between reasonable ignorance of different types of legal information.\(^{51}\)

2 *The identity of the defendant*

Courts have also limited the discovery rule by refusing to apply it even when the plaintiff did not know and could not reasonably have known the defendant's identity.\(^{52}\) This reformulation of the discovery rule has occurred in two categories of cases. In the first category, the injured party knows who caused the alleged harm but is reasonably unaware that the injurer acted as the agent of another party. She sues the immediate injurer and discovers the link between the injurer and the principal only after the end of the statutory period. Some courts, notwithstanding their allegiance to the discovery rule, have held that a new claim against the principal is time barred. In one case, for example, the plaintiffs were injured in a car accident by a federal employee and filed their claim in state court. When, after the end of the statutory period, they eventually

\(^{50}\) *Pennaulty Corp. v. Nation*, 550 A. 2d 1155 at 1167 (Md. 1988).

\(^{51}\) In *Wait v. United States*, 611 F. 2d 550 at 555 (5th Cir. 1980), for example, a motorcyclist was injured in a car accident. The injuries and a later infection led to the amputation of his leg. The plaintiff knew that his leg had been amputated because of the infection and that the infection had been caused by the treatment of the Veterans' Administration. The plaintiff did not, however, know that the VA's treatment was improper. He filed his claim more than two years after the amputation but fewer than two years from the time he learned about the hospital's negligence. Despite *Kubrick*, the circuit court held that the discovery rule tolled the statutory clock until the plaintiff knew about the 'specific acts of negligence causing his injury.' See also *Warr v. United States*, 626 F. 2d 1278 (5th Cir. 1980) (effectively rejecting *Kubrick*); C.M. Wagner, *United States v. Kubrick: Scope and Application* (1988) 120 Mil.L.Rev. 159 at 187 (discussing the connection between *Warr* and *Kubrick*).

\(^{52}\) *Kubrick* itself seems to assume that the knowledge of the identity of the defendant is a prerequisite for starting the clock, as it states that the plaintiff must know 'the critical facts that he has been hurt and who has inflicted the injury.' *Kubrick*, supra note 45 at 122 (emphasis added). The disparity between the different cases trying to narrow the discovery rule is probably another indication of their shaky logic.
learned the identity of the driver's employer, they attempted but were denied the opportunity to sue the federal government. Their reasonable ignorance of the injurer's status was perceived to be immaterial, despite the general application of the discovery rule.\textsuperscript{53}

In the second category, the plaintiff initially sues a party that she mistakenly thinks is responsible for the injury, or, mistakenly thinking that the injury was not the result of human agency, does not sue at all. Later, upon discovering the identity of the true injurer, the plaintiff seeks, respectively, to modify her original complaint or to file suit. In such situations, some courts, generally supportive of the discovery rule, have refused to permit the victims' new claims, arguing that the statutory period begins to run when the plaintiff becomes aware of the direct cause of harm, even if she is still ignorant of the injurer's identity.

In Zelaznik v. United States, for example, an illegal alien murdered the plaintiffs' son. After the murder, the plaintiffs investigated the murderer's background and learned that a state psychiatric hospital had released him shortly before the crime. Only eight years later, however, did they discover that he had unsuccessfully tried to turn himself in to the Immigration and Naturalization Service. The parents' negligence claim against the INS was rejected as time barred. The court explained that the parents knew who had killed their son at the time of his death. This knowledge of the immediate physical cause of the harm sufficed to initiate the statutory clock. '[T]he fact that a reasonably diligent investigation would not have discovered the defendant's [the INS's] involvement is no longer relevant for the purposes of accrual of the statute of limitations.'\textsuperscript{54} Likewise, in a medical malpractice case where the plaintiff learned that the doctor who had performed her surgery was not the one

\textsuperscript{53} See, e.g., Wilkinson v. United States, 677 F. 2d 996 (4th Cir. 1982), cert. denied, 459 U.S. 906 (1982). Wilkinson, a pedestrian, was run over by a soldier. While Wilkinson was aware that the defendant was a federal employee, he filed his suit in state court on the assumption that the defendant had not been acting in the scope of his employment. When, after the end of the statutory period, it was determined that the soldier's driving had in fact been within the scope of his employment, Wilkinson attempted to bring a new claim against the United States. He argued for application of the discovery rule on the grounds that he could not have known that the soldier had been acting within the scope of his employment. Butzner J., in dissent, agreed that Wilkinson's new claim accrued only 'when he or his attorney knew, or in the exercise of reasonable care should have known, that his injury was caused by a government employee acting within the scope of his employment.' Ibid. at 1004. The majority, however, refused to apply the discovery rule. In Gould v. United States (Dep't. of Health & Human Services), 905 F. 2d 738 (4th Cir. 1990), the plaintiff, the widow of a patient who had been the victim of malpractice, was not even aware that the defendants, the treating physicians, were federal workers. Nevertheless, the court, after a rehearing, refused to apply the discovery rule.

\textsuperscript{54} 770 F. 2d 20 (3d Cir. 1985), cert. denied, 475 U.S. 1008 (1986).
whom she had approved, the court rejected her request to implement the discovery rule and to allow a tardy claim against the doctor who performed the surgery. The court explained that

[the] plaintiff's contention in the present case would extend the discovery rule as applied to medical malpractice cases to mean that the cause of action accrues when the person injured learns, or reasonably should have learned, of the identity of the person responsible for his injury, even though he earlier knew of the injury itself.... [W]e are not convinced that the logic of the rule warrants such an extension by us.\textsuperscript{55}

As in the case of Kubrick, these courts' narrow construal of the discovery rule is hardly reconcilable with the fault principle that underlies the rule. Later courts, unwilling to follow this precedent, have generally allowed plaintiffs reasonably ignorant of the defendant's identity to file their tardy claims.\textsuperscript{56}

3 The extent of the harm

Some courts have proposed limiting the discovery rule by excluding the full magnitude of the harm from the knowledge required to trigger the

55 Guembard v. Jabby, 381 N.E. 2d 1164 at 1167 (III. 1978). In Gibson v. United States, 781 F. 2d 1354 (9th Cir. 1986), the plaintiffs brought suit seeking compensatory and punitive damages for an alleged conspiracy involving the burning of their garage. The plaintiffs claimed that their cause of action against the government began to run not from the time of the fire but from the time they discovered that an FBI agent had hatched the plot to burn their garage. The court rejected their suit: 'If plaintiffs would have us stretch the boundaries of the Kubrick decision so delay accrual of a federal tort claim until plaintiff knows or has reason to know of the culpability of federal agents. Language in Kubrick, emphasizing the strategic importance to the litigant of knowing whom to sue, supports plaintiffs' proposed construction. However, binding circuit precedent forecloses us from considering such an extension of Kubrick.' See also Dymon v. United States, 749 F. 2d 484 (9th Cir. 1984), where children of decedents whose car had been swept off the road in a flash flood discovered, after the limitation period had passed, that national park rangers may have been responsible for the negligent road supervision that led to their parents' death. The court found their claim time barred, holding that once a plaintiff knows both of the injury and of its immediate physical cause, her cause of action has accrued.

