ISSN 1936-5349 (print)
ISSN 1936-5357 (online)

HARVARD

JOHN M. OLIN CENTER FOR LAW, ECONOMICS, AND BUSINESS
FELLOWS' DISCUSSION PAPER SERIES

CONTRACUAL ACCIDENTS

Yuval Procaccia

Discussion Paper No. 23

9/2008

Harvard Law School
Cambridge, MA 02138

Contributors to this series are John M. Olin Fellows or Terence M. Considine Fellows in Law and Economics at Harvard University.

CONTRACTUAL ACCIDENTS

Yuval Procaccia*

Abstract

I consider a unified framework for the problem of "contractual accidents," situations in which parties contract on the basis of a common misperception of reality. Several contract law doctrines pertain to such situations, including those of mutual mistake, frustration of purpose, impossibility, commercial impracticability and misunderstandings.

Existing economic analysis has so far been largely unsuccessful in providing a theoretical justification for these legal doctrines. I argue that this emanates not from the doctrine's incompatibility with principles of economic welfare, but rather from the inadequacy of the model used for their analysis. In particular, the standard model is premised on the assumption that the parties contemplate all potential eventualities from the outset. By maintaining that fundamental premise, the model assumes away the very problem it seeks to explain – that of a possibly incomplete perception of reality, which limits the parties' ability to map their genuine wishes onto a set of binding contractual terms.

I therefore consider an alternative framework that relaxes the assumption of perfect contemplation of the realized state. The parties in this modified model contemplate a set of potential states, but they know that the set is possibly incomplete. In turn, each contemplated state is assigned only a conditional probability, whereby the condition refers to the working assumption that the set is complete. Without further departing from standard assumptions, I show that – contrary to common wisdom – this modified framework explains the merits of the existing laws of contractual accidents.

* S.J.D. candidate, Harvard Law School. For comments and discussions, I am grateful to Oren Bar-Gill, Assaf Ben Shoham, Hanoch Dagan, David Enoch, Ehud Guttel, Alon Harel, Assaf Jacob, Christine Jolls, Alon Klement, Amir Licht, Barak Medina, Ariel Porat, Uriel Procaccia, Ivan Reidel, Uzi Segal, Eyal Zamir and seminar participants at the Hebrew Univeristy and the IDC, Herzlia. I also wish to thank the John M. Olin Center for Law, Economics and Business at Harvard Law School for its generous research support.

I. Introduction **2**
II. The Assumption of Perfect Contemplation **8**
III. The Standard Model and the Existing Analysis of Contractual Accidents **12**
1. Existing Theory *12*
2. Critique *15*
IV. The Modified Model **23**
1. The First Stage *23*
A. Mapping Intentions onto Outcomes 23
B. Trade of Unquantifiable Risk 25
C. Optimal Decisions to Contract 33
D. Mitigation of Damages 35
2. The Second Stage *37*
A. General 37
B. Reliance 38
C. Imperfect Verifiability of Accidents 43
D. Renegotiation Costs 44
E. The "Finality" of Transactions as a (Non)Reason for Full Enforcement 45
F. Summary 47
VI. Conclusion **48**

I. Introduction

One of the most fundamental problems arising in the law of contracts concerns agreements formed on the basis of a factual misconception. At the time of contracting, the parties are unaware of a pertinent fact that will impact their aggregate contractual gains and its distribution. The miscomprehended fact may concern the nature of the object that is traded, the meaning of contractual terms, or the background circumstances in which the contract is contemplated. In all such settings, the contract as written does not accurately convey the rights and obligations the parties were voluntarily willing to assume. Absent their common error, they would either have consented to a different contract or refrained from contracting altogether. Their contract, as it was formed, is a result of an "accident."

As contractual accidents come in various forms, several doctrines have been developed to address them. Most notably, the doctrine of "mutual mistake" applies to misconceived facts that objectively "exist" in the world at the time of contracting; the doctrines of "impossibility" (or "commercial impracticability") and "frustration of purpose" typically concern unforeseen risks associated with future contingencies;[1] the laws of contract formation refer to the consequences of misunderstandings; the laws of contract modification pertain to the validity of renegotiated terms when the parties fail to foresee future developments; and the laws of illegal contracts govern situations in which the parties are unaware of their agreement's illegality. In all of these scenarios, the resulting agreement contains an element of accident. As the subjective world envisioned by the parties materially departs from reality, a question arises as to the agreement's desirable legal effect.

This paper develops a theory of the laws of contractual accidents. It focuses, in particular, on two of the most prominent features of these laws. The first examined feature is the legal consequence of a recognized accident – the consequence of excuse from performance. The implication of an excuse is that both parties are relieved of all outstanding obligations, and remain liable to one another only in restitution. Although courts are formally authorized to order other types of remedies in lieu of restitution, it is an option they hardly ever exercise in practice.[2] The first major question considered here is

[1] In some cases, however, these doctrines could also apply to misconceptions of an existing fact. See Restat. 2d of Contracts § 266.

[2] Restatement 2d, §158 (mistake), §272 (impracticability and frustration); E. ALLAN FARNSWORTH, CONTRACTS, §§9.9,9.3 (2d. ed. 1990). ARTHUR L. CORBIN, CORBIN ON CONTRACTS §78.6 (stating that restitution is the most common remedy actually

therefore whether the rule of excuse is indeed optimal in accident cases. Does it reflect the mutual interest of most contracting parties ex ante, and if so, how?

The second doctrinal feature examined concerns the observation that courts exercise great caution when confirming that an accident had occurred. Performance is excused rarely, only when accidents are in some sense major and severe, and when the promisee's reliance of the contract is not exceedingly large. Is this restrictive approach warranted? And if so, should these particular criteria be the ones to guide the courts in making this determination?

The first step in addressing these normative issues is of course to form some initial understanding of *what exactly accidents are* and how they come about. Interestingly, the traditional legal scholarship and the economically-oriented one offer two very different answers to that question. For lawyers, accidents typically emanate from the parties' inability to fully capture reality. As they fail to consider the possibility of all potential eventualities, the allocation of rights and obligations under the contract diverges from what they had truly intended and supposed. This failure is due to the human limits of imagination, and the limited capacity to process all available information to its finest detail.[3] Further, even when information is

---

applied, and that courts have been "reluctant to order more than a generous remedy of restitution".) For further analysis of the case law in that regard see Andrew Kull, *"Mistake, Frustration, and the Windfall Principle of Contract Remedies,"* 43 Hastings L.J.1, 10 (1991) (asserting that cases are typically decided in "whatever manner leaves the parties free of further obligations to each other, whether by way of performance or restitution."); Leon Trakman, *"Winner Take Some: Loss Sharing and Commercial Impracticability,"* 69 Minn. L. Rev. 471, 471 (1985) (stating that despite the theoretical availability of various loss-sharing remedies, courts have generally not utilized their discretion to order them.)

[3] The perception of contractual accidents as arising from an imperfect capacity to contemplate all possible eventualities is demonstrated well in the language of existing laws. Section 151 to the Restatement (Second) of Contracts (hereinafter Restatement) defines a mistake as "a belief that is not in accord with the facts." Similarly, under Sections 261 and 265 the doctrines of impracticability and frustration respectively are said to apply to an event "the non-occurrence of which was a basic assumption on which the contract was made." Section 2-615 of the Uniform Commercial Code uses the same language to define impracticability and is entitled "Excuse by Failure of Presupposed Conditions." In a similar vein, a misunderstanding is defined by Section 20 to the Restatement as a situation where "the parties attach materially different meanings to their manifestations and…neither party knows or has reason to know the meaning attached by the other." And Section 201, which concerns the law of contract formation, provides that mutual assent may fail if the parties attach different meanings to a term, and are unconscious of their common error.

Moreover, as the laws of accidents are restricted to situations in which the state of the world had not been contemplated, the Restatement explicitly denies relief if the

lacking but could be purposely acquired, the desirable scope of such an investment may be impossible to determine. For to pinpoint the type of information that is worthy of investigating, a decision-maker must rely on some prior comprehension of the possible results that the investment would yield: She must conceive of the set of all possible conclusions, assign each a corresponding probability, and evaluate the functional relationship between the expected value of the acquired information and the expected cost of acquiring it. Realistically, parties are often incapable of executing such a complex procedure optimally, either because it requires an initial level of knowledge they do not possess, or because the associated mental processing costs are prohibitive. Thus, although the parties are aware of the possibility of accidents in the abstract, they are unable to prevent them in any practical sense.[4]

The standard economic model, in contrast, rests on a very different conception of accidents. In the world of the prevailing model, the parties are assumed to be exogenously endowed with knowledge of all potential states of the world. Moreover, they costlessly assign a probability to each such state, and these probabilities are internally consistent in that they add up to unity.[5] The model thus avoids the fundamental question of how these perceptions of reality emerge, and what characterizes the critical mass of information required to form them. It simply assumes that the parties identify all possible eventualities and recognize all potential risks.

It follows from the standard setup that within the confines of the model, “real” accidents, in the sense ascribed to that term in legal

---

realized event had been contemplated ex ante. Section 154, pertaining to mutual mistake, thus provides that excuse will not be granted if "the risk [was] allocated to [the promisor] by agreement of the parties" or even when the promisor was only "aware, at the time the contract [was] made, that he [had] only limited knowledge with respect to the facts to which the mistake relates but [treated] his limited knowledge as sufficient."

[4] This hypothesis is of course also supported by behavioral and experimental findings. The biases and heuristics that will generate a general misconception of contractual facts are numerous. See e.g., Baruch Fischhoff, Paul Slovic, and Sarah Lichtenstein, *Fault Trees: Sensitivity of Estimated Failure Probabilities to Problem Representation*, 4 J Exp Psych: Human Perception & Performance 330 (1978) (showing empirically that individuals consistently under-estimate the risk of events they do not explicitly conceive of); Daniel Kahneman and Amos Tversky, *Availability: A Heuristic for Judging Frequency and Probability*, 5 Cognitive Science 207 (1973) (arguing that in actual choice processes, only the simplest and most available scenarios are typically considered, while contemplation of complex interacting factors is rare.)

[5] See LEONARD J. SAVAGE, THE FOUNDATION OF STATISTICS (1954); R. DUNCAN LUCE AND HOWARD RAIFFA, GAMES AND DECISIONS (1957), ch. 13.

doctrine, can never truly occur.[6] At most, the parties might find themselves surprised, in the sense that an event that had been assigned low probability ex ante, happens to materialize ex post. But as all contingencies are assumed to be conceived of right at the outset, accidents could by definition only represent an unlucky realization of an explicitly assumed risk, never a realization of a truly unforeseen contingency.

The two competing notions of accidents – the legal and the economic – have both been used to explain the contours of existing doctrine. However, for reasons elaborated upon below, the prevailing economic analysis has so far been largely unsuccessful in providing a satisfactory justifying theory of the law. That is indeed unsurprising, given the fundamental divergence between the legal and the economic understanding of the phenomenon. Although the two streams of scholarship speak about the same body of law, they understand it to address a very different conceptual problem. And given their disagreement about the nature of the problem, one only expects them to also disagree about the attributes of the desirable solution. Upon examination of the model's implications, one indeed finds an unbridgeable gap between the law as it is, and the existing economic theory constructed to explain it.

So what is the lesson to be learned from existing economic theory? If one is prepared to carry the standard account to its inevitable conclusions, then one must take the position that the current legal regime of contract law must be fundamentally revised, notwithstanding its prevalence throughout culture and time. This conclusion is not logically impossible, and in any event this paper does not seek to quarrel with that position. Instead, the aim of this paper is to develop an alternative theory which might provide a more convincing theoretical justification for existing legal policy. The suggested approach reconciles theory and doctrine not by suggesting that the law should be changed to fit the model, but rather by fixing the flawed model to accurately describe the problem of accidents as it occurs in reality. It is shown that once the flaw is removed, the revised theory serves well as a justifying framework, and its conclusions largely correspond to the attributes of existing law.

The modified model developed here is thus a hybrid of the traditional legal approach, and the standard economic one. It differs from traditional legal reasoning in that it relies on an economic model.

---

[6] See e.g., Eddie Dekel et al., *Standard State-Space Models Preclude Unawareness*, 66 Econometrica 158 (1998) (demonstrating that unawareness to a possible event cannot be modeled within the confines of a state-space model);

It is a model, however, which relaxes the assumption of perfect contemplation, and in that sense it departs from traditional economic reasoning as well. It is suggested that this is indeed the most compelling approach to take from a descriptive point of view. When considering problems such as mistakes, unforeseen contingencies and misunderstandings, one expects results to be qualitatively distorted if one assumes that parties are not truly mistaken, that they do in fact foresee all contingencies and that they understand each other's communications perfectly. The assumption of perfect contemplation is thus a crucial one, and by assuming it counterfactually, the derived conclusions lose much relevance with respect to the problem at hand.[7]

Under the suggested formulation, accidents arise out of occurrences that the parties do not contemplate at the time of contracting. They are aware ex ante that accidents are possible, but they do not know the particular attributes of the accident state. Furthermore, they are insufficiently aware of the type of information that they lack, and hence they cannot effectively invest in acquiring it. An accident is thus characterized as an exogenous shock that the parties have no

[7] Although the standard neoclassical model assumes away the problem of ignorance, the shortcomings of that simplifying assumption have been long recognized by economists. Most notably, Frank Knight has distinguished between risk and uncertainty, the former referring to situations where probability is known whereas the latter involves ignorance with respect to probabilities. See FRANK H. KNIGHT, RISK, UNCERTAINTY AND PROFIT 19–20, 197–232 (1921). Keynes has similarly distinguished between uncertainty and probability, defining the former as a situation in which "there is no scientific basis on which to form any calculable probability whatever." He went on to argue that in the case of uncertainty "the hypothesis of a calculable future leads to a wrong interpretation of the principles of behavior." John M. Keynes, *The General Theory of Employment*, Quarterly Journal of Economics 209, 213 and 222 respectively (1937). In a similar spirit, Hicks contended that "one must assume that the people in one's models do not know what is going to happen, and know that they do not know just what is going to happen." That leads him to the conclusion that "the usefulness of 'statistical' or 'stochastic' methods in economics is a good deal less than is now conventionally supposed. We have no business to turn to them automatically; we should always ask ourselves, before we apply them, whether they are appropriate to the problem at hand. Very often they are not." JOHN R. HICKS, CAUSALITY IN ECONOMICS pages vii. and 129 respectively (1979). Even Savage, whose classic work is often cited as a primary justification for the assumption of perfect contemplation has admitted that "a person may not know the consequences of the acts open to him in each state of the world. He might be... ignorant." The presumption that perfect contemplation is possible "carried to its logical extreme... is utterly ridiculous... because the task implied in making such a decision is not even remotely resembled by human possibility." LEONARD SAVAGE, THE FOUNDATIONS OF STATISTICS 15-16 (1954). For somewhat more recent discussions see Tony Lawson, *Uncertainty and Economic Analysis*, 95 Economic Journal 909 (1985) and Paul Davidson, *Is Probability Theory Relevant for Uncertainty? A Post Keynesian Perspective*, 5 Journal of Economic Perspectives 129 (1991).

control over. They cannot avoid it and they cannot meaningfully predict its probability or the severity of its impact. However, as the possibility of the shock is itself acknowledged, they can and do form preferences with respect to its *legal consequences*. This paper seeks to explore the nature of those preferences. By providing an answer to this question one identifies the majoritarian default rule, whose implementation is taken to represent the policymaker's objective.[8]

From all other respects, the present framework retains the assumptions of the standard model, on both the descriptive and normative levels. Namely, individuals are still taken to be rational and seeking to advance their self-interest given the information they possess, and the goal of legal policy is still taken to create the conditions that allow them to maximize their gains from trade. This rather modest modification of the model turns out to be sufficient, however, to provide a new theory regarding the effect of existing doctrine, a theory that lends significant support, on welfare grounds, for the doctrine's continued application.

