

QUANTIFYING FORESEEABILITY

OREN BAR-GILL*

ABSTRACT

This Article extends the law-and-economics literature on the foreseeability doctrine and on contract default rules more generally. It derives (numerically) the optimal default cap on contractual damages in a model with a continuum of buyer types and perfect competition among sellers. When communication costs are low, the optimal cap is significantly higher than the damages incurred by the average buyer. A better performance technology reduces the optimal damages cap. Greater homogeneity among buyers increases the optimal cap. The Article identifies conditions under which an optimally defined foreseeability threshold significantly increases welfare. It also explores the normative implications of the doctrinal preclusion of a zero-damages default.

I. INTRODUCTION.....	620
II. MODEL.....	624
A. <i>Framework of Analysis</i>	624
B. <i>The Buyer's Decision Whether to Disclose Her Type</i>	625
1. <i>The Four-Region Equilibrium</i>	625
2. <i>The Three-Region Equilibrium</i>	627
C. <i>The Welfare Maximizing Rule</i>	627
D. <i>Numeric Analysis</i>	627
III. THE OPTIMAL DAMAGES CAP.....	628
A. <i>Communication Costs</i>	629
B. <i>Performance Technology</i>	631
C. <i>Distribution of Buyer Valuations</i>	635
IV. HOW IMPORTANT IS IT TO OPTIMALLY SET THE FORESEEABILITY THRESHOLD?.....	639
V. A ZERO-DAMAGES DEFAULT.....	642
VI. EXTENSIONS.....	647
VII. CONCLUSION.....	648

* Professor, New York University School of Law. This Article greatly benefited from comments and criticisms by Barry Adler, Lucian Bebchuk, Omri Ben-Shahar, George Geis, Clay Gillette, Marcel Kahan, Ehud Kamar, Jody Kraus, seminar participants at Tel-Aviv, and participants at the Symposium on Default Rules in Private and Public Law, Florida State University College of Law, March 25, 2005, and at the 2005 Israeli Law and Economics Association Annual Meeting. Jessica Gonzales, Robin Moore, Tien Anh Nguyen and Preeti Parikh provided excellent research assistance. Finally, I thank the John M. Olin Center for Law, Economics, and Business at Harvard Law School, the William F. Milton Fund of Harvard University, and the NYU School of Law for generous financial support.

I. INTRODUCTION

Contract law limits recovery for breach of contract. In particular, harm deemed extraordinary or excessively remote is excluded from the damages measure under the foreseeability doctrine.¹ Limited recovery is not, however, an immutable tenet of contract law. It is a default rule. A contracting party can render an extraordinary harm recoverable simply by informing the other party that such harm is possible in the present case. The foreseeability doctrine in effect specifies a threshold level of harm. Harm below the threshold, ordinary harm, is always recoverable. Harm above the threshold is recoverable only if it is communicated to the other party at the time of contracting. What is the optimal threshold? What harm should be recoverable by default, and when should specific communication be required?

The foreseeability doctrine has served in the law-and-economics literature as the primary example for studying the optimal design of default rules.² I too hope that the analysis developed in this Article will be modestly informative for general default rule theory. I do not, however, develop these broader implications here. Rather, I focus solely on foreseeability doctrine. My goal is to derive the optimal default cap on contract damages and to explore the relationship between this theoretical optimum and the judicial elucidation of the foreseeability doctrine.

The analysis is motivated by the recognition that damages for breach of contract lie on a continuum—from the obvious harm that would always follow a breach of the type of contract considered to the most remote and extraordinary harm that would occur in only a very small percentage of cases. It is this continuum that foreseeability doctrine divides—into recoverable and nonrecoverable damages—by setting a cap, a threshold level of damages. The question I ask is thus somewhat different from the standard question in this literature. Most previous accounts assume two discrete levels of damages—low damages and high damages—and ask whether the default rule should allow recovery for only low damages (limited liability), or for both low and high damages (unlimited liability). Rather than asking

1. See U.C.C. § 2-715 (2005); RESTATEMENT (SECOND) OF CONTRACTS § 351 (1981).

2. See, e.g., Barry E. Adler, *The Questionable Ascent of Hadley v. Baxendale*, 51 STAN. L. REV. 1547 (1999); Ian Ayres & Robert Gertner, *Filling Gaps in Incomplete Contracts: An Economic Theory of Default Rules*, 99 YALE L.J. 87 (1989) [hereinafter Ayres & Gertner, *Filling Gaps*]; Ian Ayres & Robert Gertner, *Strategic Contractual Inefficiency and the Optimal Choice of Legal Rules*, 101 YALE L.J. 729 (1992) [hereinafter Ayres & Gertner, *Strategic Contractual Inefficiency*]; Lucian Ayre Bebchuk & Steven Shavell, *Information and the Scope of Liability for Breach of Contract: The Rule of Hadley v. Baxendale*, 7 J.L. ECON. & ORG. 284 (1991); Jason Scott Johnston, *Strategic Bargaining and the Economic Theory of Contract Default Rules*, 100 YALE L.J. 615 (1990); George S. Geis, *Empirically Assessing Hadley v. Baxendale*, 32 FLA. ST. U. L. REV. 897 (2005).

whether liability for breach of contract should be limited or unlimited, I ask where on the continuum of damages levels the limit should be placed.³

The first step in the analysis is to derive the equilibrium behavior of the contracting parties. There are two alternative equilibria identified by the number of regions into which the group of buyers is divided. In the first, four-region equilibrium, buyers divide into four groups: (1) buyers with low valuations (and correspondingly low damages) who choose not to contract because they do not want to pay the high pooling price and they do not find it worthwhile to expend communication costs in order to attain a lower price, (2) buyers with higher valuations who separate themselves in order to pay a lower price, (3) buyers with intermediate valuations who save on communication costs and pay the pooling price, and (4) buyers with high valuations who are willing to pay a higher price to guarantee their expectation interest. In the second, three-region equilibrium, buyers divide into only three groups: (1) buyers with low valuations who choose not to contract, (2) buyers with intermediate valuations who save on communication costs and pay the pooling price, and (3) buyers with high valuations who are willing to pay a higher price to guarantee their expectation interest.

The size of the different groups in both equilibria depends on the damages cap. What is the optimal cap? In the standard two-type model, the optimal rule—that is, the optimal choice between limited and unlimited liability—induces the right type of equilibrium (pooling or separating) and, if separation is efficient, guarantees that this separation is attained with minimum communication costs.⁴ In the continuous-type model, characterizing the optimal rule is more complicated. First, as mentioned above, the choice is no longer binary—limited versus unlimited liability; rather the optimal cap can be anywhere on the continuum of possible damages levels. Second, the pooling versus separating question is no longer a binary question. There are different degrees of pooling and separating.

3. Bebchuk and Shavell briefly consider a continuous-type extension to their two-type model. See Bebchuk & Shavell, *supra* note 2, at 308. Johnston studies a continuous-type model but considers only two possible rules: an unlimited liability rule and a limited liability rule with liability capped at the average level of damages. See Johnston, *supra* note 2, at 626-27, 651. Moreover, Johnston assumes that sellers have market power, while I assume perfect competition among sellers. See *id.* at 636-37. Geis uses a continuous-type model but assumes discrete care levels (and, correspondingly, discrete probabilities of performance) and considers only an unlimited liability rule and a limited liability rule defined by a fixed damages cap. See Geis, *supra* note 2. Adler studies a two-type model but characterizes each buyer type by a stochastic continuous distribution of damages. See Adler, *supra* note 2, at 1561.