56 For a case holding that knowledge of the principal-agent relationship is part of the discovery rule, see, e.g., Van Lies v. United States, 542 F. Supp. 862 (N.D.N.Y. 1982) (deciding that where the plaintiff did not know and could not have known that the defendant was a federal employee operating in the scope of his employment, his claim against the government was not time barred). For a decision holding that the identity of the defendant is part of the discovery rule see, e.g., Omer v. International Paint Company, 796 F. 2d 759 at 763 (Wash. 1990) ('The knowledge or imputed knowledge of a particular defendant's identity) is necessary for the plaintiff's cause of action against that defendant to accrue.... [T]his holding finds support in a majority of jurisdictions which have recently faced the question of whether the discovery rule encompasses the identity of the particular party responsible for the plaintiff's injury').
running of the statutory period. Under this approach, the statutory period begins to run from the moment the plaintiff knew or could have known that the defendant’s breach of duty had caused her some harm, even if the initially known injury was so small that it would not have led a reasonable person to sue.

For example, a plaintiff in a car accident case who initially suffered only ‘contusions and abrasions’ but discovered after the end of the statutory period that she had sustained a permanent injury to her neck was denied the benefit of the discovery rule. The court refused to postpone the limitation period to the time that the plaintiff could have known of the later injury and instead counted the limitation period from the time of the accident, even though it was not clear that a reasonable person would have litigated over the insignificant injuries then apparent. The court also refused to treat the two injuries (the contusions and abrasions initially manifest, and the later discovered neck injury) as giving rise to two independent causes of action. Instead, the court saw only a single injury and computed the limitation period from the moment of the accident, when that injury occurred.

Similarly, in Landerson v. Bauer, when a group of plaintiff homeowners claimed that the defendant developers and builders had negligently allowed water to seep into and damage their property, the court held that the ‘statute of limitation begins to run when the injured party knows or reasonably ought to know that some damage has resulted from the wrongful act. That is true although the damage is slight, continues to occur, or additional damage caused by the same wrongful act may result in the future.’

Like other attempted limitations of the discovery rule, this doctrine has failed to win judicial support. When the initial harm is insignificant, courts either ignore it, holding that the limitation period should start only when the known damage is such that a reasonable person would have filed a claim, or differentiate between the damages, finding that

58 For more discussion on division and unification of the damages see Part IV, section B, below.
59 Stephens, supra note 57 at 540.
60 681 P.2d 1316 at 1321 (Wyo. 1984).
61 See, e.g., Martinez-Rivera v. Richardson-Merrill, 105 Cal. App. 3d 316 (1980) [hereinafter Martinez-Rivera] (holding that statutes of limitation do not run until the plaintiff possesses a true cause of action, i.e., until events have developed to a point where the plaintiff is entitled to a legal remedy, not merely a symbolic judgment such as an award of nominal damages); Nicol v. Philip Morris, 201 F. 3d 29 (1st Cir. 2000) (concluding that Rhode Island would not follow a draconian ‘slight injury’ concept); Maria Olsen v. Kansas (State Highway Commission), 679 P. 2d 107 at 173 (Kan. 1984) (‘Our statutes of limitation were not designed to force injured parties into court at the first sight of injury, regardless of how slight it might be, just because that injury and damages resulting therefrom may be permanent in nature’).
the later harm generates a cause of action distinct from the one arising out of the initial, insignificant injuries.\textsuperscript{52}

4 Causation

Finally, some courts have sought to protect defendants in late discovery cases by excluding causation from the preliminary knowledge the plaintiff must possess to initiate the limitation period. According to this reasoning, if the plaintiff knew or could have known about the harm, and knew or could have known about the defendant’s blameworthy conduct, then the statute of limitation begins to run even if the plaintiff could not possibly have known about the causal link between the two.

The causation issue arose recently in the context of child sex abuse claims filed by the victims long after the end of the relevant limitation periods. The plaintiffs argued for the implementation of the discovery rule on the grounds that they were unable to sue until they had matured and obtained psychological assistance. Several courts decided that, to provide some protection to the defendants and to balance the parties’ interests, it was necessary to distinguish between two types of claims. The discovery rule would be applicable, these courts held, only when the plaintiff had totally repressed all memories of the sexual molestation itself and could not reasonably have known about it. However, if the plaintiff knew that she had been abused, and was unaware only of the nexus between the abuse and her current injuries, the statute of limitation would not be tolled, even if her failure to recognize the nexus had been reasonable.\textsuperscript{63}

The pattern observed above with respect to the previous attempts to restrict the discovery rule recurs in this context. Subsequent courts have generally refused to distinguish between ignorance of the abuse and ignorance of the relationship between the abuse and current injuries.\textsuperscript{64}

\textsuperscript{62} Martinez-Ferrer, ibid. at 326 (advising that courts should split causes of action to avoid injustice when a plaintiff fails to file a timely claim and the damage injury later becomes more severe).

\textsuperscript{63} See, e.g., Johnson v. Johnson, 701 F. Supp. 1563 at 1567 (N.D. Ill. 1988) ("[T]he cases which have been brought to date have fallen into two categories: (1) those where the plaintiff claimed she knew about the sexual assaults at or before majority, but that she was unaware that other physical and psychological problems were caused by the prior sexual abuse; and (2) cases such as this one, where the plaintiff claims due to the trauma of the experience she had no recollection or knowledge of the sexual abuse until shortly before she filed suit. [T]he law in other states is also limited although to date, by statute or case law, states who have addressed the issue do apply the discovery rule at least in type 2 cases"); DeRosa v. Carvell, 242 Cal. Rptr. 568 (1987) (hereinafter DeRosa) and E.W. & D.W. v. D.C.H., 754 P. 2d 817 (Mont. 1988) (hereinafter E.W.) (holding that the discovery rule is not applicable in the first category of cases).

\textsuperscript{64} See, e.g., Dunleas v. Doppes, 924 P. 2d 196 (Hawaii 1996) ("We are persuaded by the reasoning of those courts that, having considered the application of either statutory or
Legislators in some states have enacted statutes reversing decisions that endorsed the distinction. Almost all jurisdictions now either fully implement the discovery rule in child sex abuse cases or firmly reject it.

B EVIDENTIARY LOSS AS A FACTOR

Courts advancing the doctrinal exceptions to the discovery rule that we have described in section A necessarily employ terms associated with the discovery rule: the plaintiff’s knowledge, the reasonableness of the plaintiff’s ignorance, and so on. The logical weakness of these exceptions, and their failure to win general acceptance, suggests that they are in fact the product not of a narrow position on the meaning of reasonableness but of concerns external to the discovery rule. In section B, we show that manipulation of the discovery rule is, in particular, very often actuated by the rule’s failure to take into account the effect of the plaintiff’s delay upon the defendant’s evidentiary position. This section analyses situations in which a court has the freedom to decide, depending on how it characterizes the facts of the case, whether or not the statutory period has run. Since, in these circumstances, the issue is, on its face, not whether or not to extend the statutory period but, rather, how to characterize the facts, courts can and frequently do depart from the doctrinal terms associated with the discovery rule and explicitly invoke the defendant’s evidentiary loss in justifying their rulings. In other cases, this factor, while unmentioned, is, from the facts of the case, of clear importance.

judicially created discovery rules to claims of child sexual assault, have determined that the issue of when a plaintiff discovered, or reasonably should have discovered, that she or he was psychologically injured and that the injury was caused by child sexual assault is a question of fact for the jury’ ) [emphasis added]; Den v. LaBrasse, 588 A. 2d 605 at 607 (R.I. 1991).