The Article proceeds as follows. Section II briefly discusses some general ramifications of the standard model's assumption of perfect contemplation, and further explains why it is an inadequate framework by which to analyze contractual accidents. Section III reviews the existing law-and-economics literature on contractual accidents and argues that it neither offers a persuasive account of actual doctrine, nor is capable of doing so in principle, given its assumption of perfect contemplation.

---

[8] This is an attractive normative goal to pursue, not only because it leads to maximization of gains in a utilitarian sense, but, perhaps more importantly, because it is an objective that rational parties both share ex ante. As in many other applications of economic analysis to the law of contracts, if a different rule were applied, and result in greater accident costs, the parties could both benefit by opting out of that default rule and adopting the value maximizing rule instead. As they would then split the added gains through bargaining, the move towards joint value maximization would leave them both better off.

It should be noted that in some areas of contract law an optimal legal rule is a "penalty" default rather than a majoritarian one. A penalty default prescribes a term that is known to run against the parties' hypothetical agreement. Ayres and Gertner have shown that such a rule could sometimes create desirable incentives, as it induces the holder of private information to disclose it, and it might also lead to a reduction in the administrative costs of contract enforcement. In the particular context of contractual accidents, however, neither of these justifications is pertinent. By assumption, the limited perception of the world is shared by the parties, and hence an issue of private information and disclosure does not arise. Further, as accidents are taken to be unavoidable, penalizing the parties when they occur would not result in reduced administrative costs. See Ian Ayres and Robert Gertner, *Filling Gaps in Incomplete Contracts: An Economic Theory of Default Rules*, 99 Yale Law Journal, 87 (1989).

Section IV then develops the modified model, which relaxes the assumption of perfect contemplation. Within this framework, the conditions under which the rule of excuse is optimal are then derived. Section V discusses the circumstances in which courts rightfully deny accident claims, and accounts for the cautionary policy that governs that application of the doctrine. Finally, Section VI offers brief concluding remarks.

II. The Assumption of Perfect Contemplation

Prior to delving into the substantive analysis of contractual accidents under the standard model, a few words are in order regarding the deeper underpinnings of the model's assumption of perfect contemplation, which lies at the center of the analysis.

The assumption of perfect contemplation comprises of two distinct components. One is that the parties conceive of all possible *states of the world*; and the other is that they can *assign a probability to each of those states*. A person who contemplates reality perfectly can therefore never be truly surprised. If an event has occurred, it must be the case that it had previously been contemplated and accounted for.

Let us consider both components of the assumption, beginning with the one concerning knowledge of probabilities. The idea that people know probability distributions is sometimes justified on the grounds that - as a conceptual matter – *a logically valid probability assessment can always be formed*. The underlying reasoning is that if one knows something meaningful about a set of potential events, then that information could be utilized to produce a best estimate of the probability distribution corresponding to that set. Alternatively, if no information is available at all, then one must apply the well-known *"principle of insufficient reason,*" guiding him to assign equal probability to all events.[9] Hence, probability, viewed through this lens, is merely a quantitative expression of whatever one knows or does not know about the likelihood of an event. As one either knows nothing or knows something about any conceivable issue, a probability "vacuum" can never exist. Hence, the assumption that agents come into the world equipped with probability distributions seems to rest on a solid conceptual ground.

Undeniably, the claim that in the absence of positive information one ought to assign equal probability is intuitively compelling. After all, if nothing is known about the relative likelihood

[9] Thus, for example, if one knows that an urn contains a ball that might be red or black, and nothing further is disclosed with respect to probabilities, then one must assign equal probability to each of the possible colors.

of two events, it seems difficult to justify an arbitrary decision to assign one of them greater probability than the other. It turns out, however, that the statement that ignorance provides a *positive* reason to assign equal probability to the two events is not logically sound. One way to express the failure of this notion is by demonstrating that the probability estimate derived from this principle depends on the manner in which the event is framed: Namely, it can be easily shown that if one uses different words to describe the same event, the principle of insufficient reason may well assign a different probability estimate to each description. Clearly, if an estimate depends on wording, rather than on some innate property of the event itself, or the world surrounding it, it cannot be regarded as an estimate at all: At most, it may provide a comforting mental illusion of one. It follows that if one's objective is to enhance the likelihood of a correct choice, then the principle of insufficient reason offers no meaningful guidance. In the present context the implication is that if the parties truly have no indication as to the probability of an accident, utilizing the principle of insufficient reason is as logically coherent as assigning probability pursuant to an instruction of a wizard.

The logical defect of the principle is well illustrated by the following example.[10] Consider a factory producing cubes (which by definition are equal in width, length and height.) Suppose that you are told that the length of each side of the cube lies within the range of 0 and 2 but the distribution of these possible values is unknown. You are asked to set an estimate as to the probability that the length lies within the range of 0 and 1. As you have no reason to suspect that the real value is more likely to be above 1 or below it, the principle of insufficient reason guides you to set your estimate at 0.5. So far, so good. Now, however, consider the following question: The factory produces cubes whose *volume* lies within the range of 0 and 8. You are asked to form an estimate as to the probability that the volume lies between 0 and 1. Applying the principle of insufficient reason again, your prediction is now 1/8. However, recall that the volume of a cube equals its length raised to the third power. Hence, the set of cubes whose volume is less than 1 is exactly the same as the set of cubes whose length is less than 1. The probability estimate for these two events must therefore be the same. Thus, although the two questions are in fact the very same question, the principle of insufficient reason has generated two different responses to them. That divergence is purely a

[10] The example is due to BAS C. VAN FRAASSEN, LAW AND SYMMETRY (CLARENDON PRESS, 1989).

product of the manner in which the relevant event was phrased, not of any substantive difference in the information provided.[11]

The essential reason for the failure of the principle of insufficient reason is that it relies on an under-defined notion of symmetry between categories of events. It rests on the intuition that if events are symmetrical in all observable respects, there should be no reason to differentiate between them in the assignment of probability. It so happens, however, that the appearance of symmetry is misleading. When nothing is known about a phenomenon, the state space can be depicted in various different forms, each of which entailing a different representation of symmetry. Thus, by assigning a uniform distribution to states under one framing, one inevitably assigns a non-uniform distribution to the same state space under a second. At the end of the day, the probability estimate produced by this principle is determined not by the symmetry attribute at all, but rather by one's arbitrary choice of a depiction of the state space. The principle thus applied is therefore empty of meaningful substance: One is as well off utilizing it as one would be ignoring it altogether.

It transpires that knowledge cannot be logically derived from lack of knowledge. A probability estimate must grow out of something other than the void of ignorance. However intuitively appealing, a procedure that produces a prediction without resorting to substantive information about the world cannot be premised on consistent logic. It follows, in turn, that the assumption that agents have knowledge of probabilities is far more involved than what intuitively meets the eye.

Consider now the assumption that agents perceive of all possible contingencies, or states of the world. Similarly fundamental problems plague that assumption as well. In reality, it is of course evident that individuals do not conceive of all potential eventualities, and the degree of incompleteness is often quite substantial. To feel the significance of that imperfection, consider, for instance, the fact that for a complete contemplation of the state space to be formed, the condition of logical omniscience must be satisfied. This condition requires that if

[11] An additional example, due to Von Mises, is as follows: Suppose you are told that water and wine are mixed in a glass in some unknown proportion between 1:1 and 1:2. By the principle of insufficient reason we may deduce that it is as likely that the ratio of water to wine lies within the interval of 1 to 1.5 as within the interval of 1.5 and 2. Now, however, consider the proportion of wine to water. Applying the rule again, we conclude that the ratio is as likely to lie in the interval of 0.5 and 0.75 as within the interval of 0.75 to 1. But that, in turn, implies that the ratio of water to wine is as likely to lie between 1 and 1.3333 as between 1.3333 to 2. That, of course, amounts to a contradiction. See e.g., MICHAEL G. BULMER, PRINCIPLES OF STATISTICS 8-11 (1979). For further discussion see DONALD A. GILLIES, PHILOSOPHICAL THEORIES OF PROBABILITY 37-49 (2000).

one knows *x*, then if *x* logically entails *y*, then one must also know *y*.[12] If logical omniscience is violated, it is straightforward that events such as *y* might not be contemplated, in violation of the assumption of a complete state space. But such an assumption is manifestly inconsistent with what we plainly know about the limits of human cognition. Thus, for instance, it might be safe to presume that some of us do not know all the theorems of mathematics (proven and yet unproven), although they are nothing more than logical conclusions derived from a set of assumptions.[13] Unless one is willing to accept such capability as a reasonable description of reality, one must seriously doubt the plausibility of the assumption of a complete state space.

Of course, by highlighting the wide gap between the presumed level of information under the assumption of perfect contemplation and the "real" level of knowledge held by genuine people, one still does not critique the use of the assumption in economic modeling. As any economist would point out, economic models consciously rely on assumptions that are not entirely realistic. If it were otherwise, they would no longer comprise a "model" of reality, but rather provide a fully-detailed picture of reality itself. Models utilize imperfect assumptions because perfect descriptions would generate too much complexity, hopelessly undermining any prospect of achieving analytical progress. Hence, to truly critique the assumption of perfect contemplation for modeling purposes, one's burden is not merely to show that it is in some sense counterfactual, but rather that it undermines the very capacity of the model to generate useful results.

To be sure, in many different applications, using the assumption of perfect contemplation is warranted. If, say, the subject-matter of analysis is the effect of policy on choice between several contemplated alternatives, then it might well be reasonable to assume, for the sake of the analysis, that *all* possible contingencies are contemplated. Assuming otherwise would only divert attention from the questions one seeks to focus on. This logic, however, does not carry over to the case of contractual accidents. For when the very objective of legal doctrine is to identify the rules that ought to govern non-contemplated contingencies, then assuming the counterfactual, namely that all contingencies *are* contemplated, makes no analytical sense. For if all contingencies are contemplated, one cannot even define the problem at hand, let alone provide solutions to cure it. A model utilizing such an assumption must inevitably examine an entirely different problem from

[12] See Robert J. Aumann, *Musings on Information and Knowledge*, 2 Econ Journal Watch, 88, 90-91 (2005).
[13] Id, at 90.

the one addressed by legal doctrine. It may produce thoughtful and interesting answers, but they remain answers to the wrong question.

III. The Standard Model and the Existing Analysis of Contractual Accidents

1. Existing Theory

The argument made above suggests that in the context of contractual accidents, using the assumption of perfect contemplation is likely to be a bad idea. With that notwithstanding, the assumption has been embraced by actual economic analysis. This, in turn, presents an opportunity to further examine the basis of the assumption's adequacy. If the assumption truly undermines one's ability to examine the issue at hand, then one would expect to find fundamental flaws in the results of existing theory – flaws that are not attributable to a specific modeling choice, but rather to an innate conceptual problem that can only be cured if the assumption of perfect contemplation is relaxed. This is indeed the argument set forth in this section. It begins by reviewing the major insights of existing theory, and then proceeds to offer a critical examination of their validity.

The approach of existing law-and-economics analyses emanates from an early seminal article, authored by Richard Posner and Andrew Rosenfield.[14] Their argument revolves around the notion that accidents are contemplated and hence the law can desirably impact the manner in which their costs are managed. The crux of their claim is that an accident rule is desirable if, given the parties' contemplation of a possible accident, it encourages them to efficiently invest in diminishing its expected harmful impact, and if it efficiently allocates the remaining accident risk to the party who can bear it most cheaply.

The analysis constructs a two-level examination. First, the court must consider whether the risk could have been efficiently avoided had the parties taken "precautions." A "precaution" is defined in this framework as an action that directly reduces expected accident costs by lowering either its probability or the magnitude of its corresponding harm. For example, a drought might prevent a farmer from performing her obligation to supply a promised quantity of crop. However, if she puts an irrigation system in place, she can lower the detrimental effect of a drought if it occurs, or eliminate it completely. Alternatively, if she moves her farm to a different climatic region, she might enjoy a reduced probability of droughts.

---

[14] Richard A. Posner and Andrew. M. Rosenfield, *Impossibility and Related Doctrines in Contract Law: An Economics Analysis*, 6 J. Leg. Stud. 83 at (1977).

As precaution measures are costly to acquire, parties would ideally invest in them only to the extent that their marginal benefit in terms of reduced accident expenses exceed their marginal cost. One objective of an accident rule is thus to motivate the parties to invest up to the optimal level, but not more.

Once the pool of efficient precaution measures is exhausted, some risk may still remain. Thus, the second level of examination considers who can insure against that risk more efficiently. Thus, for example, when many different buyers purchase a good from a single seller, the seller might be in a better position to acquire insurance, as his insurance policy would cover all transactions collectively. Similarly, some parties may be in a better position to self-insure, by diversifying or hedging the risk. Thus, a farmer may own several farms in different climatic regions, and thereby lessen the probability of becoming unable to perform.

With the above definition of efficiency in place, the meaning of an optimal accident rule under the standard model is straightforward. An accident rule is optimal if it places incentives for the parties to take efficient precautions, to acquire efficient insurance, and to allocate the residual risk to its most efficient bearer. Posner and Rosenfield's primary assertion is that the rule of excuse, if applied correctly, generates the desired result. An excuse should be granted to a promisor, it is argued, if the promisee had failed to prevent the accident or insure against it although he could have efficiently done so, or if he had otherwise been the more efficient bearer of the residual risk.[15] By providing relief in these instances, the rule optimally causes the promisee to internalize the accident losses and invest the optimal amount to manage them. By the same token, by denying relief to a promisor who should have borne the consequences of accidents, the accident rule forces him to internalize those costs, which again produces the desired result.

Since its publication in 1977, Posner and Rosenfield's analysis has been extended by others in various ways. Christopher Bruce has highlighted the rule's effect on the parties' incentives to mitigate damages.[16] To accomplish optimal mitigation, he suggested a rule that

---

[15] Posner and Rosenfield consider two slightly different versions of this rule. One version is simply a negligence rule, which conditions release on the promisor's optimal behavior. The second version is more restrictive, and requires not merely optimal behavior by the promisor, but also sub-optimal behavior by the promisee. Posner and Rosenfield consider both versions as efficient, provided that they are applied without error. See Id., at 110-111. For a formal, detailed analysis of both rules, including a critique of the efficiency claim, see Appendix.

[16] Christopher J. Bruce, *An Economic Analysis of the Impossibility Doctrine*, 11 J.

would condition a discharge not only on the parties' ability to avoid accident harm but also on their mitigation efforts.[17] Pietro Trimarchi proceeded to incorporate the observation that insurance is often unattainable.[18] In such cases, he contended, the rule of excuse ought to be judged on the basis of ex post considerations of allocative efficiency rather than its capacity to generate optimal insurance ex ante.[19]

Gerhard Wagner has further suggested that the rule of excuse induces the promisee to select the optimal level of reliance. The reasoning is that only under the rule of excuse, the promisee fully internalizes both the upside of reliance (the investment's anticipated return) and its downside (the wasted cost of reliance if due to an accident, trade is longer desirable). Conversely, if expectation damages are awarded, the promisee keeps the return from reliance in all states of the world, and hence does not internalize the downside. The remedy of expectation damages therefore functions as an insurance mechanism, shielding the promisee from the risk of wasted reliance. The result is that the promisee is encouraged to over-invest.[20]

Victor Goldberg has offered yet another rationale for the rule.[21] He emphasizes the possibility that even when a promisor cannot perform in the manner stipulated by the contract, he might nevertheless be able to render performance in an alternative form. Thus, for example, if a provider of a computer system is unable to perform because, say, he was unable to obtain a license for a particular component, he might still be able to offer an alternative system, consisting of somewhat different features. Such an alternative is surely

---

Legal Stud. 311 (1982).