4. See, e.g., Bebchuk & Shavell, *supra* note 2.

Finally, the continuous-type model reveals an efficiency consideration that the two-type models, at least those assuming perfect competition, do not emphasize: a potentially significant group of buyers might be driven out of the market. It is well recognized that a seller with market power may choose not to serve the low end of the market.⁵ The continuous-type model shows that low-valuation buyers might leave the market even with perfect competition among sellers. And this inefficiency must also be considered in setting the optimal damages cap.⁶

The optimal damages cap achieves the right balance of pooling and separation, minimizes communication costs, and optimally controls the entry and exit of low-valuation buyers. I develop a numeric algorithm that solves for the optimal cap and study how different parameters, specifically the magnitude of communication costs, the performance technology, and the distribution of buyer valuations, affect the optimal cap.

The analysis yields the following results. When communication costs are low, as is arguably the case in most contracting environments, the optimal damages cap is significantly above the average buyer valuation.⁷ In fact, for a broad range of parameter values the optimal cap is above the eightieth percentile of buyer valuations. The optimal damages cap is lower when communication costs are larger, but even when communication costs are very large, the optimal foreseeability threshold significantly exceeds the average buyer valuation. This result supports the modern tendency to narrowly apply the foreseeability requirement.⁸

The performance technology, and specifically the cost-effectiveness of reducing the probability of breach, also has a significant effect on the optimal damages cap. Focusing on the more common, low communication costs environments, a superior performance technology requires a lower foreseeability threshold—a counterintuitive result. It is noteworthy that even this lower cap is significantly above the average buyer valuation. The remaining parameter is the

5. See, e.g., Ayres & Gertner, *Strategic Contractual Inefficiency*, *supra* note 2, at 744, 747.

6. Barry Adler notes that marginal buyers might be driven out of the market. Adler, *supra* note 2, at 1556 n.25. Adler, however, blames this result on the high pooling price that results under expansive liability. *Id.* at 1556-57. As I argue below, low-valuation buyers might leave the market because they do not find it worthwhile to secure a low price by communicating their type.

7. When communication costs are zero, full separation will obtain: every buyer will reveal her type and sellers will tailor the probability of breach and the price to every specific buyer. Clearly the legal default is of no importance when communication costs are zero. But communication costs are never zero. How small must communication costs be to render the search for the optimal default meaningless? I take up this question below.

8. See, e.g., 3 E. ALLAN FARNSWORTH, FARNSWORTH ON CONTRACTS § 12.14, at 259 (3d ed. 2004) (“The modern trend is . . . toward narrowing the limitation imposed by *Hadley v. Baxendale* by phrasing the test in terms of ‘foreseeability.’”).

distribution of buyer valuations. The numeric analysis suggests that a higher foreseeability threshold should be set when the distribution of buyer valuations is more homogeneous. The implications of asymmetric/skewed distributions of buyer valuations are also briefly explored.

How important is it to optimally set the foreseeability threshold? How much welfare will be lost if the wrong default is chosen? Not surprisingly, if communication costs are sufficiently small, the legal default matters very little. The analytical framework developed in this Article can tell us how small is *sufficiently* small. For example, if losing 1% of the maximum attainable welfare is considered to be insufficiently important in a given context, then as long as communication costs do not exceed 2% of the average buyer valuation, even significant deviations from the optimal default should not raise concern.

The optimal foreseeability threshold described above is in some cases only a second-best optimum. The optimization problem that produced this threshold was solved under the constraint that damages cannot be capped too close to zero. The numeric algorithm thus ruled out a zero- (or close to zero) damages default, even when such a default produced the highest welfare. This doctrinal-institutional constraint reflects the view that courts would resist a regime of no liability for breach of contract, even if this is only a default regime.⁹

What are the welfare implications of the doctrinal constraint? When communication costs are high, the constraint is not binding. In most real-world environments, however, where communication costs are low, the doctrinal constraint is binding. Moreover, the welfare loss incurred by ruling out a zero-damages default can be quite significant—up to 10% of the maximum attainable welfare for some parameter values. These adverse welfare implications merit a reconsideration of the doctrine behind the doctrinal constraint. While a broad “no damages” default seems institutionally untenable, zero-damages defaults for specific categories of damages, for example, lost retail profits or consequential damages more generally, does not seem unthinkable.¹⁰

9. Robert Scott and George Triantis, while developing a powerful case against the compensation principle in contract law, recognize that “[t]he dominance of the compensation principle is now unquestioned.” Robert E. Scott & George G. Triantis, *Embedded Options and the Case Against Compensation in Contract Law*, 104 COLUM. L. REV. 1428, 1446 (2004).

10. In fact, Ian Ayres has argued that lost retail profits should be governed by a zero-damages default. See Ian Ayres, *Valuing Modern Contract Scholarship*, 112 YALE L.J. 881, 891-92 (2003). More generally, since the foreseeability limitation applies only to consequential damages, see, e.g., 3 FARNSWORTH, *supra* note 8, § 12.14, at 256-57, strengthening this limitation by setting a zero-damages default would not lead to the doctrinally problematic result that breach of contract entails no remedy unless the parties specify a rem-

The analysis in this Article builds upon and extends previous work on the foreseeability doctrine. This line of research suffers from an important limitation. The theoretical results and specifically the optimal damages cap depend on parameter values that are difficult to ascertain empirically.¹¹ The practical value of this line of research depends on the ability to import—from other social sciences—and utilize more sophisticated empirical tools.¹² It is important to emphasize that the empirical assessments required to operationalize the theoretical results need not be absolutely accurate. The analysis in this Article suggests that even rough estimates of the underlying parameters can lead to a better calibration of the foreseeability threshold and thus to greater welfare.

From a policy perspective the parameter that is most important to evaluate empirically is the magnitude of communication costs. The model developed in this Article can tell us how large communication costs must be for the default rule to matter. But only empirical research can identify the contractual environments where communication costs are sufficiently large to justify an effort to optimally set the default.

The remainder of the Article is organized as follows. Part II develops the continuous-type model. Part III reports the results of the numeric analysis: how the optimal damages cap depends on the underlying parameters. Part IV identifies conditions under which an optimally defined foreseeability threshold can significantly increase welfare. Part V explores the normative implications of the doctrinal constraint that precludes a zero-damages default. Part VI discusses extensions. Part VII concludes.

II. MODEL

A. *Framework of Analysis*

A seller operating in a perfectly competitive market faces a buyer whose valuation of the offered product or service is unknown to the seller. Let v denote the value of the product or service—the value of contractual performance—to the buyer. This value, v , is the buyer's type. The seller knows only the distribution of buyer valuations, $f(v)$. The support of $f(v)$ is normalized to $[0,1]$ without loss of generality. For any specific buyer, the value of performance is private informa-

edy. The difficulty, of course, is that the distinction between regular and consequential damages is not always easy to draw.

11. See Adler, *supra* note 2; Eric A. Posner, *Economic Analysis of Contract Law After Three Decades: Success or Failure?*, 112 YALE L.J. 829 (2003).

12. Recent work by George Geis demonstrates how sophisticated empirical methodologies can be used to operationalize theoretical contract law models. See Geis, *supra* note 2.

tion, but the buyer can communicate her valuation to the seller at a cost, k .

The probability of successful performance is given by $p(x)$, where x represents the seller's precautions (in monetary terms). I assume that $p'(x) > 0$ and $p''(x) < 0$ (and, of course, $p(x) \in [0,1]$). If the seller breaches the contract, he will pay damages, d . The damage measure can take one of two values: (1) the actual full expectation measure, v , or (2) the legally determined foreseeable harm damage measure, \hat{v} .¹³ Let $d(v)$ denote the damage measure applicable to buyer v .

The seller's expected total cost when facing a subset of buyers $v \in [\underline{v}, \bar{v}]$ is $c_{\bar{v}}(x) = \int_{\underline{v}}^{\bar{v}} [(1-p(x)) \cdot d(v) + x] f(v) dv$.¹⁴ The seller will choose x to minimize his expected cost. The minimized cost, $c_{\bar{v}}(x^*) \equiv c_{\bar{v}}$, defines the price charged by the seller. The profit of buyer v is $\pi(v) = p(x) \cdot v + (1-p(x)) \cdot d(v) - c_{\bar{v}}$ (minus communication costs, k , when the buyer chooses to communicate her type to the seller).