66 ‘Fundamental Fairness in Child Sexual Abuse Civil Litigation’ Note (1997) 8 Stan.L. & Policy Rev. 205 at 209 (‘The past decade’s wave of CSA [child sexual abuse] civil litigation did not go unnoticed in state capitals. Twenty-eight legislatures, either in response to unfavorable court decisions or sua sponte, have enacted laws seeking to accommodate CSA plaintiffs unable to file suit within the standard statute of limitations. Twenty-four of these statutes contain some form of the discovery rule. Of these, twenty-two unquestionably allow the CSA plaintiff to bring her claim within a set number of years (varying by state) after she discovered, or reasonably should have discovered, both the fact and cause of her adult psychological injuries. These statutes are notable because... they allow both types of CSA plaintiffs to invoke the discovery rule...’). See also M. Frasc, ‘Tolling the Statute of Limitations in Childhood Sexual Abuse Civil Cases’ (1997) 11 J.Contemp.Legal Issues 45.
1 Division and unification of the defendant’s conduct

Often, the cause of the plaintiff’s alleged harm is not a single act but a sustained activity across a span of time. When a condition continues across a span of time, or when there continually recurs an act that, standing alone, could constitute a separate and complete cause of action, the relevant statute of limitation can be applied, at the extremes, in one of two ways. First, and to the benefit of the defendant, the statutory time can be deemed to run from the initial invasion of the plaintiff’s right. In this case, if the invasion began farther back in time than the limitation period, the plaintiff’s claim is entirely barred, despite the fact that the invasion may have continued until the moment of the suit. Second, and in the plaintiff’s favour, the statutory clock may be made to run only from the time that the defendant ceases his wrongful conduct. Thus, even if the initial infringement occurred long before the trial, as long as the condition or the repeated acts continue to a point that is within the limitation period, the plaintiff is able to collect her entire damages.67

Analysis of the choices that courts make among the different options suggests that the defendant’s evidentiary situation is the decisive factor.68 When the plaintiff’s late filing erodes the defendant’s evidence, the courts take the first approach and bar the plaintiff’s claim. For example, in National Railroad Passenger Corp. v. Krouse,69 the plaintiff, an employee of the defendant, claimed that his employer had negligently assigned him work that could reasonably have been expected to cause physical

67 There is a middle possibility: Any given segment of the continued condition, and any particular instance of the recurrent series of act, may be regarded as a separate cause of action for which suit may be brought within the period beginning with its occurrence. Under this third approach, when the initial infringement occurred farther in the past than the limitation period permits but continued into the current limitation period, the plaintiff can collect some, but not all, of her claim. Harms that occurred within the limitation period are compensable, while the rest are time barred. See ‘Developments in the Law: Statutes of Limitation’ Note (1990) 63 Harv.L.Rev. 1117 at 1205.


69 627 A. 2d 489 (D.C. 1993) [hereinafter Krouse].
damage. Although the plaintiff suffered damage throughout his employment and suspected a nexus between the work and the damage, he filed suit only after ceasing to work for the defendant. The court rejected the plaintiff's plea to postpone the running of the statutory period to the point at which the tortious act ceased and instead barred the claim. The court explained that '[p]roof relevant to reduction of damages which a defendant might be able to produce soon after the time the plaintiff discovers his injury may become difficult to obtain because of the passage of time if presentation of the claim is postponed.' Consideration of the defendant's evidentiary loss here motivated the court to bar a claim that it might otherwise have permitted.

Evidentiary concerns also produced a pro-defendant construal of the continuing tort in *Gomez v. Montana.*71 The plaintiff, a state employee whose work involved painting and maintaining paint equipment, brought a claim against his employer for permanent injuries arising from exposure to paint and chemical fumes. The plaintiff filed his claim only after terminating his employment. Concerned about the defendant's loss of exculpating evidence, the court insisted on applying the discovery rule to bar the plaintiff's claim. Although the court conceded that claims may in some circumstances be deemed to accrue only at the point at which the continuing tort ends, it argued that in this case,

applying a continuing tort theory to determine that Gomez's causes of action accrued — and the statute of limitation began to run — on the date of Gomez's last exposure to the paint and chemical fumes is contrary to the general policy underlying statutes of limitation. The primary purpose of statutes of limitations is the suppression of stale claims which, with the attendant passage of time, inhibits [sic] a party's ability to mount an effective defense.72

Here, again, the court looked to the defendant's evidentiary position in deciding to bar the plaintiff's claim.

70 Ibid. at 498. For a case in which the court was willing to apply the continuing tort doctrine in a negligent assignment case, apparently because it did not expose the defendant to any evidentiary risk, see *Carrigan v. Burlington Northern R.R.,* 612 F. Supp. 665 at 668 (D.C. Minn. 1985).

71 975 P. 2d 1258 (Mont. 1999).

72 Ibid. at 1263. In similar cases in which time running did not threaten to erode the defendant's evidence, the continuing tort doctrine was applied. See, for example, *Rowe v. Pennsylvania,* 264 F. 2d 397 (3rd Cir. 1959), in which a worker sued his employer for traumatic arthritis allegedly caused by a defective air hammer. The plaintiff used the hammer for more than twenty-five years but filed the claim only after resigning from his job. The court determined that the statutory period began to run only with the plaintiff's resignation, and thus that the plaintiff's claim was not barred. The result here suggests price-model thinking: Since the single question in the case was whether the air hammer was defectively designed, and the hammer was still available, litigation did not expose the defendant to unfair evidentiary risks.
When, on the other hand, the plaintiff’s delay does no damage to the defendant, courts construe the sequence of torts as a single action that only ‘occurs’ at the end of the sequence and thereby permit the plaintiff to collect on the entire sequence. In Havens Realty Corp. v. Coleman, for example, four plaintiffs brought action against the defendant, the owner of an apartment complex, alleging that his racial steering practices violated the Fair Housing Act (FHA). The violations underlying the claims of three of the plaintiffs were, under the relevant statute of limitation, already stale; the three tardy plaintiffs offered no explanation for their delay. Nevertheless, the Supreme Court held that none of the claims was time barred. Because the defendant’s actions, as alleged, constituted a ‘continuing violation’ of the FHA, so long as any claim grounded in this violation could be heard, all claims so grounded would likewise be permitted. It appears that the Court reached this conclusion because it believed that the defendant had suffered no evidentiary harm from the delay. Since the defendant had in any case to answer the allegation of the timely claimant, and since the evidence that underlay this allegation also underlay the three untimely claims, permitting the other three claims would not impose an evidentiary burden upon the defendant.

Lack of evidentiary damage also explains the DC Circuit’s decision in Page v. State. In this case, a veteran filed suit alleging that for nearly twenty years the Veterans’ Administration had negligently prescribed inappropriate drugs for him. Although the plaintiff had known for a long time about the defendant’s negligence and the consequent damage, he filed the claim only after completing the drug treatment. Rejecting the defendant’s request that the court apply the discovery rule and bar the plaintiff’s claim, the court found the drug treatment a continuing tort that did not accrue until the tortious conduct had come to a halt. Here, too, the court’s pro-plaintiff analysis of the continuing tort is understandable in light of the defendant’s evidentiary situation. Since the defendant’s behaviour was not at issue, and the dispute involved only a medical question unrelated to the particular facts of the case, the defendant’s ability to respond to the plaintiff could not possibly have been damaged by the passage of time.