[17] Id., at 316.

[18] Third party insurance could be impossible to attain, he argued, because the risk in question could impact the entire pool of potential insureds systematically (such as in the cases of hyper-inflation or a war). Insurance companies might not be able to effectively diversify such risks, and would thus not be situated in a superior position to bear them. Forward contracts might similarly not exist for all relevant risks. Pietro Trimarchi, *Commercial Impracticability in Contract Law: An Economic Analysis*, 11 Int'l Rev. L. & Econ. 63, 65-70 (1991).

[19] Id. Trimarchi additionally notes the possibility that parties may not be able to predict accident risk, in which case the analysis of Posner and Rosenfield may be inapplicable. That point, however, is not further developed.

[20] Gerhard Wagner, *In Defense of the Impossibility Defense*, 27 Loy. U. Chi. L.J. 55 (1995). Although Wagner acknowledges the possibility of non-contemplated contingencies, his primary argument regarding the optimality of reliance could only be valid in a world where the promisee holds an estimate of accident probability. If he does not, then it follows that the privately desirable level of reliance cannot be identified and hence one does not expect it to be overall optimal.

[21] Victor P. Goldberg, *Impossibility and Related Excuses*, 144 Journal of Institutional and Theoretical Economics, 100 (1988).

only an imperfect substitute for the agreed upon system, as it may be more or less costly for the promisor to provide and it may yield more or less value to the promisee. Nevertheless, given the circumstances, providing it might be optimal.

As Goldberg observes, however, if several such means of performance are available, choosing the appropriate one ex post is likely to trigger an informational problem: The promisor, who would bear the cost of providing the alternative system but not appropriate the benefits would seek to provide the cheapest possible alternative, regardless of its value to the promisee. Likewise, the promisee, who would realize the benefit but not incur the cost, would demand the mode of performance that maximizes his benefit regardless of the cost. If the court cannot verify which method of performance is the most cost-effective, the result would be inefficient.

The rule of excuse, it is argued, resolves that problem. If the promisor is entirely relieved from the duty to perform, the promisee would not be entitled to any alternative form of performance by the promisor, and so would have to acquire the right to obtain the desired system. But if he pays for such a system in full, he is forced to internalize not only the benefit of performance but also its cost. His choice of the system he wishes to acquire would thus be efficient, and overall gains from trade would be maximized. The rule of excuse would therefore function as an optimal default rule.

To sum up, the following list recapitulates the various arguments in support of the rule of excuse, within the context of the standard model. Under these arguments, the rule of excuse generates

(a) *optimal precautionary behavior* with respect to the adverse accident contingency; and *optimal incentives to insure*.
(b) *optimal reliance* decisions;
(c) *optimal mitigation of damages*; and
(d) an *optimal method of substitute performance*.

These four arguments are next referred to as *propositions (a) through (d).*

2. Critique

The various arguments reviewed above all seek to explain why the rule of excuse, as generally applied, is conducive to the objective of welfare maximization. This section suggests, however, that, as a matter of principle, arguments derived from the standard model cannot

succeed in establishing this claim as long as the assumption of perfect contemplation is retained.

Upon closer examination of the substantive arguments reviewed in the previous section, it transpires that all are exposed to at least one of two critiques. Some arguments – despite their appearance - do not in fact offer a valid justification for the rule of excuse under any state, accident or otherwise. But yet more importantly, those arguments that do offer a valid justification for the rule also justify a rule of no-liability for ordinary breach of contract. It follows that if one finds those latter arguments compelling inasmuch as they apply to states characterized as "accidents," then one must also recommend that a rule of no liability be applied across the board, in all instances of non-performance. Either way, the fundamental distinction made under positive doctrine between non-performance that yields liability and non-performance that yields an excuse, cannot be adequately explained by the standard theory.

This impasse is indeed not incidental. It discloses an inherent attribute of the standard model, and is a direct consequence of its assumption of perfect contemplation: As the standard analysis denies the very possibility of genuine accidents, it cannot rationalize the differential remedial treatment of accidents and ordinary breaches. As it does not recognize the distinct informational environment leading to each instance of non-performance, it cannot explain why failure to perform is sometimes excused, while in other instances leads to full liability. As a matter of principle, the framework of the standard model is thus unsuitable for the task of producing a general justification of this branch of the law.

In the remainder of this Section, the substantive arguments derived from the standard model are critically revisited in turn. For the purposes of this Section, I assume, along with the adherers of that model, that all contingencies are contemplated. With that assumption in place, I seek to show why the model is indeed unsuccessful in offering a valid justification of the doctrine. I begin by explaining why some of the major arguments fail to validly establish the rule's desirability in accident settings; and then proceed to explain why all remaining arguments are inconsistent with the fact that excuse is not the applicable default rule for ordinary breach.

Failure to Establish Optimality in Accident Settings: The rule of excuse is justified primarily on the basis of its impact on precautions, reliance and risk allocation. I now turn to reexamine these arguments. The verbal analysis presented here, is accompanied by a formal derivation of the results, developed in the Appendix.

Consider first the claim that the rule of excuse, when applied as prescribed, generates optimal incentives to take precautions and to insure (proposition (a)). The logic of this assertion is similar to that of the standard negligence rule in the torts context: If the promisor takes due care, he incurs no liability; and if he takes inadequate care, he must bear the full social cost generated by the realized risk. Thus, provided that the threshold level of due care is properly defined, taking due care is his privately optimal strategy.[22] It should be observed, however, that this happy result does not carry over to the case where the promisee happens to be the efficient harm avoider. Similarly to the previous case, if the promisee takes due care, he is free from liability. But unlike the previous case, if he takes inadequate care he remains free from liability! While he must bear his own accident costs should the promisor be excused, he is never made liable for the accident losses borne by the promisor. He therefore fails to internalize the full social cost associated with his negligent conduct. His privately optimal strategy may therefore well be to take *inadequate* care.[23]

Further, the rule of excuse allows for the possibility of inefficient failure to perform. As the promisee takes inadequate care, it follows that *in equilibrium* there are instances in which an accident occurs and the promisor is excused. As he is not liable for the loss that

---

[22] In fact, liability in expectation damages will generally exceed the external harm caused by the failure to acquire information. The reason is that precautions are valuable to the promisee inasmuch as they would have caused him to change his reliance choices (understood here to include the costs of lost opportunities). Thus, in particular, if – given the actual state of affairs - trade transpires to be undesirable (as Posner and Rosenfield assume), then the value equals the sum he would otherwise wastefully sink into the contractual relationship and the added gains that would accrue from an alternative opportunity. But expectation damages would generally exceed that amount. That is so because the expectation measure compensates a promisee not only for the costs of reliance, but also for the higher return that the promisee expected to gain from it.

A familiar result in the analysis of negligence is that setting the liability measure above the level of external harm does not generate over-deterrence, provided that courts can perfectly identify the optimal level of care. For a discussion of that point, see, e.g., STEVEN SHAVELL, FOUNDATIONS OF ECONOMIC ANALYSIS OF LAW, ch. 4 (2004). As widely acknowledged, however, the assumption that courts perfectly apply the negligence standard is quite unrealistic. If one were to incorporate the more plausible assumption that judges occasionally err in their application of the standard (and even if the expected error equals zero), it would follow that the rule induces the promisor to over-invest in precautions. Hence, under the framework of the standard model, this is yet another source of inefficiency generated by the rule of excuse.

[23] As shown in the Appendix, both versions of the suggested rule lead to the result of inadequate care by the promisee. Whereas the negligence rule produces this result as an exclusive possibility, the alternative version yields the result under fairly broad conditions, but not across the board.

failure to perform inflicts on the promisee, it is a cost he disregards. Hence, the failure to perform occurs too often, particularly in states of the world in which trade is desirable, as it produces greater social value than cost.[24]

The prospect of inefficient failure to perform, moreover, adversely impacts the parties' reliance incentives, thereby undermining the validity of proposition (b).[25] At the time of contracting, both parties know that the promisor might fail to perform even if performance is efficient. If they expect to renegotiate a new contract in such a scenario, a holdup problem is created, which drives the relying party to restrict his reliance investment to a lower, inefficient level.[26] Alternatively, if upon default the parties do not renegotiate due to high transaction costs, then reliance yields no value at all in accident states. In that case, reliance is optimal only in a second-best sense

---

[24] One might seek do defend the Posner-Rosenfield position by claiming that in impossibility cases, the danger of inefficient breach is not a significant one. The reasoning would be that in impossibility situations performance is physically prevented by an intervening event, and hence the non-desirability of trade ex post is evident and verifiable. The court can thus allow a discharge when trade is undesirable and deny it otherwise. That argument, however, is largely unconvincing. For even when performance in the exact form stipulated in the contract is made impossible by a supervening event, there are frequently alternative possible modes of performance that a promisor could take to satisfy the expectation interest of the promisee. For example, if a horse dies before it is delivered, rendering a contract impossible to perform as written, the promisor could perhaps still deliver a different horse and thereby partially satisfy the expectation interest of the promisee. When a drought renders it impossible to deliver the agreed amount of crops, performance is still possible if the farmer acquires the missing amount from elsewhere. Thus, even if the character of the supervening event is verifiable, knowing of that event is insufficient to determine whether trade is ex post efficient.

The danger of inefficient breach yet intensifies when the contractual accident takes a different form than pure "impossibility." In cases of mutual mistake, misunderstanding, impracticability and frustration of purpose, the costs and benefits of performance are typically unverifiable. In any event, there is little reason to believe that they are more likely to be verifiable than in standard breach situations.

[25] For the seminal work on the holdup problem, see OLIVER E. WILLIAMSON, MARKETS AND HIERARCHIES: ANALYSIS AND ANTI -TRUST IMPLICATIONS (1975); Oliver E. Williamson, *"Transaction Cost Economics: The Governance of Contractual Relations,"* 22 J. L. & Econ. 233 (1979); Benjamin Klein, Robert Crawford & Armen Alchian, *"Vertical Integration, Appropriable Rents, and the Competitive Contracting Process,"* 21 J. Law. Econ. 297 (1978).

[26] The holdup problem functions as follows: At renegotiations, both parties realize that the relying party can reap the fruit of his sunk reliance investment only if a contract is re-formed, but not if renegotiation break down. In such a setting, the relying party has more to lose from failure to re-contract than if he had not already relied. That, in turn, places him at a bargaining disadvantage, which raises his private expected cost of reliance ex ante. To limit his bargaining vulnerability, the relying party is thus induced to lower his level of reliance inefficiently.

Once again, the promisee is induced ex ante to under-invest in reliance.

Finally, the rule of excuse may well generate an undesirable allocation of risk. As observed by Alan Sykes, the impact of the rule of excuse on the parties' relative exposure to risk will vary with the particulars of the case.[27] To see this, suppose initially that any event of non-performance is treated as a standard breach and consider the impact of changing the rule such that in some cases of non-performance the promisor would be excused. How would that affect the promisor's exposure to risk? The answer is “we cannot say." For on the one hand, the ability to walk away, liability free, in some states of the world, would obviously relieve her of some risk; but on the other hand, if the rule is changed in that fashion, the promisee would demand to be compensated for her ensuing disadvantage by modifying the price term in her favor. But that modification in price would in turn imply that the value of the promisee’s expectation interest would now rise. Hence, in the states of the world in which the promisor breaches and is not excused, the promisor's liability would rise correspondingly. That increase in liability would in turn raise her risk-bearing costs ex ante. The ultimate impact of the rule on the promisor's overall exposure to risk is therefore inconclusive: The rule causes losses to shift from states of non-performance in which an excuse is granted to other states of non-performance in which full expectation damages must be paid. Hence, if one seeks justification to apply the rule of excuse, then its impact on risk allocation cannot serve the desired purpose.

In sum, the rule of excuse fails to produce optimal results in any of the parameters examined: it places sub-optimal incentives for precautionary actions and insurance (contrary to proposition (a)); it encourages inefficient reliance (contrary to proposition (b)); it was shown to generate inefficient trade; and it generates no obvious advantage in the allocation of risk.

Failure to Establish Non-Optimality in Non-Accident Settings: The observations set forth above should be sufficient to cast serious doubt with regard to the validity of existing justifications of the doctrine. However, as discussed next, the failure of the theory stems from an even more fundamental reason. That failure, in particular, undermines the validity of propositions (c) and (d) as well.

When a party defaults, his failure to perform is generally labeled a “breach.” The associated outcome of “breach” is generally the assignment of liability by the measure of the promisee’s expectation

---

[27] Alan O. Sykes, *"The Doctrine of Commercial Impracticability in a Second Best World*," 19 J. Legal Stud. 43 (1990)

interest. Yet, in some cases, the failure to perform is labeled an "accident", implying that no liability is assigned other than restitution. To provide a justifying theory of contractual accidents one must therefore establish two related propositions. First, that, under certain circumstances, the rule of excuse outperforms the rule of expectation damages and should therefore be preferred; and second, that there is at least a close correlation between the cases in which excuse is the superior rule under the theory, and those in which it is in fact instructed by legal doctrine.

Existing literature has focused on the former proposition, while neglecting the latter almost entirely. That is not surprising, for even if it did seek to establish the latter, it would not be an achievable task. The existing theory is indeed correct in the claim that excuse could in some instances be superior to expectation damages. However, these instances are not those in which excuse is actually applied.[28] Hence, while existing theory does shed light on the virtues of the rule excuse in a general sense, it does not provide reason to support the manner of its application. Consequently, it cannot be seen as a theory that supports the actual doctrine of contractual accidents.

This point is easily exemplified when considering the remaining justifications, (c) and (d). Inasmuch as these justifications are valid on their own terms, they are equally valid with respect to any event of

[28] Apparently, for that reason Michelle White has suggested to abolish the very distinction between accident and breach. See Michelle White, *"Contract Breach and Contract Discharge due to Impossibility: A Unified Theory,"* 17 J. Legal Stud. 353 (1988). A similar policy prescription has been suggested by George Triantis, although for a different set of reasons. George Triantis, *Contractual Allocations of Unknown Risks: A Critique of the Doctrine of Commercial Impracticability*, 42 U. Toronto L.J. 450 (1992). Triantis argues that even when certain contingencies are not specifically contemplated, they are nevertheless contractually allocated. Thus, for instance, even if a shipper and a carrier fail to consider the particular possibility that delivery would be made impracticable due to a closure of the Suez Canal, they always contemplate some broader category of risks that includes the realized contingency, such that some misfortune would prevent the carrier from reaching its destination as planned. The contract, Triantis argues, allocates this broader category of risk even if the specific possibility of the Canal's closure is not specifically contemplated. As all risks are thereby taken into account and contracted for, he concludes, contracts should always be enforced. For a similar claim, see also Eric A. Posner, *"Economic Analysis of Contract Law after Three decades: Success or Failure?"* 112 Yale L.J. 829, 845-849 (2003). This argument raises an evident difficulty, however. It is that the risk associated with the broader category cannot be correctly assessed unless all individual events are specifically contemplated. Hence, although one might accept the contention that all risks *could* be allocated, it does not follow that it must be optimal for the parties to do so if they are non-contemplated. In fact, the point of this paper is to establish that the welfare of the parties is enhanced if the governing rule allows them to allocate only contemplated risks.

non-performance, regardless of whether one labels it an "accident" or not. Hence, the theory cannot explain why excuse is allowed only in a case of accident but not in the case of breach. Thus, for example, recall proposition (c), which pertains to the rule's desirable impact on damage mitigation. The argument is valid in the sense that the rule of excuse does indeed produce an optimal incentive to mitigate, and hence, other things equal, it increases the available gains from trade in the world of the standard model. But damage mitigation is of course a significant consideration not only when non-performance is due to an "accident," (however one defines it in this context) but also when it follows an ordinary decision to breach. Clearly, if, in a particular case, the interest to control the parties' incentive to mitigate overrides other competing interests (such as optimal precautions, trade, reliance and allocation of risk), then the rule of excuse ought to be applied regardless of the causes that have led the promisor to default. However, that is hardly what the doctrine actually does. If an accident does not occur, a promisor cannot successfully argue that he ought to be released from a contractual commitment only to fulfill the objective of inducing optimal mitigation by the promisee.