B. The Buyer's Decision Whether to Disclose Her Type

There are two alternative equilibria in this model. In the first, four-region equilibrium buyers divide into four sub-groups: (1) buyers with low valuations, $v \in [0, v_1]$, who choose not to contract because they do not want to pay the high pooling price and they do not find it worthwhile to expend k in order to attain a lower price, (2) buyers with higher valuations, $v \in [v_1, v_2]$, who separate themselves in order to pay a lower price, (3) buyers with intermediate valuations, $v \in [v_2, v_3]$, who save on communication costs and pay the pooling price, and (4) buyers with high valuations, $v \in [v_3, 1]$, who are willing to pay a higher price to guarantee their expectation interest. In the second, three-region equilibrium, buyers divide into only three sub-groups: (1) buyers with low valuations, $v \in [0, v_1]$, who choose not to contract, (2) buyers with intermediate valuations, $v \in [v_2, v_3]$, who save on communication costs and pay the pooling price, and (3) buyers with high valuations, $v \in [v_3, 1]$, who are willing to pay a higher price to guarantee their expectation interest.

1. The Four-Region Equilibrium

The threshold value v_1 can be found by solving

$$(1) v_1 = c_{v_1} + k.$$

13. I abstract from the possibility of stipulated damages set at a level other than the full expectation measure.

14. In the special case where the seller knows the buyer's valuation, his expected cost is $(1-p(x)) \cdot d(v) + x$.

The marginal buyer is indifferent between exiting the market and earning a zero payoff on the one hand, and communicating her type and earning her valuation minus the tailored price (and minus communication costs) on the other hand. The threshold price is $c_{v_1} = (1 - p(x^*(v_1))) \cdot v_1 + x^*(v_1)$, where $x^*(v_1) = \text{argmin} \langle (1 - p(x(v_1))) \cdot v_1 + x(v_1) \rangle$.

The threshold value v_2 satisfies

$$v_2 - c_{v_2} - k = p(x) \cdot v_2 + (1 - p(x)) \cdot v_2 - c_{v_2v_3}, \text{ or}$$

$$(2) \quad c_{v_2} + k = c_{v_2v_3}.$$

The marginal buyer is indifferent between paying the tailored price plus communication costs on the one hand, and paying the higher pooling price on the other hand. Note that $v_2 < \hat{v}$. A buyer with $v_2 = \hat{v} - \epsilon$ would not expend k to get a marginally more favorable combination of price and care. The threshold price is $c_{v_2} = (1 - p(x^*(v_2))) \cdot v_2 + x^*(v_2)$, where $x^*(v_2) = \text{argmin} \langle (1 - p(x(v_2))) \cdot v_2 + x(v_2) \rangle$. A seller facing a pool of buyers with $v \in [v_2, v_3]$ expects a total cost of

$$c_{v_2v_3}(x) = \frac{1}{\text{Pr}(v \in [v_2, v_3])} \cdot \int_{v_2}^{v_3} [(1 - p(x)) \cdot d(v) + x] f(v) dv, \text{ where } d(v) = \begin{cases} v, & v \leq \hat{v} \\ \hat{v}, & v > \hat{v} \end{cases}, \text{ or}$$

$$c_{v_2v_3}(x) = \frac{1}{\text{Pr}(v \in [v_2, v_3])} \cdot \left\{ \int_{v_2}^{\hat{v}} [(1 - p(x)) \cdot v + x] f(v) dv + \int_{\hat{v}}^{v_3} [(1 - p(x)) \cdot \hat{v} + x] f(v) dv \right\}.$$

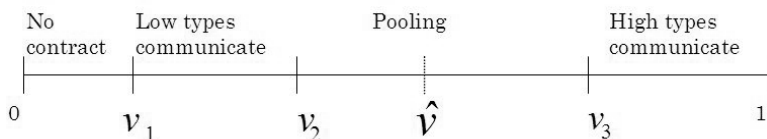
The pooling level of precaution is $x_{v_2v_3} = \text{argmin} \langle c_{v_2v_3}(x) \rangle$, and the pooling price is $c_{v_2v_3} \equiv c_{v_2v_3}(x_{v_2v_3})$. The threshold value v_3 satisfies

$$p(x) \cdot v_3 + (1 - p(x)) \cdot \hat{v} - c_{v_2v_3} = v_3 - c_{v_3} - k, \text{ or}$$

$$(3) \quad (1 - p(x)) \cdot (v_3 - \hat{v}) = (c_{v_3} - c_{v_2v_3}) + k.$$

The marginal buyer is indifferent between paying the pooling price and getting only the legally determined damages measure in case of breach on the one hand, and paying the higher tailored price plus communication costs to guarantee her expectation interest on the other. Note that $v_3 > \hat{v}$. A buyer with $v_3 = \hat{v} + \epsilon$ would not expend k to get a marginally more favorable combination of price and care. The threshold price is $c_{v_3} = (1 - p(x^*(v_3))) \cdot v_3 + x^*(v_3)$, where $x^*(v_3) = \text{argmin} \langle (1 - p(x(v_3))) \cdot v_3 + x(v_3) \rangle$. Using equations (1) through (3), I can solve for the threshold values, v_1 , v_2 and v_3 , as a function of the legally determined damages cap, \hat{v} . The four-region equilibrium is depicted in Figure 1.

FIGURE 1
THE FOUR-REGION EQUILIBRIUM



2. *The Three-Region Equilibrium*

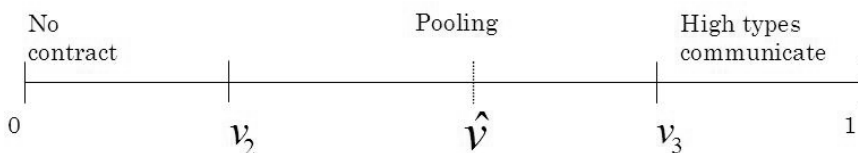
The threshold value v_2 satisfies $0 = p(x) \cdot v_2 + (1 - p(x)) \cdot v_2 - c_{v_2v_3}$, or

$$(4) \ v_2 = c_{v_2v_3} .$$

The marginal buyer is indifferent between exiting the market and earning a zero payoff on the one hand, and paying the pooling price on the other hand. The threshold value v_3 is derived from equation (3), as in the four-region equilibrium.

Using equations (3) and (4), I can solve for the threshold values, v_2 and v_3 , as a function of the legally determined damages cap, \hat{v} . The three-region equilibrium is depicted in Figure 2.

FIGURE 2
THE THREE-REGION EQUILIBRIUM



C. *The Welfare Maximizing Rule*

For any legally determined damages cap, \hat{v} , there is a unique vector of threshold values— (v_1, v_2, v_3) in the four-region equilibrium, and (v_2, v_3) in the three-region equilibrium—that defines the subgroups of buyers and implicitly the price, $c_{\bar{w}}$, paid by each buyer and the profit, $\pi(v)$, enjoyed by each buyer. Social welfare is given by $W(\hat{v}) = \int_0^1 \pi(v) f(v) dv$. The optimal \hat{v} solves $\max_{\hat{v}} \langle W(\hat{v}) \rangle$.

D. *Numeric Analysis*

The continuous-type model developed above is more realistic than the standard two-type model. The cost of realism, however, is com-

plexity. Solving the continuous-type model analytically is beyond this author's capabilities. I therefore resort to numeric analysis, solving the model (or, more precisely, asking a computer to solve the model) for different parameter values.¹⁵

First and foremost, I study the effects of communication costs by varying the magnitude of the parameter k . Second, I study the implications of different performance technologies. I use the probability of performance function $p(x) = 1 - e^{-\lambda x}$, where the λ parameter represents the performance technology, specifically the efficacy of precaution expenditures in reducing the probability of breach.