73 455 U.S. 363 (1982).
75 For a similar reading of this case see MacAyeal, ‘Discovery Rule,’ supra note 68 at 617: ‘The Havens Realty Court indicated that the rationale behind the continuing violations doctrine is that conduct which repeats itself or that is continuing in nature is subject to ready investigation and confirmation and does not present the staleness problems that statutes of limitation are primarily designed to prevent.’
76 729 F.2d 918 (D.C. Cir. 1984).
77 Ibid. at 923.
Division and unification of the plaintiff's harm

A second circumstance in which the application of the statute of limitation depends upon the court's characterization of the facts occurs when the defendant inflicts a single injury that gives rise to a series of harms at various points after the injury. When the plaintiff sues, the statutory period has already passed from the moment that the injury was inflicted. Some of the harm traceable to the injury has, however, occurred for the first time within the statutory period. By treating each harm as a distinct cause of action accruing at the moment that the harm occurs, or, conversely, by conceptualizing different harms as a unified harm occurring at the moment that the injury was inflicted, courts can, respectively, preserve the plaintiff's claim despite its staleness or block her claim in spite of its timely filing. An examination of courts' rulings again shows a connection between the probable erosion in the defendant's evidence and the characterization of the plaintiff's injuries.

When, despite the plaintiff's delay, the defendant has suffered no evidentiary harm, courts are more willing to construe each harm as a distinct cause of action and thereby preserve claims arising from harms that occurred within the statutory period. In *Martinez-Ferrer v. Richardson-Merrell*, 78 for example, the plaintiff used, in 1960, a drug manufactured by the defendants to treat his cholesterol problems. Shortly thereafter he developed an acute allergic reaction and a severe case of dermatitis, which his doctor attributed to the defendants' drug. The plaintiff did not file suit against the defendants at that time, however. More than fifteen years later, when he developed a cataract that was again traced by his doctor to the defendants' drug, the plaintiff filed suit. Letters sent by the defendants to prescribing doctors at the time the plaintiff had taken the drug showed that the defendants had been aware that use of the drug increased the risk of contracting cataracts. Despite the substantial time that had passed between the initial harm and the litigation, the court refused to find the claim time barred. Rather, the court divided, for temporal limitation purposes, the initial harm (the allergy and dermatitis) from the later harm (the cataract), thus permitting the plaintiff to collect damages for the cataract. In explaining its decision, the court noted that 'the cataract-causing potential of [the drug] became known to defendants in 1960 or 1961. They obviously have had an opportunity to gather evidence on the subject while the facts were as fresh as they could be. In addition, they have had two decades to refine the result of their researches.' 79 As the defendant suffered no evidentiary damage, the court concluded that applying the statute of limitation to bar the plaintiff's claim would cause 'injustice.' 80

---

79 Ibid. at 525.
80 Ibid. at 526. For a similar ruling, see *Anderson v. Sybron Corporation*, 299 S.E. 2d 160 (Ga. 1983). In this case, workers claimed against the manufacturer of a sterilizer machine for
The same rationale probably underlay another court’s verdict in a defamation case in which the defendant published a book accusing the plaintiff, a New York state judge, of collaborating with the Mafia. Although acknowledging the longstanding rule that repetition of the defamation does not constitute a new cause of action, the court was willing, for statute of limitation purposes, to consider each of the two published editions as a distinct harm. The majority argued that the fact that the first edition was hardcover and the second paperback provided sufficient grounds to differentiate them. The court therefore found that the plaintiff’s failure to file a timely claim against the defendant after the publication of the first edition did not bar him from filing a new claim after the issuance of the soft-cover copy. The dissenting judge pointed out, probably correctly, that ‘the majority, focusing on a changed cover format, [had] artificially split a single unified publication into two distinct parts in order to avoid the consequences of the Statute of Limitations.’ The majority opinion is reasonable, however, in light of the evidentiary calculus. Since the defendant was able, despite the passage of time, to establish the truthfulness of the book’s allegations, the court construed the defamation as two distinct harms in order to permit the plaintiff to proceed with his claim on the paperback publication.

v The general shift from binary to dividing rules

Perhaps the most notable and important change in liability law over recent decades has been the general shift from sanction rules (all-or-nothing rules under which the plaintiff either wins or loses) to price rules (rules that allow for the liability to be split). One after another, cataract damages arising from a leak of chemicals from the machine. The workers had earlier suffered other injuries that were traceable to the chemical leak. The court ruled that the statute of limitation would run not from the time at which the workers suffered these other injuries or became aware of the link between these other injuries and the chemical leak, but only from the time they could have known about the connection between the chemical leak and the cataract damage in particular.

82 Ibid. at 438.
83 When, in contrast, the defendant suffers substantial evidentiary harm, courts are less willing to circumvent the statute of limitation by splitting damages. See Kornak, supra note 69 at 458. For more cases forbidding the divisibility of a single cause of action, see Matisheav v. Celotex, 569 F. Supp. 1539 at 1542 (D.N.D. 1983) (holding that a single cause of action cannot be divided under North Dakota law); Connelly v. Paul Ruddy’s Equipment Repair & Service, 200 N.W. 2d 70 at 72 (1972) (rejecting the plaintiff’s request to divide a tort cause of action).
84 G. Calabresi & J.O. Cooper, ‘New Directions in Tort Law’ (1996) 30 Val.U.L.Rev. 859 (‘[I]n the last thirty years, after much pressure from theoretical scholars and others, our tort system has moved from the dominance of all-or-nothing recovery rules toward rules that allow and favor splitting in any number of ways’); G. Calabresi, ‘The Simple Virtues of the Cathedral’ Remarks (1997) 106 Yale L.J. 2201 (recommending as a topic for
common law binary doctrines have been replaced by rules that do not assign liability to one side or the other but, instead, divide the harm between the two sides. The understanding that all-or-nothing rules are often inefficient and morally unsatisfactory has convinced courts and legislatures to adopt legal mechanisms that allocate liability more accurately.85 The law’s general trend toward and increasing experience with dividing rules argues in favour of the proposed move from a sanction-based to a price-based system of temporal limitation.86

The shift from binary to dividing rules has occurred in a variety of legal contexts. To take some examples from the field of tort law. The increasing attention the substitution of ‘all or nothing’ norms with rules that divide harm between the parties).

85 See, e.g., Duncan v. Casna Aircraft, 665 S.W. 2d 414 at 425 (Tex. 1984) (‘Unfairness ... is not the only serious flaw of virtually ignoring plaintiff and third party misconduct in strict products liability actions. The failure to allocate accident costs in proportion to the parties’ relative abilities to prevent or to reduce those costs is economically inefficient’); Kern v. Ash Bodakhem, Ltd., 748 S.W. 2d 95–96 (Tex. 1988) (noting that in a growing number of states ‘[t]he all-or-nothing concepts of contributory negligence and assumption of risk were abandoned in an effort to create a less confusing and more efficient system of loss allocation’).

The difference between binary and dividing norms does not map on to the distinction between rules (decision norms that allow little discretion) and standards (decision norms that give legal decision makers broad latitude.) Comparative negligence, for example, offers no more discretion than contributory negligence. While it generates more flexible outcomes, comparative negligence is much more a rule than a standard, since it prescribes exactly how damages must be allocated. Similarly, while many more outcomes are possible under the price model than under the discovery rule, the price model, because it confines the court to a single parameter – the defendant’s evidentiary loss – is a rule. Indeed, the price model arguably allows courts less discretion, and is therefore more rule-like, than the discovery rule. The arguable modern shift from rules to standards (see D. Kennedy, ‘Form and Substance in Private Law Adjudication’ (1976) 89 Harv.L.Rev. 1685) should therefore not be identified with the move from binary to dividing norms.