The same point equally applies to proposition (d), pertaining to the choice of an optimal mode of performance.[29] One could indeed argue that under certain circumstances, the purpose of ensuring that an efficient mode of substitute performance be taken ought to be assigned definitive weight. However, from the lens of legal doctrine, that consideration, in and of itself, could never supply a legitimate reason to renege on a contractual obligation when an accident had not occurred. A promisor cannot ask to walk away scot-free from a contractual promise only to make sure that the efficient mode of substitute performance is ultimately chosen by the promisee.

In light of these difficulties, the adherer of the standard model must make the following inevitable choice: If, for some reason or another, he believes that arguments (c) and (d) ought to be assigned definitive weights in accident settings, then he must recommend that the rule be applied universally, even in cases where the doctrines of contractual accidents do not apply. And if he alternatively believes that propositions (c) and (d) are of secondary importance, and hence in cases of ordinary breach promisors must be liable to disappointed promisees, then he must also object to the rule of excuse in accident settings under similar circumstances. Either way, he cannot explain why the occurrence of an accident in its doctrinal sense should be

[29] See Goldberg, supra note 21.

deemed a relevant consideration in the determination of the appropriate remedy.

Viewing the matter more generally, it is evident that the same critique would equally hold with respect to any conceivable justification of the rule of excuse, as long as it relies on the assumption of perfect contemplation. The reason is simply that if all contingencies are similarly contemplated, there is no basis on which to distinguish between different acts of non-performance. Under the standard model, different remedies affect different incentives, and the applied rule must strike a balance between them according to their relative importance. That approach accords no weight to the label of non-performance as an "accident" or a "breach," and, in fact, it ascribes no definitional sense to that distinction. In the absence of a valid measure for distinction, a good argument raised in favor of the rule of excuse has a general appeal, and is applicable to cases of ordinary breach as well. [30]

So what exactly is in need of revision? Is it the rules of law that govern accidents, or is it the model we use to theorize about those rules? In the remainder of this paper I explore the latter possibility, and find that when the assumption of perfect contemplation is relaxed,

---

[30] Accordingly with the above conclusion, the economic analysis of mutual mistake has not been more successful in explaining the rule of excuse. Smith and Smith have shown that the doctrine of mutual mistake viewed in combination with the doctrine of unilateral mistake might promote disclosure of private information. It should be observed, however, that such an information forcing quality constitutes a relevant consideration only when a mistake is not truly mutual, for if it were, information would be symmetric to begin with. It therefore remains an open question whether the doctrine is ever independently desirable when information is absent on both sides. See Janet K. Smith & Richard L. Smith, *Contract Law, Mutual Mistake and Incentives to Produce and Disclose Information*, 19(2) J. Leg. Stud. (1990) 467.

The difficulty to justify the mutual mistake doctrine has been further demonstrated by Ian Ayres and Eric Rasmussen. Their analysis too, in most part, considers the issue of disclosure, and is therefore not pertinent to mutual mistake. They do have one main argument, however, that does rest on the premise that parties share a common misperception. Namely, they observe that in a case of scrivener's error, and in the absence of reliance, granting an excuse allows the parties to enter renegotiations and recreate their contractual understanding as they had originally intended it. That, in turn, is desirable to both parties ex ante as it relieves both of them of the risk that such an error would frustrate their original intentions. While the argument is persuasive, it is rather narrow in scope. For in general, unlike the case of a scrivener's error, an accident causes parties to see the world differently in a material respect, and therefore adjust their valuation of the transaction. Hence, renegotiations do not typically mimic the parties' original intent in the first contract but instead recreate an entirely new contractual agreement. That argument, therefore, does not provide a general justification for the doctrine. See Ian Ayres and Eric Rasmussen, *Mutual and Unilateral Mistake in Contract Law*, 22 J. Leg. Stud., 309 (1993).

theory does in fact provide substantial justification for the existing legal practices. The modified model is thus presented next.

IV. The Modified Model

The modified model admits the possibility that the parties might not conceive of a potential eventuality. Despite their awareness to the possibility of accidents, they cannot affect either their likelihood or their impact through deliberate action. However, they may nevertheless form preferences with respect to the legal consequences that should attach to accidents. This Section examines the attributes of an accident rule that most closely satisfies the preferences of most contracting parties with respect to those legal consequences.

The analysis proceeds in two stages. In the first, it is demonstrated that when several assumptions are satisfied, the rule of excuse is endowed with several major advantages over any alternative. In the second stage these assumptions are gradually relaxed, and the rule's optimality features are reconsidered. It is then observed that the rule remains optimal only in some accident settings but not in others. The analysis thus provides an insight into the criteria that determine the rule's desirability in each particular case. These normative results are found to correspond to the actual practice of courts, and therefore serve as a justifying theory of the law.

1. The First Stage

The setting considered in the first stage of the analysis, incorporates the following assumptions: (1) The parties do not engage in reliance; (2) If a contract is discharged, the parties can renegotiate a new contract at no cost; and finally, (3) If an accident-related dispute reaches litigation, courts are able to discern whether the realized risk had been contemplated or not. As mentioned above, all three assumptions are later relaxed. They are made at this stage for their simplifying quality only.

A. Mapping Intentions onto Outcomes

One of the most fundamental functions of a contract is to effect an accurate translation of the parties' desired intentions into a set of legally binding rights and obligations. The parties wish their contract to regulate their cooperative behavior in a manner that both maximizes their joint gains from trade, and distributes those gains according to their agreed-upon division rule. In the world of the standard model, designing a contract that accurately maps those intentions onto legal

outcomes is a trivial matter: The parties create the optimal legal framework simply by stating their desired outcome for each possible state of the world. As the state space is itself common knowledge, and as it includes the state of reality by assumption, the contract can perfectly reflect the parties' wishes.

Less obviously, that fundamental feature of the standard model remains intact even if the contract does not assign an outcome for each state of the world due to positive drafting costs. To see this, consider the objectives of efficiency and distribution in turn. From the perspective of efficiency, drafting costs produce an exogenous constraint on the feasible gains that are attainable. Some terms that would enhance efficiency if included are nevertheless omitted due to the costs of drafting them. The parties, however, respond rationally to that constraint by omitting a term only if its efficiency value is less than the cost of stipulating it. The joint gains from trade are therefore still maximized within the boundaries of feasibility. The contract is therefore as efficient as it could possibly be.

The objective of distributing gains according to the agreed-upon division rule is similarly preserved. The parties, who are assumed to possess a computational ability that is both costless and perfect, fully comprehend the distributive effect generated by any omitted term. Thus, even if due to high drafting costs a particular term is excluded, the distributional implications of that omission are recognized and can be fully undone by a parallel adjustment in price. The distribution of gains is thus maintained as well. The standard model guarantees that there would be no divergence between the outcomes the parties wish to assign to each state of the world, and those that the contract ultimately produces.

This attribute of the standard model is no longer sustained, however, under the premises of the modified framework, in which the parties know ex ante that they might fail to contemplate the state of reality. When reality is in fact not contemplated, enforcement of the contract's instructions might lead to very different outcomes from those originally intended. The parties' attempt to divide gains in the agreed-upon manner may well transpire to be unsuccessful. The contract becomes an imperfect device for mapping intentions onto outcomes, and the parties know it to be so.

That attribute of the contractual form is of course undesirable to the parties: If the divergence between intentions and outcomes could be avoided, the parties ex ante would by definition wish to avoid it. Accordingly, if a particular accident rule minimizes that divergence, they would be inclined to adopt it.

Given the assumptions made in this part of the analysis, it is

easy to observe that the rule of excuse facilitates the desired result. When a contract is discharged upon the occurrence of an accident, the parties are legally committed to the contract only under a subset of the possible states, namely those that they had explicitly contemplated ex ante. Obviously, for that subset of events, the contract serves as a perfect mapping device from intentions onto legal outcomes, for the same reasons for which this observation holds under the assumptions of the standard model. Somewhat less obvious is that the rule of excuse also facilitates the desired result indirectly when the realized contingency had not been contemplated. For when the contract is discharged, the stage is set for renegotiations, in which the parties are placed in a position that most closely approximates that which they have occupied originally, at the time of contracting. As neither side is liable for damages or entitled to it, the parties are pulled back to square one (recall that, for the time being, we assume that there is no reliance.) The outcome of bargaining in renegotiations is thus likely to mimic the result that would have obtained in the original round of negotiations, had the realized state been contemplated at the time.

It is important to observe that this attribute of the rule of excuse is shared by no other imaginable remedy. For under any alternative rule, non-performance is accompanied by payment of damages, typically from the non-performing promisor to the disappointed promisee. Hence, under any alternative, the promisor's gains following renegotiations are less than what he could have captured at the original round of bargaining, had the accident contingency been contemplated. Accordingly, the parties' hypothetical wishes with respect to the consequences of non-contemplation are not fulfilled. It follows that if such an alternative were assigned as a default, the parties would wish to opt out of it in favor of the rule of excuse.

B. Trade of Unquantifiable Risk

General: Suppose you were offered to participate in a gamble you knew nothing about. It might yield a gain, it might produce a loss, and you otherwise have no information regarding its possible outcomes or their corresponding probabilities. Would you find such an offer attractive?

The risk arising from non-contemplated contingencies is unquantifiable, just as in the proposal described above. The parties cannot meaningfully evaluate either its associated probability or the set of possible consequences that might transpire. Accordingly, the question this Section addresses is whether parties generally wish to engage in the trade of such unquantifiable risk, in addition to the

anticipated, calculable risk that is the subject of their agreement. It is observed that the choice between the remedies of excuse and expectation damages presents precisely that question: As elaborated below, the remedy of expectation damages entails trade of all risk, quantifiable or not; conversely, the rule of excuse allocates only contemplated, measurable risk, while keeping the incalculable kind outside the boundaries of contractual obligation.

To begin, it should be emphasized that the question of whether trade of unquantifiable risk generates a benefit or a cost cannot be determined on the basis of the rationality assumption per se. Rational preferences are formed with respect to known alternatives, not with respect to those that are unknown. When one is walking entirely in the dark, one cannot regard a step to the left as any more or less rational than a step to the right. Thus, in order to address that question one must resort to reasoning based on something other than rationality alone.

In the present context, I suggest two primary reasons for which the desirable default rule is one that does not include the trade of unquantifiable risk as an integral part of the contract. First, it is observed that when the attributes of a traded risk are unknowable, the transaction yields no discernable surplus in either an absolute or an expected sense. But while no visible value is generated, there are certain costs associated with entering such a deal, as the terms of trade must be determined through a costly process of bargaining and contracting. Second, a considerable body of empirical evidence suggests that decision-makers are indeed reluctant to enter agreements for trade of unquantifiable risk. That result is independent of the parties' risk preferences, as it applies even in cases where the party who would be relieved from the risk is unquestionably more averse to standard risk than the party who would assume it.

I next turn to discuss these two arguments in greater detail.

Avoiding Trade of the Unquantifiable Risk of Accidents: A contract governed by an accident rule of expectation damages could be viewed as consisting of two separate sub-agreements. The first specifies the terms that the parties wish to govern contemplated states, such as when the cow indeed transpires to be barren, as originally assumed; and the second concerns the formation of entitlements in accident scenarios, and is thus, in essence, a contract for the transfer of unquantifiable risk. In particular, the second sub-agreement provides that the parties must follow the contractual instructions literally even if they entail a division of entitlements that materially departs from their mutual intention ex ante. The promisee thereby assumes the risk that his money would buy something other than what he intended to acquire, and the promisor, likewise, assumes the risk of incurring a

different cost than anticipated to satisfy his obligation to perform. Although both parties find themselves engaged in a different transaction from that which they had anticipated, they remain committed to it as if the accident scenario were entirely compatible with what they had originally envisioned.

Whereas the rule of expectation damages entails the formation of both sub-agreements - the acquisition of the anticipated good and the transfer of accident risk - the rule of excuse entails only the formation of the former. When an excuse is allowed, the parties' exchanged obligations become binding only on the condition that the realized state had been contemplated. Depending on whether the value of the traded good rises or falls relative to the contemplated state, either the seller or the buyer chooses to withdraw from the deal by seeking an excuse. For example, in *Sherwood*, the buyer withdraws from the contract if the cow is less valuable than anticipated whereas the seller does the same if it is more valuable. The seller thus resumes his ownership of the cow in all accident states, and is thus the one to absorb the accident's impact on value for better or worse.[31] In contrast, under expectation damages, the cow must be delivered and the payment must be rendered notwithstanding the accident, and hence the risk of non-contemplated changes in value is shifted to the buyer.[32,33] The question is then which of those allocations of accident risk better corresponds to the wishes of

[31] Note that the economic definition of risk, as used here, may be inconsistent with the common usage of the term. In standard language, it may be said that under the rule of excuse Sherwood is insured, as he is not required to give up a cow that is worth more than the price for which it was sold. That is not so under the economic definition of risk, however. For under the economic definition of the term, being insured implies that one's payoff remains constant regardless of the realized state of the world. But if in an accident state Sherwood is allowed to avoid the deal and retain possession of the more valuable cow, his payoff in the accident scenario is higher than his payoff in the anticipated state. That variance, in and of itself, is a source of risk rather than a mechanism of insurance.

[32] For an accident concerning the cost of performing an obligation, consider the case in which a closure of the Suez Canal prevents a carrier from delivering goods from the United States to Iran through the planned route. The first agreement specifies the terms of trade under the contemplated state, in which the Canal can be accessed. Absent the second agreement, the promisor's obligation would not have been binding in a non-contemplated state, and hence the shipper would be the one to bear the accident risk. The second agreement however provides that the obligation undertaken by the carrier to deliver in a non-accident state must also be satisfied if an accident occurs. Hence, in that case as well, the rule of expectation damages shifts accident risk from one party to the other.

[33] Note that expectation damages situate the promisee at the same position he would have occupied absent the breach. Hence, if Sherwood chose to breach, he would have to pay Walker the price of a fertile cow, not a barren one. That, in turn, would imply that the option of paying damages in lieu of performance would not be exercised.

most contracting parties ex ante. Is it more likely that parties wish their contract to transfer unquantifiable risk to the buyer by default, or do they rather have it remain with the seller?