Finally, I study the role of the distribution of buyer valuations. The distribution of buyer valuations is assumed to follow a Beta distribution. The Beta distribution has a domain of $[0, 1]$.¹⁶ The distribution's two parameters, A and B , determine the variance and skewness of the distribution. The benchmark distribution in this study is a Beta distribution with parameter values $A = B = 1$, which is a Uniform distribution. The implications of greater homogeneity among buyers is explored by increasing the values of the distribution parameters while maintaining $A = B$. Removing the $A = B$ constraint allows me to explore the implications of skewed distributions.

The broad range of parameters studied represents many real-life contracting scenarios. Nevertheless, it should be recognized that the generality of the results is limited by the specific functional forms used in the numeric analysis. A different performance technology or a different distribution of buyer valuations will clearly affect the quantitative results. Still, many of the qualitative insights derived from the numeric analysis are robust to variations in the assumed functional forms. Moreover, the general framework developed in this Article can be used to reassess the quantitative results for alternative functional forms.

III. THE OPTIMAL DAMAGES CAP

The model developed in Part II can be used to study the characteristics of the optimal damages cap. How do different parameters—specifically the magnitude of communication costs—the performance technology, and the distribution of buyer valuations affect the optimal cap?

15. The numeric solutions were derived using Matlab (Version 7).

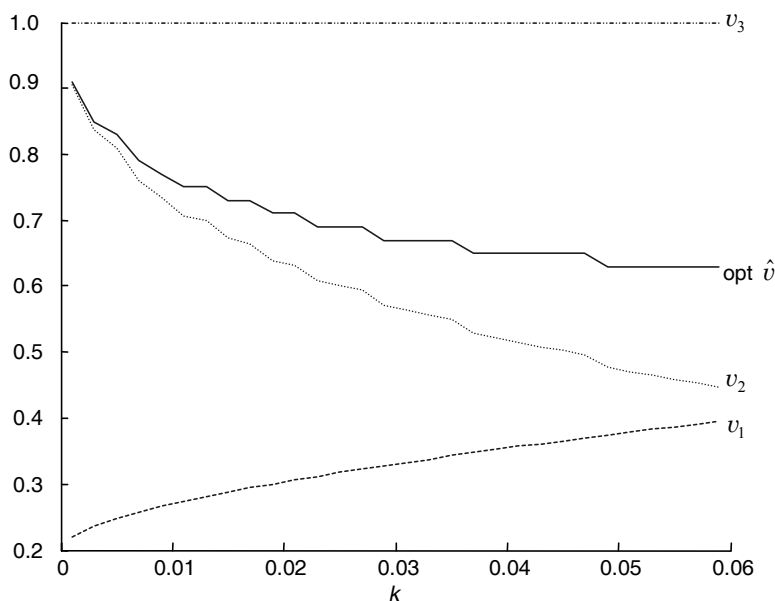
16. A more general version of the Beta distribution has parameters that determine the location and breadth of its domain.

A. Communication Costs

The magnitude of communication costs, as captured by the k parameter, significantly affects the optimal damages cap, \hat{v} . Figure 3 depicts the functional relationship between \hat{v} and k , when buyer valuations are uniformly distributed and the performance technology is characterized by $\lambda = 5$. Figure 3 also depicts the threshold values, v_1 , v_2 , and v_3 , as a function of k .

FIGURE 3

THE EFFECTS OF COMMUNICATION COSTS ($\lambda = 5, A = B = 1$)



Starting at the low communication costs end of the scale, the optimal foreseeability threshold is relatively high, at the ninetieth percentile of possible valuations of contractual performance. An increase in communication costs leads to a decline in the optimal damages cap. The reduced pooling price achieved by the lower cap allows more low-value buyers to join the pool and save on communication costs: v_2 decreases. The optimal cap falls monotonically, further increasing the size of the pool, and saving on communication costs.

It is not until communication costs reach a fairly high level—14% of the average buyer valuation—that the decreasing v_2 graph intersects with the increasing v_1 graph and the four-region equilibrium gives way to the three-region equilibrium (not depicted in Figure 3). At this point there is no more communication by low-value buyers:

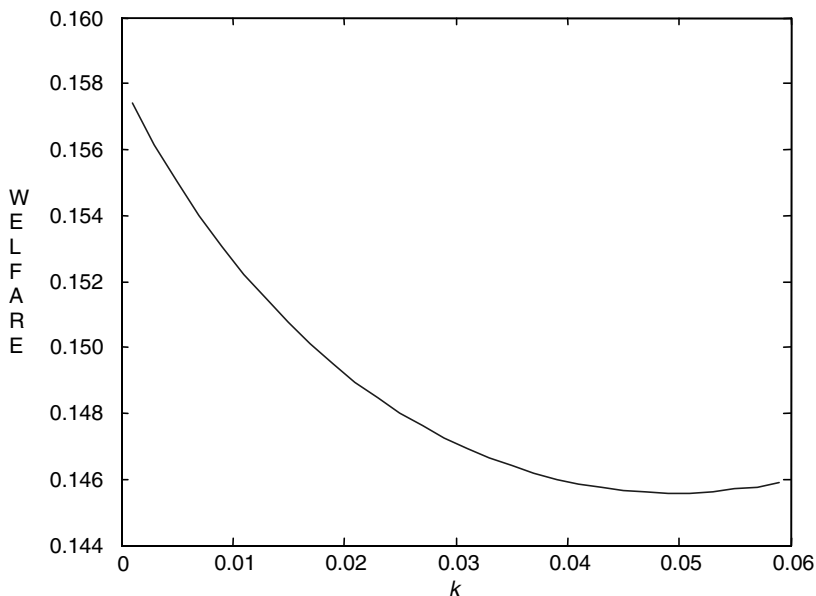
they either join the pool or exit (the communication option is always dominated by at least one of these two alternatives). Switching from the four-region equilibrium to the three-region equilibrium generates a discontinuity in the level of the optimal damages cap.

As communication costs continue to increase, the optimal cap resumes its monotonic decline. This decline lowers the pooling price and induces more low-value buyers to enter the market. At this stage communication costs are not borne by anyone. Still, higher communication costs prevent higher-value buyers from leaving the pool (even when the damages cap falls) and, thus, facilitate the inclusion of more low-value buyers through a reduced pooling price brought about by a lower damages cap.

Figure 3 highlights three results. First, the optimal damages cap is consistently and significantly above the average buyer valuation. Second, relatively small communication costs, 0.001, or 0.2% of the average buyer valuation (0.2% of 0.5 = 0.001), already generate substantial pooling. Specifically, 10% of buyers, those with the highest valuations, choose not to communicate their type. Third, even with relatively small communication costs, a substantial number of low-value buyers are excluded from the market. Specifically, over 20% of buyers, those with the lowest valuations, withdraw from the market.

Not surprisingly, attainable welfare decreases as communication costs increase. Figure 4 plots welfare as a function of communication costs.