86 The argument from historical evolution is especially strong here because the universal application of the discovery rule seems to have been something of an historical mistake. It appears that the rule spilled over from a narrow exception in the traditional regime of temporal limitation. In fraud and intentional concealment cases, the traditional rule was that the statutory period would begin to run not from the moment of the injury but from the day on which the plaintiff learned or could have learned about the facts of the case. See, e.g., Pennwalt Corp. v. Nasia, 550 A. 2d 1155 at 1158 (Md. 1988) (‘The roots of [the discovery rule] can be traced to this court’s early cases in equity involving claims of fraud. In those cases, the discovery rule was applied to bar the running of the statute of limitations’). In this context, the discovery rule generates no problematic windfalls. Since the defendant has committed an intentional harm, he is aware of the possibility of suit, and therefore will most likely retain evidence. Moreover, fraud and intentional concealment are harms that the plaintiff is unlikely to detect immediately, so that any evidentiary loss that the discovery rule does generate for the defendant is probably offset by evidentiary harm to the plaintiff. When the discovery rule spread to other contexts, however, it led to the windfalls described above.
doctrine of contributory negligence dictated that a plaintiff who contributes to the occurrence of the harm can collect nothing, unless she satisfies one of the exceptions to the rule, in which case she is eligible to collect her entire claim. This doctrine was replaced by the more flexible system of comparative negligence. According to the latter, when parties together cause harm by their concurrent negligence, liability is divided rather than attributed to only one of the litigants.87 The courts have similarly modified the rule of assumption of risk, once a binary defence, so that the plaintiff’s consent to engage in the risky activity is no longer decisive. Instead, the courts consider the plaintiff’s approval as only one factor in defining her level of fault. Likewise, strict immunities that previously barred claims against certain defendants (such as charities, the plaintiff’s family members, or federal officials) have been nullified, and courts today incorporate consideration for the special status of these defendants into the remedy judgment. Another example is the replacement of the traditional ‘but for’ standard of causation with doctrines such as market share liability and industry liability, which enable a court to find that the defendant was only a partial cause of the damage.

The shift from binary to dividing rules has occurred also in the context of joint tortfeasors. The traditional common law rule held that where damage was the result of more than one defendant’s negligence, the plaintiff could file her claim against whichever tortfeasor she wished, and the latter had no right to obtain reimbursement from other responsible individuals. Most jurisdictions today, however, apply a contribution regime. Thus, while in the past joint tortfeasors cases typically ended with a verdict against one wrongdoer, who then had to cover all the damages, today most legal systems allow division of liability among all responsible individuals.

Finally, some courts have replaced the traditional preponderance-of-the-evidence standard with a more flexible doctrine that recognizes loss

87 The movement from contributory negligence to comparative negligence occurred gradually. Bridging the binary system of contributory negligence and the dividing system of comparative negligence was a middle period in which the pro-defendant consequences of contributory negligence were softened by such doctrinal exceptions as ‘last clear chance’ and ‘last wrongdoer.’ See, e.g., Calabresi & Cooper, ‘New Directions in Tort Law’ supra note 84 at 872-3 (‘The doctrine of last clear chance, which ameliorated the harshness of the all-or-nothing contributory negligence rule, typically disappears under comparative negligence’). In this middle phase, the system of binary consequences was retained, but courts enjoyed more flexibility in determining which of the two consequences to apply. Our current system of temporal limitation is, likewise, currently in a middle phase. Traditional statutes of limitation offered binary consequences and gave courts little discretion in determining which of the two consequences to apply. The discovery rule preserves the binary consequences but allows courts some flexibility in deciding whether or not to bar a late claim. Moving to the price model would complete this evolutionary process.
of chance as compensable. The loss-of-chance doctrine makes it possible for a plaintiff to collect even when she cannot prove more probably than not that the harm was caused by the defendant’s wrongdoing. For example, a plaintiff who would have had a 10 per cent chance of recovering had her disease been diagnosed on time, and whose doctors negligently failed to discover the illness, can collect 10 per cent of the damages that she suffers, even though there is a 90 per cent chance that a timely diagnosis would not have benefited the plaintiff.

In fact, the proposed price model can be viewed as a combination of two existing price doctrines: comparative negligence and loss of chance. Though the wrongdoing by the plaintiff that entails the application of comparative negligence usually occurs prior to the time of the injury, it can also occur afterward. For example, a plaintiff who unreasonably refuses to submit to a surgery that would improve her condition loses her right to full compensation. Hence, this article’s proposal that the plaintiff’s compensation right be reduced when her filing (which necessarily occurs after the injury) is late conceptually follows existing liability principles. The particular remedy advanced here – proportional reduction of the plaintiff’s collection right according to the evidentiary damage that the defendant suffers – applies the rationale behind the loss-of-a-chance doctrine.

VI. The evidentiary function of temporal limitation

We have demonstrated to this point that the price model responds better than the current sanction models to the problem of evidentiary decay. This demonstration argues for the superiority of the price model only on the assumption that evidentiary decay ought to be the central concern of temporal limitation. Courts and academics have offered a host of other

reasons, however. Statutes of limitation, it is argued, are intended to promote repose, to reduce transaction costs, to protect reliance interests, or to reinforce our status quo bias. Moving from a sanction regime to a price model that does not impose any temporal restriction on filing undermines the ability of temporal limitation to achieve these goals. While the price model, supplemented by the peppercorn penalty, encourages prompt filing, it nevertheless permits the possibility that a plaintiff in a late discovery scenario will file her claim long after the time permitted under sanction rules.

The alternative rationales offered for temporal limitation rely upon the ability of statutes of limitation to establish a definitive end date for a given cause of action. Therefore, while they argue against the price model, they do not provide any reason to favour the discovery rule, since the discovery rule, like the price model, sets no absolute chronological limit to a given claim. Under the discovery rule, a defendant can never know that the threat of a potential suit has passed, since it may turn out that the plaintiff remains unaware of certain essential facts and, thus, that her delay is reasonable. Our project in Part VI is therefore less to establish the superiority of the price model over the discovery rule than to show that temporal limitation rules need not provide absolute chronological boundaries.

The basic argument of Part VI is that other rules than temporal limitation limit the chronological scope of potential liability. These rules


90 See, e.g., Parron v. Hohn, 607 N.W. 2d 8 at 14 (S.D. 2000) ('Statutes of limitation have allowed people, through "legislative grace," to be "freed from the consequences of their actions after a statutory period of time resulting in peace of mind for the individual"); Walker v. Armao Steel, 100 S.Ct. 1978 at 1980 (1980) ('The statute of limitations establishes a deadline after which the defendant may legitimately have peace of mind').

91 See, e.g., R.A. Epstein, Past and Future: The Temporal Dimension in the Law of Property (1986) 64 Wash. U.L.Q. 661 at 678 ('The statute of limitation enhances the marketability of title by shortening the period during which prospective purchasers and lenders (both noted for their squeamishness) need examine the state of the title... The statute of limitation generally avoids these title-clearing costs').