When risk is quantifiable, transferring it from a party who is more averse to risk to a party who is less so can always create a Pareto improvement. Namely, there exists a price for the provision of such insurance, which allows both parties to benefit from the transaction. Conversely, when risk is unquantifiable, the terms of trade that would enhance the welfare of both parties cannot be ascertained. As the risk of an accident is unknown, the parties can only assign a price arbitrarily, in a manner that may or may not be welfare enhancing. The transaction, therefore, can no longer be assumed to bear the qualities of an insurance contract or create a Pareto improvement. In turn, there is little reason to assume a-priori that most contracting parties wish to engage in such transactions by default. Such a conclusion certainly does not follow from the parties' mere interest in contracting for the anticipated good.

Furthermore, to execute the trade of unquantifiable risk, the parties must exert considerable effort. They must formulate a negotiation strategy and engage in costly bargaining until they reach agreement about the terms of trade. Taking that course therefore seems to involve a potentially substantial cost without producing a visible benefit. The rule of excuse, on the other hand, relieves the parties from that imposition and allows them to restrict their contractual relationship only to states of the world whose value is assessable.

To further illustrate this latter point, consider the following simple example. Suppose that a buyer and a seller negotiate the sale of a widget, which they believe would cost 50 for the seller to produce, and generate a value of 100 to the buyer. Suppose further that the parties wish to divide the surplus equally. In an accident-free world, they would simply set the price at 75 and thereby accomplish their desired allocation. However, in reality, they acknowledge that the cost and valuation might change due to an accident. If the contract remains in full effect under an accident contingency, it follows that the price of 75 must be either lowered or raised for the equal division of surplus to be restored. But in order to derive the appropriate modification, the parties must initially agree on the expected accident impact and probability. As that is unknown, however, it is a task they cannot meaningfully perform.

This difficulty is wholly averted under the rule of excuse.[34] Namely, if upon the occurrence of an accident each party goes his own way without bearing further liability, then setting the price at the level of 75 is consistent with the parties' preferences *regardless* of the true accident probability or impact. The buyer must only determine his reservation value of the contemplated widget, and can free himself from assigning an evaluation to the different widget that it might turn out to be. If an accident eventually transpires he would no longer hold an entitlement to the acquired commodity and would no longer carry an obligation to render the contractual price. The seller too needs only to determine the value of the contemplated widget, since in an accident contingency he resumes ownership of it. The rule of excuse therefore saves the parties the cost of negotiating an arbitrary price for an unquantifiable measure. Whereas risk is still borne by one of the parties, it does not play a role in the contractual negotiations.

From this perspective as well, the rule of excuse is not merely superior to that of expectation damages but also the unique accident rule generating the desired result. The reason is that it is the only rule relieving the parties from all rights and obligations under the contract. Under any alternative, an accident would necessarily require that some transfer be made from one party to the other, whose expected value would need to be assessed ex ante. But in order to determine that value, the parties would once again need to engage in the exercise of evaluating the non-measurable risk of an accident.

Empirical Evidence: The disinclination of most individuals to trade unquantifiable risk is also supported empirically. Both in laboratory experiments and in the field, parties exhibit reluctance to trade risk that is unknown or substantially ambiguous. That is so even in contexts in which trade would clearly be desirable if risk were quantifiable, as the potential insurer is less averse to calculable risk than the potential insured.

The insurance market is a fertile ground for examining this hypothesis. As the exchanged commodity in this market is pure risk, it is a setting in which the desirability of risk-shifting can be examined in isolation from other possible motivations to contract. In line with the stated hypothesis, lack of risk predictability is indeed widely recognized as reason for which insurance is not voluntary provided.[35,]

---

[34] In the presence of reliance investments, that statement must be qualified. See Section V.1 *infra* for a discussion.

[35] See e.g., BARUCH BERLINER, LIMITS OF INSURABILITY OF RISKS (1982) (analyzing the various reasons for which insurance might not be voluntarily provided, among them the unpredictability of risk); Goran Skogh, *Development Risks, Strict Liability and the Insurability of Industrial Hazards*, 87 Geneva Papers on Risk and Insurance, 237

[36] When evidence regarding risk is lacking or insufficiently reliable, insurance companies are found to either refuse to offer coverage, or to demand exorbitantly high premiums that only few consumers are willing to pay.

Terrorism risk in the United States is one recent, prominent example of that phenomenon. In the wake of the attacks of September 11th 2001, insurance companies have found the risk arising from future acts of terror incalculable, and as a result refused to continue offering coverage for such risks. To reinstate the provision of terrorism insurance, Congress soon thereafter enacted the Terrorism Risk Insurance Act of 2002. The Findings and Purpose section of that Act accordingly explains that "widespread financial market uncertainties have arisen following the terrorist attacks of September 11, 2001, including the absence of information from which financial institutions can make statistically valid estimates of the probability and cost of future terrorist events, and therefore the size, funding, and allocation of the risk of loss caused by such acts of terrorism." The Act thus provides that insurance would be offered by the government, alongside a mandatory requirement of private insurers to offer some coverage despite their unwillingness to do so voluntarily.[37]

Additional examples of unpredictable risks for which insurance is under-provided are abundant. Among them are insurance for environmental pollution,[38] of losses arising from climate change,[39] of

---

(1998) (explaining the risk associated with certain industrial hazards is unquantifiable and thus uninsurable.); Howard Kunreuther & Robin M. Hogarth, *How Does Ambiguity Affect Insurance Decisions?*, in CONTRIBUTIONS TO INSURANCE ECONOMICS (GEORGES DIONNE ed., 1992) (providing experimental support to the claim that ambiguity in the measure of risk explains insurance firms' reluctance to offer insurance); Howard Kunreuther et al., *Insurer Ambiguity and Market Failure," 7 Journal of Risk and Uncertainty*, 71 (1993) (reporting experimental results indicating that demanded insurance premiums rise sharply when probability is ambiguous; Michelle E. Boardman, *Known Unknowns: The Illusion of Terrorism Insurance*, Geo. L.J. 783 (2005) (describing the risk of terrorism as a "known unknown," rendering it impossible to insure in a non-regulated market.) Goran Skogh, *Risk-Sharing Institutions for Unpredictable Losses*, 155 Journal of Institutional and Theoretical Economics, 505 (1999) (identifying alternative methods of coping with unpredictable risks when insurance is unavailable);

[36] That is so despite the fact that in a free market the potential insurer is by definition less averse to risk than the potential insured. In the context of contractual accidents, reluctance to shift accident risk to the promisor should be even more forceful, as the promisor is not thought to be systematically less averse to risk than the promisee.

[37] For further discussion, see Boardman, supra note 35 and Jane Kendall, *The Incalculable Risk: How the World Trade Center Disaster Accelerated the Evolution of Insurance Terrorism Exclusions*, 36 U. Rich. L. Rev. 569 (2002).

[38] See Howard Kunreuther, *The Role of Actuaries and Underwriters in Insuring Ambiguous Risks*, 9 Risk Analysis 319 (1989) (explaining the unavailability of

certain industrial hazards,[40] of certain chemical and nuclear disasters,[41] of auditor liability,[42] of floods[43] and of losses associated with instable political regimes in developing countries.[44] In all of these cases, inability to quantify risk is regarded at least as a significant factor that prevents a fully-fledged insurance market from emerging.

This tendency has also been clearly observed in controlled experiments. In one particularly relevant one, reported by Robin Hogarth and Howard Kunreuther, MBA students at the University of Chicago were asked to state their reservation values with respect to an insurance contract.[45] The stimulus concerned an owner of a small business who considered purchasing insurance against a loss of $100,000 due to claims concerning a defective product. Subjects were asked to consider various alternative probability values for the loss, namely 0.1, 0.35, 0.65 and 0.9. In one group, subjects played the role of the business owner and asked to state their willingness to pay for insurance under each probability measure, whereas in the second group, they played the role of the insurance company, and asked to state their minimal supply price. To compare reactions to those questions as a function of the weight of available information, the statement of each

---

insurance against environmental pollution by the ambiguity associated with that risk.) See also Kenneth S. Abraham, *Environmental Liability and the Limits of Insurance*, 88 Columbia Law Review, 942 (1988) (discussing the unpredictability of liability imposed for harm associated with environmental pollution.)

[39] Richard S.J. Tol, *Climate Change and Insurance: A Critical Appraisal*, 26 *Energy Policy*, 257 (1998) (arguing that risks associated with climate change are unquantifiable and therefore uninsurable.)

[40] See Skogh, *supra* note 35.

[41] See, e.g., Martin T. Katzman, Pollution *Liability Insurance and Catastrophic Environmental Risk*, 55 Journal of Risk and Insurance, 75 (1988) (chemical disasters are exceedingly rare, and hence insurers lack historical data on which to base a probabilistic prediction. Disasters that occurred in the past often reflect an outmoded safety technology. Finally, non-historical methods of risk analysis provide little further insight as to their objective probability).

[42] Peter Moizer and Lisa H. Smith, *UK Auditor Liability: An Uninsurable Risk?*, 2 Int. J. Audit, 197 (1998) (explaining that difficulty to insure arises from lack of evidence regarding the magnitude of risk);

[43] Inability on the part of private insurers to estimate the risk of floods has triggered the enactment of the National Flood Insurance Act of 1968, under which the government undertakes the role of an insurer.

[44] See Howard C. Kunreuther & Paul R. Kleindorfer, *Insuring Against Country Risks: Descriptive and Prescriptive Aspects*, in MANAGING INTERNATIONAL RISKS (R. HERRING, ed., 1983). The problem of uninsurability has been mitigated through the provision of government insurance through the Overseas Private Investment Corporation. See 22 U.S.C. 2191-2200a (2000).

[45] Robin Hogarth and Howard Kunreuther, *Ambiguity and Insurance Decisions*, 75 The American Economic Review 386 (1985).

probability measure was accompanied by an additional comment regarding the decision-maker's level of confidence with respect to the stated probability. In some cases subjects were told that they "felt confident" that the probability truly reflected the objective likelihood of the triggering event (the "non-ambiguous case"), whereas in others, they were told that they "experienced considerable uncertainty" with respect to the true probability value (the "ambiguous case").

In the non-ambiguous case, the median willingness to pay was either greater or equal to the median supply price under all probability measures. Hence, in all cases, one would expect that insurance contracts would be formed under market conditions, as suggested by the general analysis of insurance when risk is quantifiable. In the ambiguous case, in contrast, the buyers' median willingness to pay was found in all cases to be lower than the sellers' median supply price. Hence, under those conditions, if the subjects were to meet and negotiate the trade of risk, an insurance market would either not be created at all, or at most be a very thin one. The result was significant, with a probability of only 0.0004 that it could have been derived by chance.

This result is not unique to this particular experiment, and has been replicated in additional studies, which varied both the empirical technique and the identity of the subjects. In particular, consistent results were obtained when the subjects were professional actuaries, underwriters and reinsurers.[46]

[46] In a similar vein, Laure Cabantous asked 78 French insurance underwriters to price a premium for two types of risk: In the first, a loss of L would occur with a probability of 0.2%, whereas in the second a loss of the same magnitude would occur with a probability of either 0.1% or 0.3%, each with equal likelihood. As the two risks are actuarially the same, one would expect the premium to be set identically. However, the underwriters priced the second risk at a much higher rate: Whereas the first risk was priced with a loading factor of 35% of the actuarial value, the second was priced at more than twice that amount, at 78%. See Laure Cabantous, *Ambiguity and conflict aversion in the field of insurance: Insurers' attitude towards to imprecise probabilities*, unpublished manuscript, University of Toulouse (2003). See also Howard Kunreuther et al., supra note 35 (finding that actuaries, underwriters and reinsurers recommend considerably higher premiums to be charged for ambiguous risks, and some express outright reluctance to offer any insurance for such risks); For similar, additional results see Robin M. Hogarth and Howard Kunreuther, *Risk, Ambiguity and Insurance*, 2 Journal of Risk and Uncertainty (1989); Robin M. Hogarth and Howard Kunreuther, *Pricing Insurance and Warranties, Ambiguity and Correlated Risks*, 17 Geneva Papers on Risk and Insurance Theory, 35 (1992). See also John C. Hershey et al., *Sources of Bias in Assessment Procedures for Utility Functions*, 28 Management Science 936 (1982) and Richard Thaler, *Towards a Positive Theory of Consumer Choice*, 1 Journal of Economic Behavior and Organization, 39 (1980).

To sum up, the rule of expectation damages, under which the trade of a contemplated good is necessarily accompanied by the transfer of an unquantifiable risk, is not likely to trace the preferences of most contracting parties. Not only does it allocate that risk to a party who has no visible advantage of sustaining it, but it also forces the parties to assign it an arbitrary value and exert effort in reaching agreement for its exchange. Moreover, the reluctance to trade unquantifiable risk is demonstrated empirically, with evidence suggesting that trade of such risk generates disutility, not positive value.

C. Optimal Decisions to Contract

One of the most prominent objectives of policy in the realm of contract law is to encourage parties to trade whenever a positive surplus can be produced. Under the assumptions of the standard model, the parties possess the necessary information to accomplish the desired outcome. As states and probabilities are known, the parties can evaluate the prospect of gains, compare it against the risk of losses, and proceed to contract if and only if the former outweighs the latter. If some of that information is missing, however, this balancing procedure does not yield a definitive result. The desirability of contracting can therefore not be fully ascertained.

An obvious question then arises as to how beneficial opportunities could be identified, and how legal policy could be used to facilitate contracting when trade could produce positive gains. To be sure, an accident rule, be it what it may, does not produce information that is otherwise unavailable. However, as it determines the consequences of accidents, it does affect the parties' contracting incentives. From this perspective as well, the rule of excuse is endowed with desirable qualities.

To amplify, suppose that two parties, Aaron and Betsy, consider whether to enter a contract. They are aware that an accident might occur, and that if it does, trade might transpire to be undesirable, as performance would generate more cost than benefit ex post. The probability of such an accident, however, remains unknown.

In the anticipated, non-accident scenario, both parties of course expect to gain. Let *Ga* denote Aaron's expected surplus given that the realized state is contemplated, and define *Gb* similarly for Betsy. Let *p* and *q* denote the (unknown) probabilities that Aaron or Betsy respectively would eventually fail to perform due to an accident.

Let us initially assume that accidents are not recognized, and so in the event of an accident a defaulting promisor is liable for

expectation damages. Given this assumption, the parties' overall expected gains from forming the contract are *(1-p)Ga-pGb* for Aaron, and *(1-q)Gb-qGa* for Betsy. Observe that each party's payoff consists of both a positive element and a negative one. The relative magnitude of the two elements depends on the value of *p* for Aaron and of *q* for Betsy. Namely, as the probability of an accident rises, the promisor's expected gains fall until they become negative after surpassing a certain threshold value.[47] The parties thus choose to enter the contract only if they form some optimistic belief regarding accident risk, under which the overall value of the expression remains positive.

For Aaron, the prospect of a loss is all the more likely the larger Betsy's anticipated gain of *Gb* and the lower is his anticipated gain of *Ga*. For if Betsy's anticipated surplus is large, he is liable for more in damages in an accident scenario; and if his own surplus is small, the anticipated benefit may not be worth the potential loss. For Betsy, the contract is more likely to be a losing prospect under the opposite conditions. As the parties' uninformed assessments of risk may plausibly take many different values, and may differ from one another, one expects that in certain cases they would choose not to enter the contract. But observe that despite these possible disincentives to contract, as long as the probability of an accident is greater than zero, trade is socially optimal *always*. The prospect of a loss that inhibits contracting arises out of transfer payments, not social costs. Hence, any failure to contract due to the private risk of incurring them amounts to a wasted opportunity to create value. It follows, in turn, that under the rule of expectation damages, the decision to trade may be inefficient.