FIGURE 4
WELFARE AS A FUNCTION OF COMMUNICATION COSTS
($\lambda = 5, A = B = 1$)

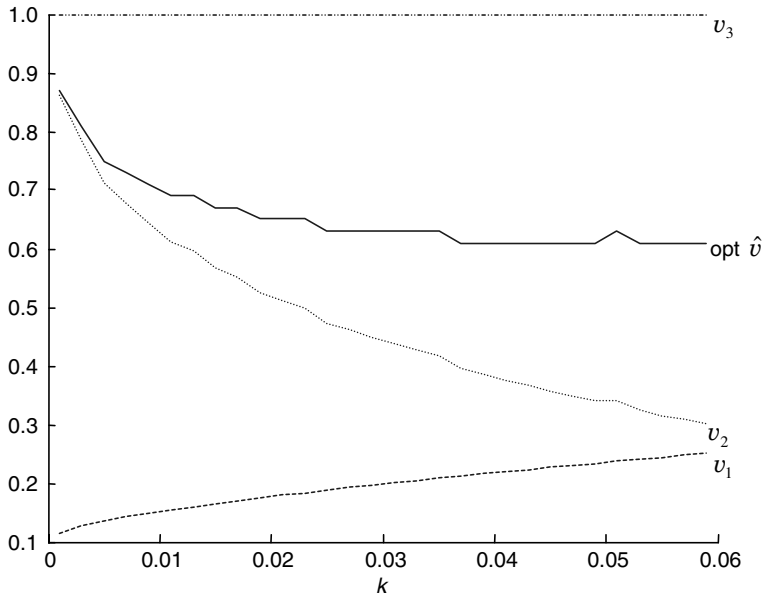


Starting from very low communication costs, an increase in communication costs from 0.001, or 0.2% of the average buyer valuation, to 0.01, or 2% of the average buyer valuation, leads to a 5.3% reduction in welfare. The marginal effect of communication costs on welfare decreases as communication costs increase. In fact, for sufficiently high communication costs, specifically $k = 0.05$, or 10% of the average buyer valuation, an increase in communication costs *increases* welfare. A possible explanation is that the high communication costs are not incurred by buyers, but rather serve to support an efficient pooling equilibrium that minimizes communication costs.

B. Performance Technology

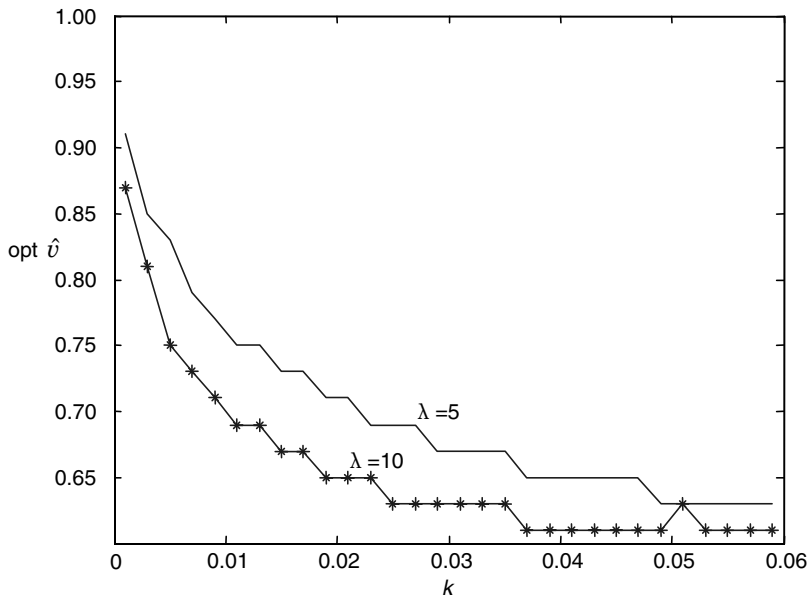
Are the results stated in Part III.A robust to changes in the performance technology? I begin by studying more efficient performance technologies. For concreteness, compare the performance technology studied in Part III.A, which is characterized by $\lambda = 5$, and a more efficient performance technology characterized by $\lambda = 10$. Figure 5 replicates Figure 3 with a $\lambda = 10$ technology.

FIGURE 5

THE EFFECTS OF COMMUNICATION COSTS ($\lambda = 10, A = B = 1$)

When communication costs are $k = 0.001$, the optimal foreseeability threshold given the $\lambda = 5$ technology is at the ninetieth percentile of possible buyer valuations. With the more efficient $\lambda = 10$ technology, the optimal threshold is below the ninetieth percentile. The more efficient technology makes pooling more attractive to high-valuation buyers. It is thus possible to reduce the foreseeability threshold, increase the size of the pool, and lower communication costs without driving high-valuation buyers out of the pool. The size of the pool increases further, compared to the size of the pool with the $\lambda = 5$ technology, as communication costs rise. The optimal foreseeability threshold with a $\lambda = 10$ technology is always lower than the optimal threshold with a $\lambda = 5$ technology. These results are depicted in Figure 6.

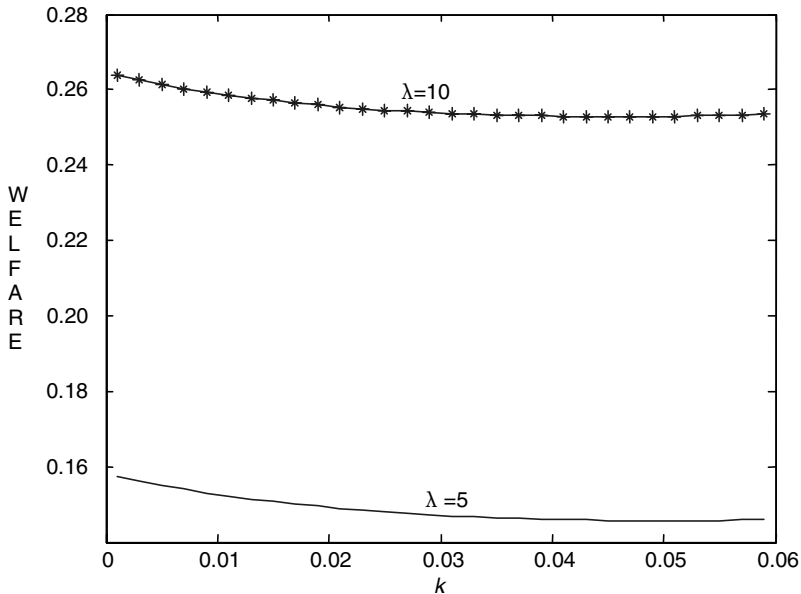
FIGURE 6

 \hat{v}^* AS A FUNCTION OF k FOR $\lambda = 5$ AND $\lambda = 10$ 

The more efficient performance technology also significantly reduces the number of low-valuation buyers who withdraw from the market. Specifically, for communication costs of $k = 0.001$ only 12% of all buyers withdraw from the market with the $\lambda = 10$ technology, as compared to over 20% of all buyers who withdraw from the market with the $\lambda = 5$ technology.

Not surprisingly, the more efficient performance technology produces higher welfare at every level of communication costs. Figure 7 graphs welfare as a function of communication costs for both the $\lambda = 5$ technology and the $\lambda = 10$ technology.

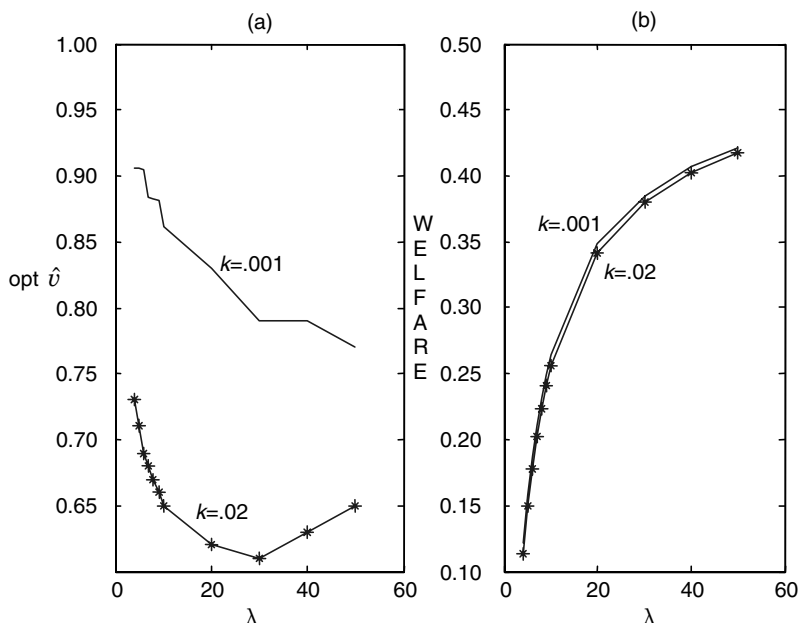
FIGURE 7
WELFARE AS A FUNCTION OF COMMUNICATION COSTS FOR
 $\lambda = 5$ AND $\lambda = 10$ ($A = B = 1$)



The welfare effects are quite large: moving from a $\lambda = 5$ performance technology to a $\lambda = 10$ performance technology entails a welfare increase of over 60%.