92 This explanation is associated with Oliver Wendell Holmes's short comment on temporal limitation: 'I should suggest that the foundation of the acquisition of rights by lapse of time is to be looked for in the position of the person who gains them, not in that of the loser... It is in the nature of man's mind. A thing which you have enjoyed and used as your own for a long time, whether property or an opinion, takes root in your being and cannot be torn away without your resenting the act and trying to defend yourself, however you came by it. The law can ask no better justification than the deepest instincts of man.' O.W. Holmes, The Path of the Law (1897) 10 Harv.L.Rev. 457 at 477. See also R.C. Ellickson, Bringing Culture and Human Frailty to Rational Actors: A Critique of Classical Law and Economics (1989) 65 Chi.Kent L.Rev. 25 at 39 (explaining adverse possession in terms of the endowment effect).
ensure that shifting to a price model of temporal limitation will not unleash a flood of stale claims. In contrast to these other temporally sensitive rules, statutes of limitation start to run from the injury, the moment at which the plaintiff first has reason to sue. Since the matter of securing and obtaining evidence then becomes most urgent, it is at that point that the legal system can most efficiently intervene to create incentives designed to maximize evidence and avoid distortion. The other temporally sensitive rules either come into play before the harm occurs or operate independently of the harm. A properly constructed system should therefore design statutes of limitation to fulfill evidentiary goals, making use of other rules to keep out stale claims.

An example illustrates the ability of other legal rules to limit the chronological scope of liability. Suppose a customer drops a piece of fruit onto the floor of the grocer’s shop at T1. At T2, a second customer kicks the fruit to another location. At T3, the fruit is squashed and oozes juice. At T4, a third customer slips on the juice and injures herself. If, as in traditional statutes of limitation, there is a defined period within which suit may be brought, then the limitation rule will provide an obvious end date, measured from the time of the injury, to litigation arising from the injury. But other rules also provide rough end dates, measured from pre-injury events. The tort law requirement that the cause of the injury be close in time to the injury means that the first customer will, at some point after T1 (perhaps as early as T2), be free of the threat of litigation. Likewise, insofar as it bars a claim against the first customer, this rule facilitates adjudication by limiting the relevant evidence to more recent, and presumably more easily accessible, facts.

Rules that restrict the chronological scope of liability can be found across all areas of private law. Consider again, for example, proximate cause, the doctrine illustrated by the above hypothetical. Proximate cause works to protect parties that, while causes in fact of the relevant injury, are so distant from the injury that fairness and efficiency require that they be free of liability. Distance in time and remoteness in space are the two main factors that the proximity standard takes into consideration. Thus it acquits parties whose actions took place outside the physical arena in which the harm occurred and, more importantly, those whose actions were committed at some chronological distance from the moment of injury.93

Similarly, the defence of contributory negligence limits the adjudication of stale claims. Since contributory negligence by the plaintiff completely bars her claim, and since the plaintiff’s negligence typically continues up to the moment of the accident, application of the defence

renders it unnecessary to investigate the events preceding the accident. Proof that the plaintiff’s behaviour just prior to the accident constituted negligence is in and of itself sufficient to reject the claim.94

Expanding the grounds of liability can also, counter-intuitively, reduce the chronological scope of a lawsuit. When the plaintiff can prevail by proving several simple facts from the recent past, scrutiny of the more distant past becomes unnecessary. The doctrine of vicarious liability illustrates this point.95 Generally, when the plaintiff is injured by an employee in the scope of his employment, the employer is a more attractive defendant than the employee. The employer, typically possessing ‘deep pockets,’ is not only less likely than the employee to be judgment-proof but also less likely to attract a jury’s sympathy. Without vicarious liability, the plaintiff could not collect from the employer without showing that the employer had been negligent in hiring the injuring employee, or that the employer had given inadequate directions to the employee, or that the employer had failed to offer the necessary training. Under vicarious liability, however, the plaintiff can ignore these matters of the more distant past and, instead, collect from the employer merely by showing that the employee negligently inflicted an injury upon her.

In the law of contract, rules that give legal effect to post-contractual behaviour reduce the relevance of events of the more distant past.96 Instead of looking to the contract itself, a court may refer to the parties’ more recent behaviour. Thus contract law, for example, generally privileges the course of performance over the terms of the contract. If a service provider contracts to provide a weekly service on Fridays, but in fact renders the service on Sundays, and continues to do so for three

94 Ibid. at 1191 (‘[T]he theory that regarded contributory negligence as an absolute bar to recovery had the effect of reducing the length of the causal chains, for except in the rare cases the plaintiff’s negligence occurred at or persisted until the very instant of the accident’). It is not unlikely that the law’s recent abandonment of contributory negligence in favour of comparative negligence is linked to the current trend toward shorter statutes of limitation. Since, once contributory negligence is out of the picture, the chronological scope of liability extends further backward from the point of the injury, we compensate by reducing the chronological scope of liability forward from the point of the injury. For instances of the trend toward shorter limitation periods in common law jurisdictions see D.W. Louisell & H. Williams, *Trial of Medical Malpractice Cases*, looseleaf (Albany, NY: M. Bender, 2000) at ¶13.01 (2); PLR (2000), supra note 29 at 7451; Law Commission, *Limitation of Actions*, supra note 40 at 18; Choh Ha’Hischshrot (Department of Justice, Israel, 1994).

95 Epstein, ‘Temporal Dimension,’ supra note 93 at 1192.

96 See U.C.C. §2-208 at para. 5 (1961) (‘[C]ourse of performance shall be relevant to show a waiver or modification of any term inconsistent with such course of performance’). See, e.g., T.J. Stearnum & Co. v. 81, 195 Barge Filler, 629 F. 2d 335 at 364 (5th Cir. 1980) (‘[P]arties’ course of performance after execution of the contract can operate as a waiver of specific contractual provisions’).
years, the beneficiary of the services will not afterward be permitted to sue for breach of contract. A statute of limitation might bar the claim on the ground that too much time has elapsed since the breach, but contract law provides the service provider the same protection by permitting post-contractual behavior to modify the terms of the contract.

Other contract rules limit the chronological scope of litigation by reducing the relevance of pre-contractual negotiations. Under the common law’s mirror image rule, for example, the last correspondence between the parties defines the contracts. 97 Earlier negotiations are ignored, even though they might be probative of intent. By privileging the final communication, the mirror image rule achieves the same result as a statute of limitation: A suit based on actions too chronologically distant will not be heard. Similarly, the parole evidence rule, which precludes prior written or oral statements from modifying or adding to the terms of a written contract, permits contractual liability to extend no further into the past than the moment at which the contract was integrated. 98

As in torts, expansion of the forms of contractual liability may limit the relevance of the past. Under traditional common law rules, the plaintiff could prove a valid contract only by showing that the parties had exchanged consideration at the time the contract was formed. The emergence of such doctrines as promissory estoppel and quantum meruit allows the plaintiff to establish liability even in the absence of a valid contract, on the basis of the more recent behavior of the parties: the plaintiff’s reliance on the defendant’s promise, in the case of promissory estoppel, and unjust enrichment by the defendant, in the case of quantum meruit. Furthermore, whereas the extent of traditional contract liability is determined by the expectation of the plaintiff at the time the contract was formed, these alternative grounds of liability measure the extent of liability using events of the more recent past: the plaintiff’s expenses, in the case of promissory estoppel, and, in the case of quantum meruit, the defendant’s gain.