This drawback of sub-optimal trade is straightforwardly avoided under the rule of excuse. For if the parties simply part ways upon the occurrence of an accident, Aarons' expected gains become *(*1-*p*-*q)Ga* if 1-*p*-*q*>0 and 0 otherwise; and similarly, Betsy's expected gains are given by *(*1-*p*-*q)Gb* if 1-*p*-*q*>0 and 0 otherwise. Those payoffs contain no negative element, and hence they also guarantee to each of the parties a private gain that is non-negative. As the contract requires no transfer payments to be rendered, it shields the parties from potential losses. The parties are thus mutually motivated to enter the contract regardless of the incalculable values of *p* and *q*. The condition of optimal decisions to contract is thereby restored.

The underlying intuition of this result is a simple one. The rule of excuse allows the parties to become legally bound only in contemplated states. Hence, in accident scenarios, the promisee does not receive a windfall benefit and the promisor does not incur a

[47] Namely, *Ga/(Ga+Gb)* for Aaron and *Gb/(Ga+Gb*) for Betsy.

windfall loss. That unequivocally increases the likelihood of contracting when trade is efficient: For the promisor, contracting becomes all the more desirable, as he is no longer exposed to the risk of uncompensated losses; and for the promisee, contracting remains sufficiently desirable, as she too can still profit from the contractual opportunity if it is indeed efficient. While the rule of excuse does not create new information, it renders the incentive to contract independent from the unknown value of accident probability. Contracting is thereby consummated if and only if it is efficient.

The rule of excuse, in this case as well, is not merely superior to the rule of expectation damages, but is also the unique rule to generate this desired result.[48] The reasoning is by now a familiar one: Any rule other than that of excuse would require some transfer to be made from one party to the other in the event of an accident. The expected value of such liability, in turn, would depend on the unquantifiable accident probability. For sufficiently high values of accident risk, the contractual prospect would yield a negative surplus to the liable party. The possibility of an efficient opportunity to trade that is nevertheless unconsummated, would thereby reemerge.

D. Mitigation of Damages

The fourth and final argument advanced here in support of the rule of excuse pertains to its impact on the promisee's incentive to mitigate damages. When a promisee learns of the error underlying the agreement, there might be means available to her to lessen her ensuing costs. Namely, she may cut losses by ceasing any ongoing reliance activity, by entering into a substitute contractual arrangement, or perhaps by engaging in new investments that have just become cost-effective given the discovered state of reality.[49] From a social perspective, such mitigation efforts are desirable as long as on the margin the loss they prevent exceeds their cost.

Under the rule of expectation damages, the promisee carries no incentive to engage in mitigation, as her entire loss is effectively insured by the non-performing promisor. The promisee's moral hazard thus generates inflated losses ex post, and diminishes the expected value of contracting ex ante. In some cases the problem of inadequate mitigation could be alleviated directly, by limiting recovery only for

[48] See, however, section V.1 *infra*, in which reliance investments are incorporated into the analysis.

[49] For a systematic analysis of damage mitigation see Charles J. Goetz & Robert E. Scott, *The Mitigation Principle: Toward a General Theory of Contractual Obligation*, 69 Va. L. Rev. 967 (1983).

losses that could not be efficiently avoided. Indeed, inasmuch as such a rule is practically enforceable, it generates the optimal result: By restricting the promisee's entitlement only to the portion of the loss she could not avoid, she is driven to internalize both the cost and the benefits of her mitigation efforts.

In reality, however, enforcing such a rule requires courts to verify information that is often not in fact available to them. The expected costs and benefits associated with mitigation often takes the form of private information, known to the promisee but unprovable to a court at reasonable cost. The court is thus not an effective monitor of the efficiency of the promisee's mitigation opportunities and is not well-positioned to cure the problem of moral hazard. It follows that in order to control mitigation incentives a different policy instrument is required, one that does not rely on the availability of evidence and the verifiability of information. The accident rule, if applied desirably, could serve precisely that end.

From this perspective as well, the rule of excuse functions optimally. When a contract is discharged upon the promisor's failure to perform, the promisee bears her own accident losses and appropriates the gains of her own mitigation efforts. As she internalizes both the costs and the benefits, her private incentive converges with the social one, causing her to invest in mitigation if and only if the benefit exceeds the cost. But unlike the solution of limiting recovery for mitigated losses, which requires the court to reach a positive factual finding, the optimal result under the rule of excuse does not hinge upon the verifiability of information.[50]

At first glance it might seem that this advantage of the rule of excuse applies equally to all instances in which a contract is not performed, not merely to those categorized as accidents. To be sure, after a decision not to perform has been reached, mitigation of losses is indeed desirable regardless of whether the realized contingency had been originally contemplated or not. Consistently with that observation, several authors have suggested that the rule's impact on mitigation serves as a justification for existing doctrine, even under the premises

[50] It should be noted that the rule of excuse is not unique in its optimal effect on mitigation. Namely, any accident rule under which the promisee's entitlement for damages is independent from her accident losses similarly creates optimal incentives. Thus, for example, the desired result would survive under a liquidated damages rule, providing for a fixed compensatory amount that is independent of actual losses. Optimality would be then attained for a similar reason, namely that the promisee would bear the full cost of mitigation and appropriate its entire corresponding benefit.

of the standard model.[51] That argument, however, suffers from several major drawbacks. As explained in Section III, in non-accident settings, the rule's desirable impact on mitigation comes at a heavy price of encouraging inefficient breach, under-investment in reliance and inadequate investment in precautions against the realized risk. Thus, in order to argue in favor of a policy that allows a contract to be discharged in non-accident cases, one must advance the claim that the advantages of improved mitigation outweigh these various costs. Furthermore, if one contends that the court ought to perform this cost-benefit analysis, rather than simply discharge all contracts upon default, then one must presume that the court has sufficient means to determine the balance of this tradeoff in any particular case. Such a supposition seems highly dubious, however, especially given the fundamental reason for which a discharge is thought to be superior to a simple limitation on the amount of recovery, namely that the necessary information that courts need to adequately administer the rule is often unverifiable.

It is thus for good reason that in non-accident settings the rule of excuse has not become the standard remedy, and courts have not been granted the discretion to award it at all. It is for an equally good reason that it has become the standard remedy in accident settings, where the drawbacks of distorted incentives do not arise, while the benefits of improved mitigation remain intact. The advantage of the rule of excuse on mitigation incentives thus joins its other advantages in the accident setting, and provides additional support for its general application.

2. The Second Stage

A. General

Contractual accidents are to some extent endemic to all agreements. As the human cognition is imperfect, the consequences of human actions always depart to some extent from their prior imagined form. This is especially so when the imagined state involves a development that will materialize only in the future. Indeed, it is difficult to conceive of a transaction in which the future does not play a crucial role. Even in the simplest spot transaction, in which the deal is executed immediately and simultaneously, future developments are of vital importance. This is because the gains made from a traded thing are

[51] See, e.g., Bruce, supra note 16; Wagner, supra note 20; Goldberg, supra note 21; See also the discussion by Sykes, supra note 27 at 63-66.

derived from the thing's stream of future yield. Its present value, which is what a property right actually buys, is therefore never entirely knowable unless one holds the necessary information about probabilities and states.

And yet it is clear that not any divergence between the parties' prior beliefs and the actual state of reality warrants the annulment of all contractual obligations. To be sure, existing doctrine too applies the rule of excuse sparingly and with considerable caution.[52] The question is then what should be the relevant criteria to determine when the rule of excuse should be applied. What distinguishes an "actionable" accident, which frees the promisor from all future obligations, and a "non-actionable" one, in which he continues to fully bear the burden of contractual liability? What aspects of a particular case should be counted as reason to apply the rule of excuse and what should be counted as reason against it?

The analysis so far has abstracted from several issues that are pertinent to that question. Incorporating them into the framework provides reason to refrain from applying the rule of excuse in some accident settings. The foregoing arguments in favor of the rule of excuse remain intact when accidents are "severe," in the sense that enforcement of the contract would result in a major disparity between anticipated payoffs and actual ones.[53] On the other hand, these arguments are substantially weakened if the promisee's reliance is very substantial, if renegotiations are very costly, or if courts are highly

---

[52] See e.g., J. WHITE & R. SUMMERS, UNIFORM COMMERCIAL CODE 132-33 (2D ED. 1980); Robert A. Hillman, *"An Analysis of the Cessation of Contractual Relations,"* 68 Cornell L. Rev. 617, 650-655; Iowa Elec. Light & Power Co. v. Atlas Corp., 467 F. Supp. 129 (N.D. Iowa 1978).

[53] It is indeed a well-recognized principle that relief is granted only if the gap between the promisor's anticipated payoff and his actual gains is large. Thus, for instance, the commentary to the Restatement provides that: "Performance may be impracticable because extreme and unreasonable difficulty, expense, injury, or loss to one of the parties will be involved… [But] a mere change in the degree of difficulty or expense… unless well beyond the normal range, does not amount to impracticability since it is this sort of risk that a fixed-price contract is intended to cover." See RESTATEMENT SECOND §261 cmt. d. Similarly, the commentary to the Uniform Commercial Code provides that "[A] rise or a collapse in the market is itself [not] a justification, for that is exactly the type of business risk which business contracts made at fixed prices are intended to cover. But a severe shortage of raw materials or of supplies… which either causes a marked increase in cost or altogether prevents the seller from securing supplies necessary to his performance is within the contemplation of this section." UCC 2-615 cmt. 4. This principle also governs the doctrine of mutual mistake, under which relief is awarded only if the effect on the agreed exchange of performance is "material." See RESTATEMENT SECOND §152.

prone to error. The following paragraphs offer an explanation for why that is so.

B. Reliance

General: The first assumption to be relaxed is the one stating that the parties do not engage in reliance investments prior to the discovery of an accident. For reasons discussed below, the presence of reliance alters the general conclusions developed in the previous Section.

Reliance may come in various forms. Parties may take actions in preparation for the other party's performance; they may forego alternative contractual opportunities; or they may engage in new relationships that would yield value only upon the performance of the original contract. Such actions are characterized as relationship specific investments: They require either an out-of-pocket cost or an opportunity cost to be incurred before performance is rendered, but they generate the hoped-for benefit only if the contract is performed.[54]

Accidents may impact the value of reliance in differing degrees. In some cases, a reliance investment yields the same return regardless of whether an accident eventuates. Let us refer to such reliance as "transferable" (as the return for reliance can be "transferred" from the contemplated state to the accident state without loss of value). In most instances, however, the value of reliance declines in the event of an accident, and hence reliance is at least to some extent "non-transferable."

To illustrate this distinction, suppose that after concluding a sale of a tract of land, the parties discover that the lot is smaller than they had both initially believed. The buyer had already relied on the sale by hiring an architect to plan a house, which would be specifically tailored to the attributes of the land. That reliance investment is said to be "transferable" if despite the accident the house plan would still generate the same value to him as if it were as large as originally anticipated. Reliance might be non-transferable, however, for two primary reasons: First, in view of the accident, trade might no longer be desirable at all, as the buyer would value it by less than the seller. Assuming the plan could not be sold to a third party, its value would fall to zero. Second, even if trade is still desirable ex post, the reliance investments might be of lesser use. Thus, for example, the buyer might still wish to own the land, but his house plan would need to be amended or recreated to fit its

[54] If the parties relied without obtaining a contractual commitment from their counterpart, they would be exposed to holdup: once their reliance investments are made, and the costs of reliance become sunk, the other party would be in a better position to extract a larger share of the contractual pie through bargaining.

smaller size.

Whereas both types of reliance provide grounds to depart from the conclusions of the analysis without reliance, the underlying reasoning in each case differs. In the discussion that follows, the two types of reliance are therefore addressed separately.

Transferable Reliance: When reliance is transferable, the investment by definition yields its anticipated return if (and only if) trade is consummated despite the accident. Under the rule of excuse, therefore, the realization of that return depends on a successful renegotiation of a new contract. Such renegotiation, however, creates a "holdup problem," which, in turn, undermines the desirability of the rule of excuse.

To amplify, when a party expects to renegotiate the terms of the contract, his bargaining power depends on the extent to which he is already invested in that relationship. In particular, when a party has already sunk his reliance expenditures, he has more to lose from the breakdown of negotiations than if he had not already relied. As both parties acknowledge the bargaining vulnerability this entails, the non-relying party is in a position to hold the relying party up, and thereby extract a more favorable agreement. The ultimate outcome of bargaining is therefore likely to be skewed against the interests of the relying party.[55]

In light of such potential for holdup, the rule of excuse no longer generates a perfect mapping of intentions onto outcomes. As the promisor's bargaining position is stronger in the original round of negotiations than in the second, it follows in particular that the terms of trade in the second round are likely to differ from what they would have been had the accident contingency been contemplated. Under the rule of expectation damages, in contrast, the problem of holdup is avoided. A well-known feature of this remedy is that it induces promisors to perform if and only if trade is optimal ex post.[56] Under this rule, therefore, renegotiations do not occur in equilibrium, and hence the holdup problem does not emerge.

Of course, the general advantage of expectation damages in avoiding holdup must be weighed against its drawbacks, as discussed in the foregoing analysis without reliance. Namely, as accident losses are not allocated by the original contract, the rule of expectation damages

---

[55] See [xxx], supra note 25.

[56] As a breaching promisor must compensate the promisee for his foregone gains, he internalizes the entire social cost of breach and thus defaults if and only if trade is undesirable ex post. For a detailed analysis of this function of expectation damages, see Steven Shavell, *Damage Measures for Breach of Contract*, 11 Bell J. Econ. 466 (1980).

allows the accident to assign them arbitrarily, with no apparent relation to the parties' anticipated gains. Choosing the optimal remedy in a particular case therefore involves a tradeoff. It depends on the relative magnitude of the holdup problem on the one hand, and the severity of the accident on the other. Thus, for instance, if performance would result in a dramatic reallocation of gains, while the promisee's exposure to holdup would be relatively small, then the rule of excuse would remain highly attractive, for the same reasons set forth in the analysis without reliance. Conversely, if the opposite conditions obtain, the rationales underlying the rule of excuse would no longer hold.

Non-Transferable Reliance: When reliance is non-transferable, renegotiations can no longer salvage its anticipated return. Hence, as the cost of reliance transforms into a net loss, the accident rule determines its allocation. The question is therefore whether there exists a particular allocation of accident losses that fits in the mold of a majoritarian default rule.

The hypothetical agreement the parties would reach with respect to the allocation of losses is by and large a function of their relative bargaining power. That, however, highly varies among different contracting pairs, and in any event it is typically unverifiable to courts ex post. The discussion that follows therefore does not aim at identifying a majoritarian default rule in this case. Rather, the more modest objective is simply to make the point that the justifications offered for the rule of excuse in the previous Section do not apply as decisively when parties engage in non-transferable reliance.

To see why that is so, consider first the issue of mapping intentions onto outcomes. When the parties were assumed not to rely, the rule of excuse allowed them to resume their precontractual positions if an accident were to occur. Once in those positions, they were motivated to re-divide the potential gains from trade in a manner that mimicked their original intentions. In the presence of non-transferable reliance, however, those precontractual positions can no longer be resumed, as losses have already been irreversibly incurred. The potential pie has shrunk, and hence the division the parties would have selected initially could no longer be physically obtained. It follows that neither the rule of excuse, nor any alternative, can generate an accurate mapping from intentions onto outcomes.

Next consider the issue of trade of unquantifiable risk. If the rule of excuse is applied, then in the event of an accident, each party simply bears his own reliance cost, regardless of his capacity as promisor or promisee. It follows that, unlike the analysis with no reliance, one cannot evaluate the value of reliance or its desirable magnitude, unless one forms some estimate of accident risk.