The comparison between the $\lambda = 5$ technology and the $\lambda = 10$ technology demonstrates the impact of the performance technology on both the optimal foreseeability threshold and the attainable level of welfare. Figure 8 presents a broader comparative statics analysis. Figure 8(a) plots the optimal threshold as a function of the performance technology ($\lambda \in [1, 50]$) for $k = 0.001$ and $k = 0.02$. Figure 8(b) plots the maximum attainable social welfare as a function of the performance technology ($\lambda \in [1, 50]$) for $k = 0.001$ and $k = 0.02$.

FIGURE 8
THE FORESEEABILITY THRESHOLD AND
WELFARE AS FUNCTIONS OF λ ($A = B = 1$)



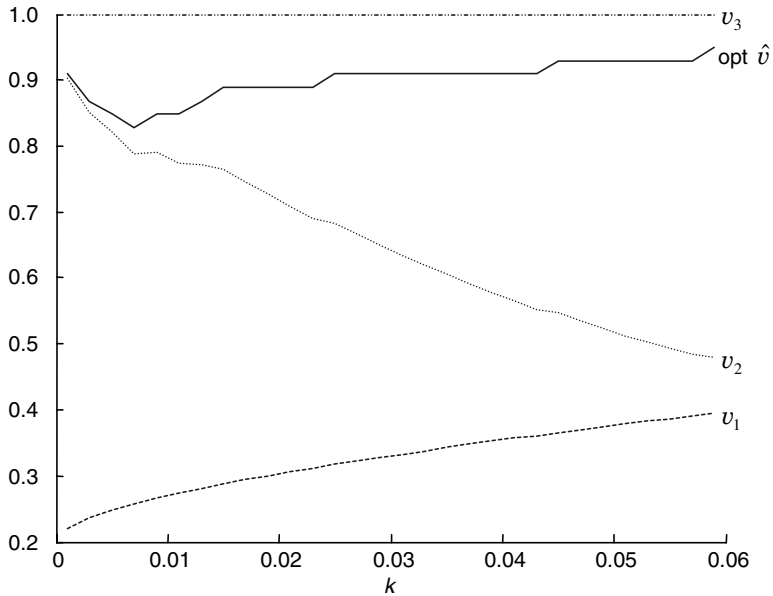
C. Distribution of Buyer Valuations

Another important factor affecting the optimal damages cap is the distribution of buyer valuations. Recall that with heterogeneous buyers, specifically with buyer valuations following a Uniform distribution, for very low communication costs the optimal threshold was approximately 0.90. And when communication costs increased, the optimal threshold declined monotonically.¹⁷ As buyers become more homogeneous, these results change. First, the initial decrease in the optimal threshold when communication costs increase is smaller. Second, after a certain level of communication costs is reached, this decrease is reversed, and the optimal foreseeability threshold increases with communication costs. The level of communication costs at which this change in trajectory occurs is decreasing in the degree of homogeneity among buyers.

These results are demonstrated for a distribution of buyers characterized by $A = B = 5$ in Figure 9.

17. Only at a very high level of communication costs, $k = 0.05$, and at a relatively low foreseeability threshold, approximately 0.63, did the trajectory change and the optimal threshold began to increase as communication costs rose further. The reason for this change in trajectory was the switch from the four- to the three-region equilibrium.

FIGURE 9

THE EFFECTS OF COMMUNICATION COSTS ($\lambda = 5, A = B = 5$)

Overall, the foreseeability thresholds are higher when the distribution of buyer valuations is more homogeneous. Moreover, the absolute thresholds underestimate the impact of increased homogeneity. With a uniform distribution of buyer valuations, a 0.9 foreseeability threshold implies that 90% of buyers would be fully compensated (for damages from breach of contract) under the default. With the more homogeneous $A = B = 5$ distribution, a 0.9 foreseeability threshold implies that over 99% of buyers would be fully compensated for breach of contract under the default.

What are the welfare implications of the degree of homogeneity? Figure 10 plots welfare as a function of communication costs for both $A = B = 1$ and $A = B = 5$.

FIGURE 10
 WELFARE AS A FUNCTION OF k
 FOR $A = B = 1$ AND $A = B = 5$ ($\lambda = 5$)

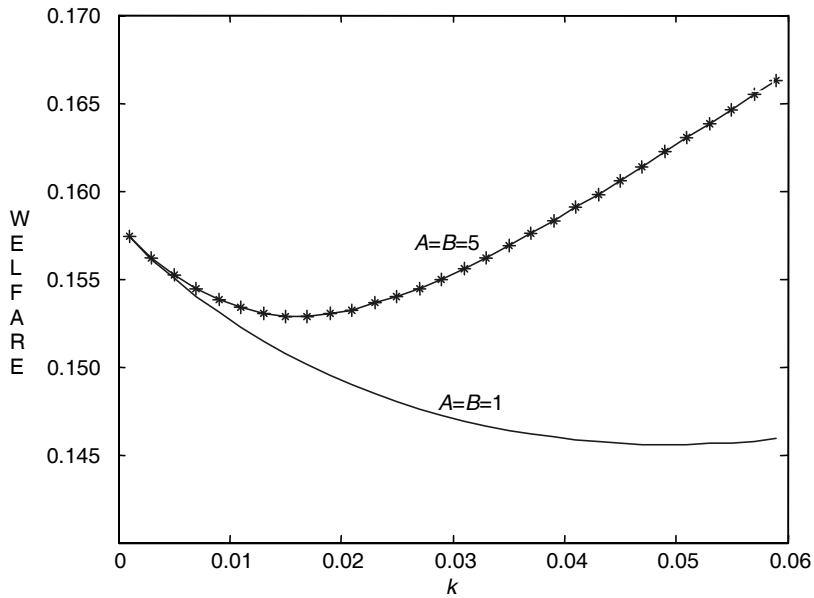
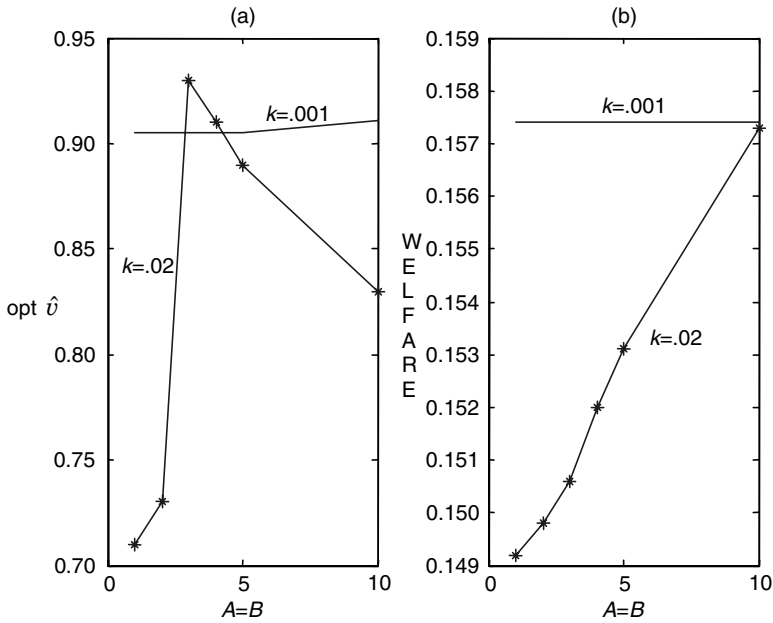


Figure 11 presents a broader comparative statistical analysis. Figure 11(a) plots the optimal threshold as a function of the degree of homogeneity ($A = B \in [1, 2, 3, 4, 5, 10]$) for $k = 0.001$ and $k = 0.02$. Figure 11(b) plots the maximum attainable social welfare as a function of the degree of homogeneity ($A = B \in [1, 2, 3, 4, 5, 10]$) for $k = 0.001$ and $k = 0.02$.