Even property law, which is usually reluctant to extinguish rights because of the scarcity and high price of land, employs rules that reduce the need to refer to the past and terminate otherwise valid claims. The

97 See, e.g., Kohler Flour Mill v. Linden, 230 Mass. 119 at 123 (1918) (‘An offer made by one party must be accepted without qualification, or, if any variation from that offer is made by the accepting party, that variation in turn must be unequivocally adopted by the party making the first offer, before a contract can be made. Ordinarily the annexing of a condition to the acceptance of an offer is regarded as the rejection of the original offer and the making of a new counter offer. Such counter offer must be accepted without departure from its terms or there is no contract’). But see U.C.C. §2-207 (1951) (abolishing the common law rule).

98 See, e.g., Thompson v. Libby, 26 N.W. 1 at 2 (Minn. 1885) (‘The rule forbids to add by parol where the writing is silent, as well as to vary where it speaks’).
expiration of servitudes as a result of a change of circumstance provides one example. A legal system that utilizes a long statute of limitation can nevertheless avoid stale claims with respect to servitudes if it makes small changes of circumstance sufficient to invalidate the servitude. The delaying plaintiff’s claim to enforce a right will most likely fail not because she has waited too long, but because a sufficiently significant change of circumstance has occurred in the interim. The recording system for deeds and mortgages also reduces the importance of the past. At common law, cases of competing rights in the same plot of land were governed by the principle of ‘first in time, first in right.’ Under this principle, the legally relevant moment is the point at which the party’s right comes into existence. The recording system instead gives preference to the party that registers its right, regardless of when the right was created. Finally, the rule against perpetuities ensures that no interest in land can come into existence more than twenty-one years after a life in being at the time of the instrument purporting to create such an interest. This rule, too, limits the temporal horizons of property law.

VII Conclusion

We have argued, in essence, that the current models of temporal limitation, in focusing solely upon the plaintiff’s behaviour in filing her claim, are both unfair and inefficient. Since there is no necessary relationship between such behaviour and the evidentiary damage that the passage of time may inflict on the defendant, the plaintiff’s timing should not play a role in determining the justiciability of her claim.

Under current regimes, windfalls may occur in both directions. The plaintiff gains when the limitation rule allows her to file a late claim despite the fact that the defendant has lost evidence. This result arises in two scenarios. First, a plaintiff, depending on the length of the limitation period and the extent to which the court is willing to construe delays beyond that period as reasonable, can strategically withhold her claim.

99 See Restatement (Third) of the Law of Property, Servitudes §7.10 (2000), Modification and Termination of Servitudes because of Changed Conditions: ‘When a change has taken place since the creation of a servitude that makes it impossible as a practical matter to accomplish the purpose for which the servitude was created, a court may modify the servitude to permit the purpose to be accomplished. If modification is not practicable, the court may terminate the servitude.’

100 Thus a party that violates the terms of a servitude wins repose as soon as the circumstances change enough so that enforcement of the servitude is no longer legally possible. Likewise, the rule reduces transaction costs because a buyer need only search for servitudes backward to a point at which the conditions of the neighbourhood substantially changed.

She can thereby erode the defendant’s probative material and shift the probabilities in her favour. Second, in a late discovery situation, a long limitation period or application of the discovery rule allows the plaintiff to file her claim even when the defendant has lost evidence. Since the plaintiff’s evidentiary loss, in contrast, is typically redressed by an assortment of rules outside the temporal limitation regime, she enjoys an unfair advantage in the late discovery situation. In the opposite direction, current forms of temporal limitation produce a windfall for the defendant when they bar a dilatory plaintiff even though the defendant has suffered no evidentiary loss.

In the language of corrective justice, these windfalls are problematic because they unfairly advantage one litigant over another. In efficiency terms, they are undesirable because they induce the wrong litigant to internalize the damage. Unless the improper allocations of liability offset each other – and there is no reason to suppose that they do – the current limitation regime, in the end, either under-deters or over-deters certain categories of behaviour.

The price model, in contrast, avoids such windfalls by yoking temporal limitation to the defendant’s evidentiary harm. Since, under the price model, the plaintiff’s collection right shrinks to match the extent to which the passage of time has reduced the defendant’s chance of winning, the litigants remain with the same number of statistical dollars that they would have had if all the evidence had been presented. In corrective justice terms, therefore, neither party enjoys an unearned advantage. The price model allows the plaintiff to file her claim at any point (thereby eliminating the defendant’s windfall under the discovery rule and the traditional limitation rule) but prevents her from benefiting from the corrosive effect of time’s passage (thus eliminating any windfall to the plaintiff). By avoiding windfalls, the price model also achieves, in the aggregate, optimal deterrence. Because it distributes error costs evenly between the litigants, a given class of actors – manufacturers, doctors, drivers, and so on – internalizes all of and only the damage that its activities actually cause.102

102 One can achieve some of the benefits of the price model, while avoiding some of its administrative costs, by implementing a hybrid of the traditional rule and the price model. In many jurisdictions, the discovery rule, too, occurs in a hybrid with the traditional rule. Typically, the law defines one point before which a claim may be filed regardless of the plaintiff’s reasonableness, and a second point after which the claim is completely barred. The discovery rule comes into play within the period between these two points. It is possible to structure the price model similarly. In, say, the first two years after the alleged damage, the plaintiff would be able to file the claim without having to suffer any reduction in her collection right. Beyond fifteen years no claim would be allowed. For claims filed between two and fifteen years from the alleged damage, the court would evaluate the evidentiary harm to the defendant.
We have defended the argument for a shift to the price model by situating our proposal within the law's larger trend away from sanction rules and toward price rules. We observe, in conclusion, that this theoretically attractive shift is feasible in practice. Under the price model, a court must first evaluate the defendant's evidentiary loss and then determine, counterfactually, the extent to which the lost evidence would have improved the defendant's litigation prospects. The model thus assumes that courts can answer counterfactual questions and that the costs of administering the price model's two-step inquiry do not render it inefficient.

Both assumptions are realistic. Courts' ability to answer counterfactual questions is evident not only in such routine contexts as 'but for' cause, where courts determine what would have happened had the defendant not behaved negligently, and future lost earnings, where courts calculate what the plaintiff would have earned had she avoided injury, but also, more relevantly, in the area of evaluation of evidence.\(^{105}\)

First, in a malpractice suit against a lawyer who fails to file his client's claim, the court, to award damages, must find that, had the litigation against the underlying defendant taken place, the plaintiff would have prevailed. Such an inquiry raises the same issues implicated by the price model: determining what evidence was available to the parties and evaluating the parties' relative evidentiary positions.\(^{104}\) The vast collection of verdicts in these cases shows that courts have effectively developed techniques for evaluating a plaintiff's odds, even though often only part of the evidence concerning the underlying claim is available.\(^{105}\) An even