It should be noted, however, that from this perspective, the rule of excuse does retain some advantage over its alternatives. For under the rule of excuse, each party is exposed exclusively to the risk associated with his own investment, and hence trade of unquantifiable risk is avoided.[57] That desirable result obtains under no other alternative rule. For under any such alternative, liability in accident states is a form of insurance against accident risk.[58] But clearly, as such insurance is costly to provide, it requires some incalculable risk to be traded for a price.

Finally, consider the issue of optimal trade. In the absence of non-transferable reliance, contracting is always desirable, as it always yields non-negative gains from trade. In turn, the rule of excuse encourages efficient contracting, as it induces the parties to enter contracts regardless of their evaluation of accident risk.

In the presence of non-transferable reliance, however, contracts can also yield losses, not only gains. The risk of accidents thus inevitably impacts the value of contracting, and the parties' estimates of such risk may well, in turn, impact their decision to contract. Depending on these estimates, the parties might decline to enter efficient contracts, and they might choose to enter inefficient ones.

As in the case of transferable reliance, the desirability of the rule of excuse in this setting depends on a tradeoff. It is more objectionable as the parties' losses from non-transferable reliance increase. It becomes more appealing, however, as the accident becomes more severe.

The Effect of Time: A particular implication of the foregoing discussion is worth explicit mentioning. As a promisee's reliance typically accumulates in time, the desirability of the rule of excuse will often also depend on the timing in which the accident is discovered. Thus, for example, one expects that if Sherwood and Walker discover the cow's true state only long after it has been delivered to the buyer, the force of the seller's legal claim would substantially diminish. The reasoning is simple: As the buyer assumes his ownership of the cow, he has all the reason to invest generously in enhancing its value, and to generally arrange his conduct around the notion that the cow is his to keep. Conversely, if the accident is discovered only, say, a few moments after the contract is formally formed, one expects Walker's

[57] More accurately, the produced estimate of accident risk affects the negotiated terms, only to the extent that it alters the parties' reservation values, through its effect on their evaluation of reliance.

[58] Thus, for example, under the rule of expectation damages, each party provides accident insurance to his counterpart in his capacity as promisor, and is insured by his counterpart in his capacity as promisee.

prospects of regaining his entitlement to it to dramatically rise. The buyer, in that case, has most likely not accumulated substantial reliance up to that point in time, and hence the rule of excuse does not lose its appeal.

This prediction is consistent with the important positive claim advanced by Andrew Kull.[59] Kull engaged in a historical study, examining the actual practices of courts in cases of contractual accidents. He argues that, as a descriptive matter, courts tend to deny accident claims when the traded good has already been delivered, but tend to accept them if the promisor has not already given up possession of it. In his words, "[t]he party who has balked at performing will not be forced to proceed, but the completed exchange will not be recalled."[60] Kull interprets these observations as reflecting the courts' policy of allowing windfall gains and losses to lie where they fall: If the cow has been delivered, let it stay with the buyer; and if not, let it remain with the seller.

The modified model developed here illuminates these observed practices in somewhat different light. The timing in which the accident is discovered often affects the promisee's level of reliance. The longer a period the promisee is allowed to accumulate his reliance investments, the less appealing the rule of excuse becomes. Hence, the courts' practice of awarding an excuse only in the preliminary stages of the contractual relationship is consistent with the theory's prescription.

C. Imperfect Verifiability of Accidents

The next assumption that is relaxed is the one stating that courts can always distinguish accident states from realizations of contemplated risks. I now assume that this factual determination can be problematic and examine the effect of this assumption on the results of the analysis.

When the contract fails to mention a contingency that is eventually realized, the court must interpret the meaning of that omission. One possibility is that due to positive transaction costs, the parties have simply failed to include a term that addresses the desired consequences in that state, although its corresponding risk had been taken into account in the determination of the price. The court's task of interpretation is then tantamount to deciphering the parties' intent ex ante. The court's ultimate goal is to enter into the parties' minds, while presuming that the thoughts going through those minds are informative.

---

[59] Andrew Kull, supra note 2.
[60] Id., at 5.

A second option is that the contingency is not mentioned because it had not been contemplated, and its realization therefore constitutes an accident. Attempting to delve into the parties' implicit intentions in these circumstances is not a fruitful exercise, for the problem is not one of uncovering the concealed aspects of the bargain, but rather that intent has never been formed. The court must then apply the rules pertaining to non-contemplated contingencies, which might be materially different from those which would apply under the hypothetical agreement scenario. But can a court realistically distinguish between these two possibilities?

A significant part of existing contract law doctrine relies on the notion that they can. Perhaps the most significant role of an adjudicator in the area of contracts is to interpret the contract by discerning that early intent from the language used, from the surrounding circumstances, and by applying some form of experience and common sense. On the other hand, it is clear that in at least certain circumstances the parties' state of mind at the time of contracting is exceedingly difficult to detect ex post. In fulfilling its interpretative task, the court is therefore bound to err at least occasionally.

Given the potential for judicial error, it is evident that the rule of excuse should not always be applied. It is clear, for instance, that when the realized contingency results only in a minuscule deviation from the parties' anticipated payoffs, discharging all contractual obligations would be unwise, given the positive probability that the contingency was in fact contemplated and contracted for. The desirability of the rule of excuse thus remains in this case as well a positive function of the accident's severity: the more substantial is the frustration of the parties' hoped for gains, the more appropriate its application becomes. Its appeal diminishes, however, the greater is the court's uncertainty as to whether the realized risk indeed constitutes an accident.

D. Renegotiation Costs

In certain settings, the accident does not entirely eliminate the potential for gains from trade. In such circumstances, the rule of excuse suffers from an evident drawback in that it compels the parties to either incur the costs of renegotiations, or forego the benefits of contracting altogether. The rule of expectation damages, on the other hand, allows trade to be consummated when a surplus can be made, without need to renegotiate. It follows that when there is reason to believe that renegotiation costs are substantial, and that the continued contractual relationship would produce value despite the accident, expectation damages become more appealing than otherwise.

This drawback of the rule of excuse must be qualified, however, in two significant respects. First, it should be observed that even under the rule of expectation damages some level of renegotiation is likely to occur simply because the newly discovered information generates new bilateral concerns that are best addressed by the formation of a new agreement. Thus, for example, suppose that a carrier undertakes to ship cargo by sea from the United States to Iran. While en route, it transpires that the Suez Canal has been closed, and he must decide on a course of action. Various options present themselves: It may proceed to deliver the goods via a different route; it may drop it off at a nearby port from which it could be shipped by air or land; if time is not pressing, it might simply keep it on board and deliver it only when the canal is reopened; or it can simply do nothing and allow the associated losses to be borne. As the contingency has not been explicitly negotiated prior to the formation of the original contract, the parties often cannot escape new negotiations if they are to determine which of those options is optimal. Furthermore, if the optimal choice is unverifiable to a court, then in the absence of renegotiations, the promisor would often be motivated to select a non-optimal course of action even if he knows what the optimal choice is.[61] The advantage of expectation damages over a discharge is thus not a general one. Rather, it is limited to cases in which renegotiations that occur under the latter rule are more costly than under then former.

The second qualification is that the rule of excuse not only raises transaction costs but also has the capacity to reduce them. As discussed in Section IV.1, an excuse renders the mapping of intentions to outcomes independent of accident risk. By enforcing the contract only when contemplated contingencies eventuate, the parties can conduct negotiations while disregarding the possibility of accidents. In contrast, under the rule of expectation damages, the parties' contractual payoff is a function of the probability and impact of accidents. In the course of bargaining, they must therefore agree on a manner in which accident risk would be incorporated into the terms of trade. That, in turn, creates an impediment to bargaining, which raises its costs. The rule of excuse, which avoids this complication, therefore lowers the costs of contracting.

E. The "Finality" of Transactions as a (Non)Reason for Full Enforcement

A final brief note is due to an argument that is often invoked

[61] See discussion of Goldberg's argument, *supra* note 21 and accompanying text.

against the award of rescission even in the face of accidents. It is the argument that a discharge undermines the principle of "finality" of the contractual agreement. E. Allan Farnsworth, for instance, expresses this principle as follows: "One who seeks to be excused on [the ground of mistake] must contend, at the outset, with the general rule that duties imposed by contract are absolute. The idea that finality is desirable in consensual transactions, lest justifiable expectations be disappointed, is expressed in the maxim, *pacta sunt servanda* ("agreements are to be observed"), rendered by the Seventh Circuit as "a deal's a deal."[62] To illustrate this principle, he continues to offer the following example: If an acquired land is found to contain mineral deposits that make it far more valuable than either party had supposed, then the original owner should not be entitled to avoid the contract. That is despite the fact that the mistake is assumed to be mutual and all other conditions prescribed by the doctrine are satisfied.[63] The reasoning, as stated, is that the sale of land must be made "final."

But why is finality an important principle to observe? The concept of finality admits several interpretations, and each of these merits closer examination. One possible argument is that the principle of finality protects the reliance of the promisee. The promisee wishes to be assured at the time she relies that her sunk investment would yield its anticipated return rather than go to waste. If transactions are "final," in the sense that the return for reliance can be recovered always, either through performance or through entitlement to damages, the promisee's investment is protected.

The premise underlying this claim, however, requires further elaboration. For to argue in favor of the promisee's right to reap the fruit of reliance in accident states, is to imply that accident losses ought to be entirely borne by the promisor. But as the accident by definition cannot be prevented by either party, one must provide a reason for such a-priori preference of the promisee. In some cases, such a reason can be deduced from the discussion above. Namely, rescinding the contract creates noise in the mapping of intentions onto outcomes and generates renegotiation costs. And if the court might mistakenly find an accident where the contingency had in fact been contemplated, the holdup problem might generate under-investment in reliance. However, if that is what one means by arguing in favor of the principle of finality, then finality is merely a manifestation of more fundamental interests, not an interest in its own right.

---

[62] See Farnsworth, supra note 2, §9.1.

[63] Namely, it involves a "basic assumption" on which the contract was made, and it has a "material effect on the agreed exchange of performances." Id., at §9.3.

A second possible rationale for the principle of finality is that accidents occur in virtually all contractual relationships, and so if any accident would entitle a promisor to a remedy of rescission, essentially all contracts would be avoided rather than performed. Intuitively, that would clearly undermine the parties' common intention that contractual commitments would be deemed enforceable at least in some states of the world.

One might further ask, however, what accounts for that intuition that parties do not wish all accidents to be governed by the rule of excuse. The discussion above is again instructive. When accidents are relatively minor, discharging the contract creates a risk of judicial error and generates further costs of renegotiations. And in the presence of reliance, the mapping of intentions onto outcomes is further disrupted. For these reasons, the parties wish their agreement to be "final" as long as the costs of a discharge exceed their corresponding gain. Thus, under this interpretation as well, the considerations already discussed explain the virtues of finality. Moreover, they trace its more fundamental roots, and thereby clarify its appropriate limits.

Finally, it might perhaps be thought that an enforcement policy is somehow superior if its applied remedy remains invariant to the realized state of the world. However, surely, parties do not derive independent value from the uniformity of the legal standard as such; in fact, the value of contracting is often derived precisely from the freedom to attach different consequences to distinct contingencies. That is the core value arising from the courts' readiness to interpret the contract rather than mechanically apply its literal instructions. That reason alone, therefore, cannot justify a general presumption against the desirability of the rule of excuse.

The emerging conclusion is therefore that the principle of finality does not represent a compelling rationale for full contractual enforcement. It either rests on dubious reasoning, or is synonymous with other, more fundamental rationales, that have been separately discussed and evaluated. In either case, it does not provide an independent basis for denying an excuse in the presence of accidents.

F. Summary

The discussion above suggests several reasons for why it would be ill-advised to apply the rule of excuse in all accident settings. Namely, if the accident entails only a non-substantial departure from the parties' expected gains under contemplated states, the advantages of the rule of excuse may weaken, or be outweighed by the costs of holdup, renegotiations and judicial error. To be sure, the optimal balance of these countervailing interests may at times be difficult to

discern in a particular case. The significance of these results is therefore not in that they provide a clear thumb rule under which the determination could be easily reached. Rather, it is that it clarifies the relevant grounds on which the determination ought to be based.

VI. Conclusion

The law and the economics of contractual accidents have so far not been speaking about the same phenomenon. Economic modeling has been pursued under the standard model, which does not recognize the existence and significance of non-contemplated states of the world. Legal analysis, on the other hand, has perceived accidents as precisely those instances in which contingencies not originally contemplated happen to arise. This fundamental conceptual divide has been manifested, among other things, in the exceeding difficulty to explain the merits of legal doctrine through economic reasoning.

This paper has taken a new approach to the economic modeling of this problem. It has accepted the traditional legal view of what accidents are, and altered the standard model in a manner that allows this view to play out in a meaningful form. Although the standard model has been altered only to the necessary, minimal extent, it provides a set of insights about the nature of the desirable remedy of non-performance, the value of existing doctrine in general, and the criteria that should inform its application in any particular case.

Appendix

The model developed below formalizes arguments made in the critique section of the paper.

The standard justification for the rule of excuse rests on the notion that the rule induces optimal precautions and reliance as well as a superior allocation of risk when the promisor is the party more averse to risk. The model developed next, however, establishes formally that, even within the confines of the standard model, the rule does not, in fact, carry these desirable properties. In particular, it demonstrates that the induced incentives pertaining to precautions, reliance and trade are all generally inefficient. The result that the rule also fails to generate a superior allocation of risk is derived formally by Alan Sykes, and so, for the development of that argument, the reader is referred to his paper.[64]

General Setup of the Model

Consider a buyer and a seller who enter a contract in view of uncertainty regarding the ultimate cost of performance, but can take measures to mitigate the risk of a high cost or insure against it. Thus, for example, one may imagine a contract between a shipper and a carrier for the delivery of a valuable by sea, acknowledging that the shortest sailing route may be closed down due to some unfolding contingencies. The risk of closure may be mitigated if they acquire information before the ship sets sail and plan accordingly; or by taking a second-best, alternate route that is sure to be open. Similarly, one may imagine a contract for the supply of crop, where the cost of production may rise if weather conditions are unfavorable. While the parties cannot affect the weather itself, they may take steps to reduce their dependency on it, or to insure against the impact of those adverse conditions.

The model's basic notation is defined as follows:

q - The contract price;

x^S , x^B - Precautions undertaken by the seller and buyer respectively, to avoid the high-cost contingency, or insure against it;

---

[64] Alan O. Sykes, *"The Doctrine of Commercial Impracticability in a Second Best World,"* 19 J. Legal Stud. 43 (1990)

r^S, r^B - Relationship-specific reliance investments undertaken by the seller and buyer respectively, before the cost of performance is realized;

$p_L\left(x^B, x^S\right)$ - The probability of low cost of performance, assumed to be concave in both elements.[65]

$v\left(r^B\right)$ - The value of performance to the buyer, assumed to be concave;

$c_L\left(r^S\right), c_H\left(r^S\right)$ - The seller's cost of performance, given the low and high contingencies respectively. It is assumed that $c_H\left(r^S\right) > q$, so that the seller seeks to avoid performance if the cost is high. Both cost functions are assumed to be convex.

I follow Posner and Rosenfield in examining two liability rules, whose instruction is depicted in the Table below. The first is the familiar negligence standard, under which the seller is granted an excuse if and only if he takes at least due care in preventing the high-cost contingency. The alternative is a more restrictive standard, in that it conditions an excuse not only upon the optimal behavior of the seller, but also upon the sub-optimal behavior of the buyer. Hence, the negligence rule releases the seller from liability in the settings corresponding to boxes 1 and 3, whereas the alternative rule does the same only for that corresponding to box 3.