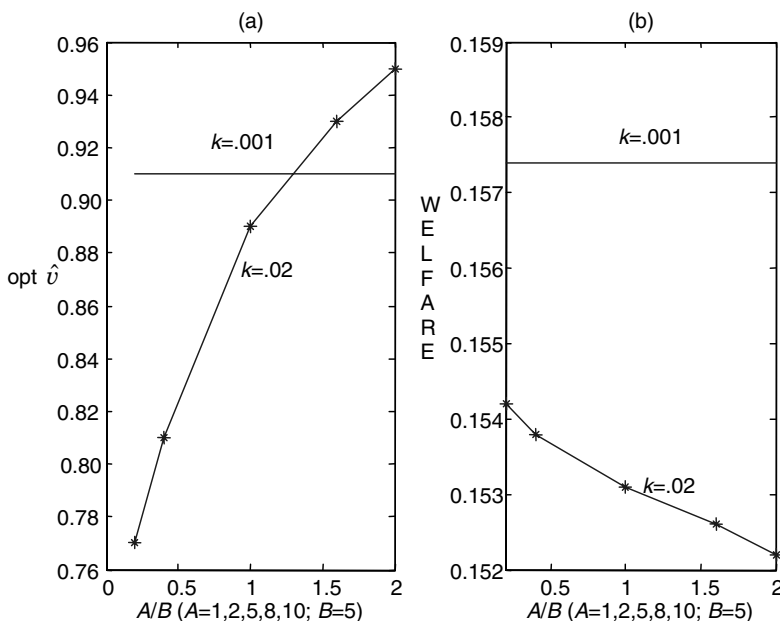
FIGURE 11

THE FORESEEABILITY THRESHOLD AND WELFARE AS FUNCTIONS OF $A = B$ ($\lambda = 5$)



The analysis thus far focused on symmetric distributions of buyer valuations. How would the results change with skewed distributions? Figure 12 begins to answer this question. It plots the optimal threshold (Figure 12(a)) and social welfare (Figure 12(b)) as a function of the ratio A/B , where a ratio $A/B < 1$ represents a distribution that is skewed to the left and a ratio $A/B > 1$ represents a distribution that is skewed to the right.

FIGURE 12
 THE FORESEEABILITY THRESHOLD
 AND WELFARE AS FUNCTIONS OF A/B ($\lambda = 5$)



IV. HOW IMPORTANT IS IT TO OPTIMALLY SET THE FORESEEABILITY THRESHOLD?

In many contractual environments communication costs are low. If it is easy to opt out of the default rule, then it should not matter very much whether the default is set optimally.¹⁸ Part IV attempts to quantify this observation. Exactly how much welfare is lost when the default rule is not set optimally? How small must communication costs be to render the choice of default immaterial?

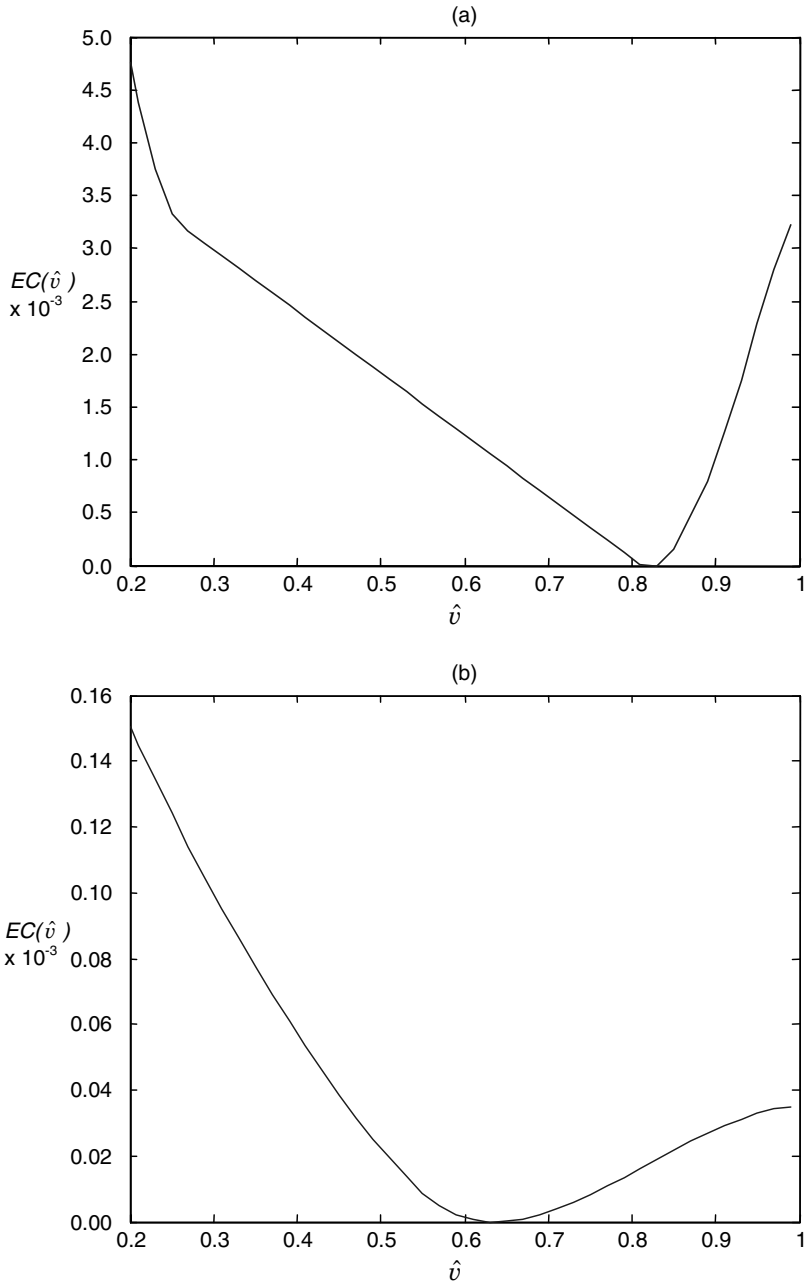
I begin by defining the efficiency cost of a default rule \hat{v} that deviates from the optimal default \hat{v}^* as $EC(\hat{v}; \hat{v}^*) = [W(\hat{v}^*) - W(\hat{v})] / W(\hat{v}^*)$. Figure 13(a) plots $EC(\hat{v}; \hat{v}^*)$ for $\hat{v} \in [0.2, 1]$ when $\lambda = 5$, $A = B = 1$, and $k = 0.005$. Figure 13(b) plots $EC(\hat{v}; \hat{v}^*)$ for $\hat{v} \in [0.2, 1]$ when $\lambda = 5$, $A = B = 1$, and $k = 0.05$.¹⁹

18. Recall the perfect competition assumption. In imperfectly competitive markets default rules can significantly affect welfare even when communication costs are very small. See, e.g., Ayres & Gertner, *Strategic Contractual Inefficiency*, *supra* note 2.