---

103 R.N. Straussfeld, 'If ...: Counterfactuals in the Law' (1992) 60 Geo.Wash.L.Rev. 359 (observing the pervasiveness of counterfactual inquiries in the legal context).
104 See 'Developments in the Law, II: Lawyers' Responsibilities to the Client-Legal Malpractice and Tort Reform' (1994) 107 Harv.L.Rev. 1557 at 1567 (explaining the evidentiary issues courts confront in conducting legal malpractice cases); J.E. Thuman, 'Measure and Elements of Damages Recoverable for Attorney's Negligence in Preparing of Conducting Litigation: Twentieth Century Case' (1991) 90 ALR 4th 1033. Especially pertinent to this discussion are malpractice cases in which the lawyer's alleged misfeasance is the failure to file the plaintiff's claim within the time limit of the applicable statute of limitation. Courts' ability to decide the parties' chances in the underlying claim, despite the erosion of some of the evidence, indicates that evaluation of lost evidence is possible. See, e.g., J.L. Rigelhaupt, 'Legal Malpractice by Permitting Statutory Time Limitation to Run Against Client's Claim' (1979) 90 ALR 5d 293 (presenting legal malpractice cases involving lawyers who negligently fail to file their clients' claims on time).
105 The largest evidentiary loss is often the absence of the underlying defendant from the litigation between the lawyer and her client. As the underlying defendant does not participate, he is not subject to discovery obligations. Courts have developed several ways to tackle this problem; see 'Developments in the Law', ibid. at 1569–70 (discussing such judicial solutions as 'settlement value' and 'loss-of-a-chance'). For an example from England, see Gregory v. Twid (1964), S.J. 219 (holding that the defendant – a
closer analogy occurs in the insurance context. Even if an insured person files her claim later than the contractually mandated deadline, courts in some jurisdictions allow the claim if the insurer cannot prove that it actually suffered substantial prejudice as a result of the insured’s delay.\textsuperscript{106} The insurer must show that timely notice would likely have allowed it to access evidence no longer extant that would have supported its position. In this instance, too, courts are asked to gauge the value of evidence that no longer exists.\textsuperscript{107}

While ascertaining the counterfactual consequences of missing evidence is possible, the task is arguably a costly one. The price model, however, also dispenses with various costly calculations required by current statutes of limitation. A precise comparison of the competing models’ administrative costs must await application of the price model. But a cursory analysis suggests that, in a number of ways, a move from the discovery rule to the price model entails substantial cost reductions. First, the discovery rule requires courts to decide when the harm occurred and whether, subsequently, the plaintiff acted diligently in filing her claim. As the price model never bars the plaintiff from filing, it would eliminate this inquiry. Second, the different exceptions to the discovery rule that toll or postpone the statutory time in special circumstances become superfluous under the price model.\textsuperscript{108} For example, it is irrelevant whether the plaintiff was disabled or insane at the time the injury occurred.

The inadequacies of our current regime of temporal limitation are becoming evident. Jurisdictions have increasingly begun to abandon the discovery rule and return to statutes of repose.\textsuperscript{109} Unsurprisingly, these new/old limitation statutes are troubled by the same inequity that initially led to the introduction of the discovery rule, namely, the barring of blamelessly tardy plaintiffs. A number of courts have found the new

\textsuperscript{106} See, eg., Shell Oil Co. v. Winterthur Swiss Ins. Co., 12 Cal. App. 4th 715 at 760 (1993) (hereinafter Shell Oil) ("The insurer must prove that the lack of timely notice had an adverse effect on the ability of the insurer to investigate and prepare a defense in the underlying claim"); UNUM Life Ins. Co. of America v. Ward, 119 S.Ct. 1380 at 1386 (1999) (examining California’s notice-prejudice rule).

\textsuperscript{107} Shell Oil, ibid. at 761 (determining that the defendant had not adequately demonstrated the significance of the eighteen potential witnesses who were no longer available).

\textsuperscript{108} See note 8 supra.

\textsuperscript{109} "The Fairness and Constitutionality of Statutes of Limitation for Toxic Tort Suits" Note (1983) 96 Harv.L.Rev. 1683 ("More importantly, in the last few years several states have replaced their discovery-of-damage rules with "statutes of repose." [S]tatutes of repose represent a return to the traditional form of time-bar statutes, even if they purport to mitigate the inequity of traditional statutes of limitations by extending the period during which a litigant may bring suit").
generation of repose statutes unconstitutional, holding that, by preventing the suits of reasonably tardy claimants, these statutes infringe upon their constitutional right to free access to the courts and discriminates against them by not granting them ‘equal protection’ vis-à-vis timely plaintiffs. The shortcomings of the current alternatives in temporal limitation argue for a move to a new, price-based approach that effectively protects the two litigants.

APPENDIX: Error costs

\[ Pf = \text{Plaintiff's chances to win when all feasible evidence (E(f)) is presented before the court.} \]

\[ Pi = \text{Plaintiff's chances to win considering the actual state of the evidence (E(i)).} \]

As explained, \( Pi > Pf \) because in both the immediate-discovery cases and the late-discovery cases, only the defendant is exposed to the risk of evidentiary loss. Since, in the case of equality, there is no erosion in the defendant’s evidence, we ignore this case and assume for the sake of simplicity that \( Pi > Pf \).

\[ D = \text{The damages alleged by the plaintiff.} \]

\[ A = \text{The modified damages according to the price model.} \]

\[ A = (Pf/Pi)D < D \]

The following analysis compares the error costs, namely, the deviation from the baseline’s distribution of damages, under the discovery rule and the price model. It shows that the price model produces fewer error costs when, due to the erosion in the defendant’s exculpatory evidence, the plaintiff’s chances rise by more than 100 per cent (\( Pi > 2Pf \)) and that the discovery rule generates fewer error costs when the plaintiff’s chances rise by less than 100 per cent (\( Pi < 2Pf \)).

110 In more than one-third of US states, statutes of repose have been declared unconstitutional. See J. Herring Hicks, ‘The Constitutionality of Statutes of Repose: Federalism Reigns’ (1985) 38 Vand.L.Rev. 627 at 657, and S.J. Werker, ‘The Constitutional Dimension of a National Products Liability Statute of Repose’ (1995) 40 Vill.L.Rev. 985 at 1053 (both listing states in which courts have ruled statutes of repose unconstitutional). See also P. Zablotsky, ‘From a Whimper to a Bang: The Trend Toward Finding Occurrence Based Statutes of Limitations Governing Negligent Misdiagnosis of Diseases with Long Latency Periods Unconstitutional’ (1999) 103 Dick.L.Rev. 455 at 457 (‘[N]early one-third of the states have either found their statutes of limitations governing medical malpractice to be unconstitutional as applied to victims of medical malpractice who could not have discovered their injuries until well after the statutory period had expired, or have interpreted their statutes to provide for open discovery. In many instances, it was the first time a court had ever found a statute of limitations to be unconstitutional’).
1. Error costs under the discovery rule
   \((P_i - Pf) \times D\)

2. Error costs under the price rule
   \((P_i - Pf) \times A + Pf \times (D - A)\)

3. Error costs under the discovery model are larger than error costs under the price model if and only if
   \((P_i - Pf) \times D > (P_i - Pf) \times A + Pf \times (D - A)\)
   \((P_i - Pf) \times D > Pf \times (D - A)\)
   \((P_i - Pf) > Pf\)
   \(P_i > 2Pf\)

4. Error costs under the discovery model are smaller than error costs under the price model if and only if
   \((P_i - Pf) \times D < (P_i - Pf) \times A + Pf \times (D - A)\)
   \((P_i - Pf) \times (D - A) < Pf \times (D - A)\)
   \((P_i - Pf) < Pf\)
   \(P_i < 2Pf\)