Notice that the latter rule is a hybrid of the familiar negligence and strict liability standards: it is more favorable to the seller than pure strict liability, as it does not hold the seller liable under all states of the world. However, it is less favorable than a pure negligence standard, as it may deny an excuse even when the seller had acted efficiently. I therefore refer to it as the "hybrid rule."[66]

---

[65] I assume for simplicity that the probability does not depend on the parties' respective reliance investments. Assuming otherwise would not affect the qualitative results.

[66] Note that the hybrid rule also differs from strict liability combined with a defense of contributory negligence, as it does not allow the seller to avoid the contact merely on the grounds that the buyer had taken inadequate care. Rather, an additional, necessary condition for release is that the seller had taken due care.
Posner and Rosenfield nevertheless refer to the hybrid rule as a standard of "strict liability." As it is clear from the context, however, that the rule is intended to release the promisor from liability in some occasions, the term "hybrid rule" more accurately describes the rule's instruction.

Table: Excuse under the Negligence Rule versus the Hybrid Rule

		Seller	
		Due Care	*Inadequate Care*
Buyer	*Due Care*	1 Negligence Rule: Excuse Hybrid Rule: Liability	2 Liability
	Inadequate Care	3 Excuse	4 Liability

Posner and Rosenfield consider both rules as efficient, but express some preference to the latter, based on the hypothesis that it is less prone to judicial error.[67] Upon reexamination of these two rules, however, it is revealed that they are not equivalent in the incentives they induce, and in fact neither of them generates optimal behavior with respect to any of the examined parameters.

I initially consider the rules' effect on precautions and trade, and then proceed to examine their effect on reliance.

Precaution and Trade

Social Optimum

Let us begin by considering the condition for efficient trade. From a social standpoint, it is of course efficient that the parties proceed with the exchange if and only if the value the buyer attaches to performance exceeds the seller's cost of providing it. While it is assumed that $v\left(r^B\right) \geq c_L\left(r^S\right)$ holds in general, it is not generally assumed that $v\left(r^B\right) \geq c_H\left(r^S\right)$. Hence, given a high realization, trade may or may not be desirable.

Assuming that the condition for optimal trade is satisfied, the social cost associated with a high realization is given by:

[67] Richard A. Posner and Andrew. M. Rosenfield, *Impossibility and Related Doctrines in Contract Law: An Economics Analysis*, 6 J. Leg. Stud. 83, at 110-111 (1977).

$$Min\left(\left(v\left(r^B\right)-c_L\left(r^S\right)\right),\left(c_H\left(r^S\right)-c_L\left(r^S\right)\right)\right).$$

Thus, for purposes of identifying the social benefit of avoiding the adverse realization, we may restrict attention, without loss of generality, to a setting in which:

$$c_H\left(r^S\right)\in\left[q,v\left(r^B\right)\right]. \tag{1}$$

We may now turn to define the social value of contracting. Assuming (1), the expected gains from trade are a weighted average of the surplus captured under the two potential states, less the cost of reliance and precautions:

$$\begin{aligned}&p_L\left(x^B,x^S\right)\left(v\left(r^B\right)-c_L\left(r^S\right)\right)+\left(1-p_L\left(x^B,x^S\right)\right)\left(v\left(r^B\right)-c_H\left(r^S\right)\right)\\&-x^B-x^S-r^B-r^S\end{aligned} \tag{2}$$

Condition **Error! Reference source not found.**, in turn, yields the following first-order-conditions:

$$\frac{\partial p_L}{\partial x^B}\left(x^B,x^S\right)\left(c_H\left(r^S\right)-c_L\left(r^S\right)\right)=1 \tag{3}$$

$$\frac{\partial p_L}{\partial x^S}\left(x^B,x^S\right)\left(c_H\left(r^S\right)-c_L\left(r^S\right)\right)=1 \tag{4}$$

$$v'\left(r^B\right)=1 \tag{5}$$

$$-\left(p_L\left(x^B,x^S\right)c_L'\left(r^S\right)+\left(1-p_L\left(x^B,x^S\right)\right)c_H'\left(r^S\right)\right)=1 \tag{6}$$

I assume that a solution to equations (3) through (6) exists and is unique. Let $\left(x^B*,x^S*,r^B*,r^S*\right)$ denote the variable values that comprise that solution.

Private Optimum

Throughout the analysis it is assumed that the applicable remedy in case of a non-excused breach is that of expectation damages. The seller, therefore, holds a general option to avoid performance if he pays the amount of $v\left(r^B\right)-q$. As can be easily verified, the seller exercises the option if and only if $c_H\left(r^S\right)>v\left(r^B\right)$, and so his production cost is effectively capped by $v\left(r^B\right)$. Hence, again without loss of generality, I restrict attention, to cases where $c_H\left(r^S\right)\in\left[q,v\left(r^B\right)\right]$.

The Negligence Rule

Let us now turn to examine the parties' respective incentive for trade and precautions. Under the negligence rule, taking "due care" is the seller's dominant strategy. To see this, suppose, by way of contradiction that it were optimal for him to take inadequate care. In that case, he would not be excused in the event of a high realization, and his expected payoff from contracting would therefore be given by:

$$p_L\left(x^B,x^S\right)\left(q-c_L\left(r^S\right)\right)-\left(1-p_L\left(x^S\right)\right)\left(c_H-q\right)-x^S-r^S. \qquad (7)$$

But observe that the first-order-condition of (7) is the same as (4), implying that the privately optimal level of care is x^S *. That, however, is contrary to the presumption of inadequate care.

Similarly, suppose, by way of contradiction, that the seller chooses an excessive level of care. As he would then be excused from performance in the event of a high cost, his private value from contracting would be given by:

$$p_L\left(x^B,x^S\right)\left(q-c_L\left(r^S\right)\right)-x^S-r^S \qquad (8)$$

and the corresponding first-order-condition would be:

$$\frac{\partial p_L}{\partial x^S}\left(x^B,x^S\right)\left(q-c_L\left(r^S\right)\right)=1. \qquad (9)$$

But as $q - c_L\left(r^S\right)$ represents only part of the social cost associated with a high realization, the level of care maximizing (4) cannot exceed that which maximizes (9). Hence the privately optimal level of care cannot exceed $x^S *$. The seller, therefore, takes due care regardless of the buyer's choice.

Since the buyer takes due care, he is excused under the negligence rule. Hence, observe that regardless of the buyer's behavior, the seller fails to perform when the cost is high, whether or not trade is socially desirable. Thus, in particular, inefficient breach occurs whenever $q < c_H\left(r^S\right) < v\left(r^B\right)$.

Let us now turn to consider the buyer's optimal precaution strategy, given that the seller takes due care, and chooses to avoid performance whenever the realized cost is high. The buyer's best response in this setting is to take inadequate care. Intuitively, the reason is that the buyer knows that in the event of a high cost the contract would be rescinded, and each party would be made to bear merely his own accident losses. As he would therefore not incur the full social cost, his effort to prevent it would be inadequate.

To see this formally, observe that the buyer's expected value from contracting is given by:

$$p_L\left(x^B, x^S *\right)\left(v\left(r^B\right) - q\right) - x^B - r^B. \tag{10}$$

and the associated first-order-condition is accordingly:

$$\frac{\partial p_L}{\partial x^B}\left(x^B, x^S *\right)\left(v\left(r^B\right) - q\right) = 1. \tag{11}$$

As the social cost of an accident is $v\left(r^B\right) - c_L\left(r^S\right) \geq v\left(r^B\right) - q$,[68] the buyer's incentive to take precautions falls below the socially optimal level.

Let us refer to the chosen level of precaution as $\overline{x}^B$.

[68] Note that the social cost is $v\left(r^B\right) - c_L\left(r^S\right)$ rather than $c_H\left(r^S\right) - c_L\left(r^S\right)$, since the seller is always excused in equilibrium, and hence the social cost of an accident is the entire contractual surplus.

The Hybrid Rule

While the hybrid rule generates somewhat different incentives, the analysis of its impact relies on similar reasoning. The discussion here will therefore be brief.

The hybrid rule, like the negligence rule, induces optimal precautions by the seller as a dominant strategy, although the proof slightly varies. To see this, initially suppose that the buyer takes due care, and examine the seller's best response to that strategy. In the described setting, an excuse is never granted under the hybrid rule, and hence the seller's expected payoff is given by (7). As shown above, the level of care maximizing (7) converges with the social optimum, and hence the seller takes due care.

Now assume, alternatively, that the buyer takes inadequate care. Once again, the seller's best response is to take due care. The proof here is by contradiction, and is identical to the one presented in the case of the negligence rule (see conditions (7) through (9) and accompanying text.) Hence, "due care" is once again a dominant strategy for the seller.

Next consider the buyer's best response to the seller's choice to take due care. If the buyer invests adequately in precautions, then excuse is never granted. The buyer is therefore sure to capture his contractual surplus, either by way of performance or by way of collecting damages for breach. His payoff is thus given by:

$$\left(v\left(r^B\right)-q\right)-x^B*-r^B. \tag{12}$$

If, alternatively, he takes inadequate care, and a high cost is realized, then the seller is excused, and thus the buyer bears his own accident losses. His expected payoff thus becomes:

$$p_L\left(x^B,x^S*\right)\left(v\left(r^B\right)-q\right)-x^B-r^B. \tag{13}$$

The buyer's choice between due care and inadequate care will therefore depend on the relative magnitudes of (12) and (13). Inadequate care would thus be the chosen strategy if and only if:

$$x^B*-x^B>\left(1-p_L\left(x^B,x^S*\right)\right)\left(v\left(r^B\right)-q\right). \tag{14}$$

Note that two contradicting forces affect the buyer's decision in this setting. On the one hand, the private benefit arising from his investment falls below its social value, because he internalizes only his own foregone surplus in the event of a high realization, taking no account of the seller's loss $\left(v\left(r^B\right)-q \le v\left(r^B\right)-c_L\left(r^S\right)\right)$. On the other hand, however, there is a sense in which the private value exceeds the social value: by taking optimal precautions, the buyer acquires full insurance against accidents, although the social risk of a high cost is not generally eliminated.

Given these conflicting forces, the buyer's choice would thus depend on case-specific parameters. Namely, if $v\left(r^B\right)-q$ is sufficiently less than $v\left(r^B\right)-c_L\left(r^S\right)$, or $p_L\left(x^B, x^S\right)$ is sufficiently close to 1, then he would take inadequate care, and the precaution level would be given by $\bar{x}^B$, as under the negligence rule (observe that (13) and (10) are the very same condition.) However, if the opposite conditions prevail, then the chosen level of care would be optimal.

Note that excessive care by the buyer cannot be part of an equilibrium: as mentioned, by taking "due care" he already acquires full insurance against accident risk. Thus, investment beyond x^B* carries no private return.

Finally observe that if the parameter values are such that the buyer takes inadequate care, then the seller is excused whenever the cost is high. Thus, similarly to the case of a negligence rule, the seller breaches inefficiently whenever $q < c_H\left(r^S\right) < v\left(r^B\right)$.

Reliance

Consider next the parties' respective reliance incentives. As the equilibrium induced by the negligence rule is the same as one of the possible equilibria associated with the hybrid rule, I initially discuss these two cases conjointly. In what follows, I examine the remaining equilibrium associated with the hybrid rule.

If the seller is induced to breach inefficiently when a high cost is realized, then the contractual surplus could be captured only if the parties renegotiate a new contract. Initially suppose that renegotiation is

prohibitively costly. In that case, the socially optimal level of reliance, for the seller and buyer respectively, is given by:

$$-p_L\left(\bar{x}^B, x^S*\right)c_L'\left(r^S\right)=1 \tag{15}$$

$$p_L\left(\bar{x}^B, x^S*\right)v'\left(r^B\right)=1. \tag{16}$$

In view of conditions (8) and (10), which state the parties' respective gains from trade, it is evident that reliance, in this case, is efficient.

The same is not true, however, if renegotiations are possible. Denoting by $\theta\in[0,1]$ the seller's relative share of the surplus in renegotiation, the parties' expected gains from trade ex ante are given by:

$$p_L\left(\bar{x}^B, x^S*\right)\left(q-c_L\left(r^S\right)\right)+\theta\left(1-p_L\left(\bar{x}^B, x^S*\right)\right)\left(v\left(r^B\right)-c_H\left(r^S\right)\right) \tag{17}$$
$$-x^S*-r^S$$

$$p_L\left(\bar{x}^B, x^S*\right)\left(v\left(r^B\right)-q\right)+(1-\theta)\left(1-p_L\left(\bar{x}^B, x^S*\right)\right)\left(v\left(r^B\right)-c_H\left(r^S\right)\right) \tag{18}$$
$$-x^B-r^B$$

Accordingly, the first-order-conditions are:

$$-p_L\left(\bar{x}^B, x^S*\right)c_L'\left(r^S\right)-\theta\left(1-p_L\left(\bar{x}^B, x^S*\right)\right)c_H'\left(r^S\right)=1;\text{ and} \tag{19}$$

$$p_L\left(\bar{x}^B, x^S*\right)v'\left(r^B\right)-(1-\theta)\left(1-p_L\left(\bar{x}^B, x^S*\right)\right)v'\left(r^B\right)=1. \tag{20}$$

Comparing these conditions to the socially optimal ones (see (15) and (16)), it is easy to verify that both parties under-invest in reliance. This result is a manifestation of the classic holdup problem.

Finally suppose that, under the hybrid rule, precautions and trade are efficient. In that case, it is a familiar result that the buyer will invest excessively, if $c_H\left(r^S\right)>v\left(r^B\right)$.[69] To see this formally, note that under

[69] See Steven Shavell, *Damage Measures for Breach of Contract*, 11 Bell J. Econ. 466 (1980).

the specified conditions, the social expected gains from trade are given by:

$$p_L\left(x^B*, x^S*\right)\left(v\left(r^B\right)-c_L\left(r^S\right)\right)-x^B*-x^S*-r^B-r^S\,, \tag{21}$$

and hence the socially optimal investment must satisfy:

$$p_L\left(x^B*, x^S*\right)v'\left(r^B\right)=1\,. \tag{22}$$

However, as the buyer's private gains are given by (12), he invests up to the point where:

$$v'\left(r^B\right)=1\,. \tag{23}$$

which establishes his tendency to over-invest. Intuitively, the buyer relies too much because his private return for reliance exceeds its social value: as his choice to take optimal precautions provides him with full immunity against accident risk, he relies as if the investment would yield its hoped-for return under all states of the world.

***

In summary, the attractive properties attributed to the rule of excuse in existing literature have been found to be severely over-stated. In general, neither versions of the rule supports optimal precautions, trade or reliance. And as demonstrated by Sykes, it cannot be justified even on the basis of risk allocation, not even under the restrictive assumption that the promisor is the more risk-averse party. As explained in the body of the paper, the remaining virtues attributed to the rule are equally applicable to cases of ordinary breach, and hence cannot stand as an exclusive reason for the rule's application only in situations of accidents.

Two alternative conclusions can be drawn from these findings: if one believes in the descriptive validity of the model developed above, then one is inclined to conclude that the rule is in fact inefficient, despite its central role in American contract law, as in many other legal systems around the world. If, however, one's point of departure is that the rule is intended to apply only when the parties lack knowledge of probabilities and states, and hence the model does not adequately describe its

operation, then it follows that the model's results have little bearing on the rule's actual welfare effects.