19. For an explanation why the domain in Figure 13 is $\hat{v} \in [0.2, 1]$, rather than $\hat{v} \in [0, 1]$, see *infra* Part V (and specifically footnote 21).

FIGURE 13

EFFICIENCY COST AS A FUNCTION OF \hat{v} FOR
 $k = 0.005$ AND $k = 0.05$ ($\lambda = 5, A = B = 1$)



To facilitate a broader comparative statistical analysis, I use a more specific definition of efficiency costs. I calculate $EC(\hat{v} = 0.5 \cdot \hat{v}^*)$ and $EC(\hat{v} = 1.5 \cdot \hat{v}^*)$, and define

$EC(\hat{v}^*) = \max\langle EC(\hat{v} = 0.5 \cdot \hat{v}^*), EC(\hat{v} = 1.5 \cdot \hat{v}^*) \rangle$ as the efficiency cost of failing to accurately set the default rule at the optimal level \hat{v}^* . This definition allows me to study how the different parameters affect the magnitude of the efficiency cost generated by inaccurate defaults.²⁰

I begin by studying the implications of communication costs. Figure 14 plots $EC(\hat{v}^*)$ as a function of k .

FIGURE 14

EFFICIENCY COST AS A FUNCTION OF k ($\lambda = 5, A = B = 1$)

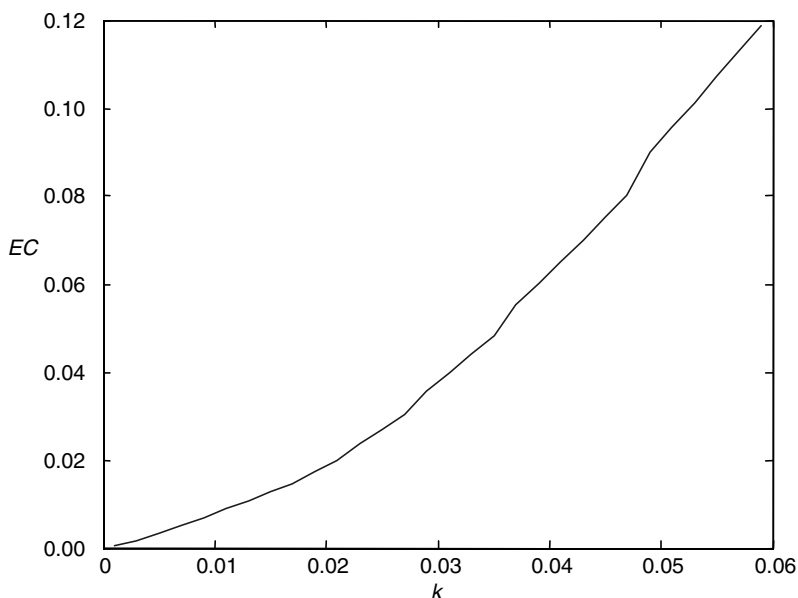


Figure 14 answers the question how small must communication costs be to render the choice of default immaterial. When communication costs are substantial, around 10% of the average buyer valuation, the efficiency cost of choosing the wrong foreseeability threshold is also substantial; almost 10% of the maximum attainable welfare is lost. But even for more reasonable levels of communication costs, say 2%

20. The assumption here is that the court gets the default wrong, but the parties anticipate this wrong default rule *ex ante*. The alternative assumption that *ex ante* the parties are uncertain about the default is discussed in Part VI *infra*.

of the average buyer valuation, the efficiency cost is nonnegligible, almost 1%.

The magnitude of the communication costs also determines the importance of tailoring the default to the performance technology and to the distribution of buyer valuations. When communication costs are high, such tailoring produces significant welfare gains. When communication costs are small, even large deviations from the optimally tailored default result in an insignificant welfare loss.

V. A ZERO-DAMAGES DEFAULT

The optimal foreseeability thresholds derived above were not always optimal. For some parameter values, a very different default rule—a zero-damages default—would produce greater welfare. In other words, the foreseeability thresholds studied thus far were in some instances only second-best optimal, namely, optimal given a feasibility constraint imposed on the optimization problem. The imposed feasibility constraint is doctrinal or even institutional in nature. It posits that courts cannot set a zero-foreseeability threshold.²¹ This would mean that absent a specific contractual provision to the contrary, there would be no liability for breach of contract.²²

To understand the potential importance of the doctrinal constraint, it is useful to study the relationship between social welfare and the foreseeability threshold for different levels of communication costs. Figure 15 depicts welfare as a function of the foreseeability threshold for $k = 0.1, 0.2, 0.3, 0.4, 0.5$ and 0.6 , assuming $\lambda = 5$ and $A = B = 1$.

21. Since the second-best optimal thresholds studied above are in some cases outperformed not only by the zero-damages default, but by a range of foreseeability thresholds between zero and 0.2, the imposed doctrinal constraint is more accurately defined as ruling out foreseeability thresholds below one-half of the average buyer valuation.

22. The posited doctrinal constraint is a corollary of the compensation principle in contract law. See Scott & Triantis, *supra* note 9, for a critical discussion of this principle. There may, however, be efficiency reasons to reject a zero-damages default, especially if default rules are sticky. See Omri Ben-Shahar & John Pottow, *On the Stickiness of Default Rules*, 33 FLA. ST. U. L. REV. 649 (2006); Johnston, *supra* note 2.

FIGURE 15

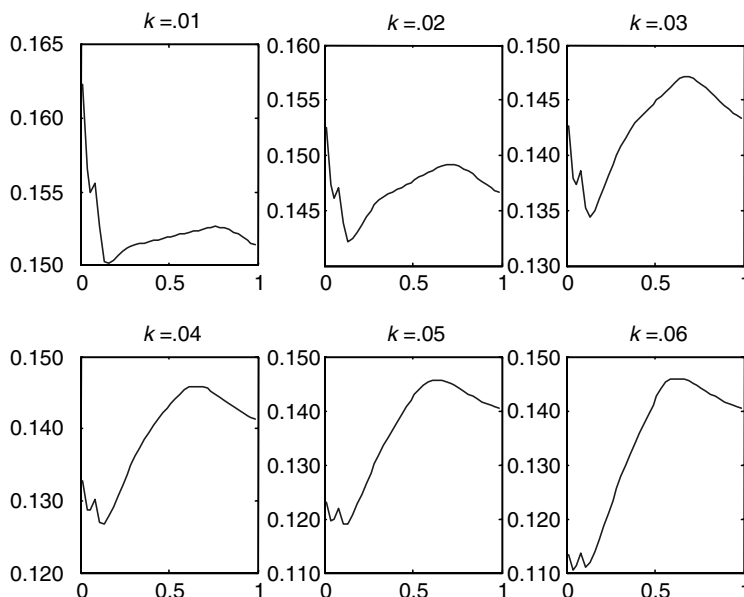
WELFARE AS A FUNCTION OF THE FORESEEABILITY THRESHOLD ($\lambda = 5, A = B = 1$)

Figure 15 shows that the welfare function has two local maxima: one at a foreseeability threshold above the average valuation (0.5) and another at a foreseeability threshold of zero. The intuition for this bimodal distribution is as follows. A three-region equilibrium is obtained when the foreseeability threshold is sufficiently low (see Figure 2, *supra*). Low-valuation buyers either withdraw from the market or join the pool. When the foreseeability threshold approaches zero, most of the low-valuation buyers remain in the market. Moreover, these low-valuation buyers do not incur communication costs. The reduction in communication costs incurred by low-valuation buyers is, however, counteracted by the increased communication costs incurred by high-valuation buyers. On balance, a foreseeability threshold near zero produces relatively high welfare, especially when communication costs are low such that the inclusion of low-valuation buyers is the dominant consideration.

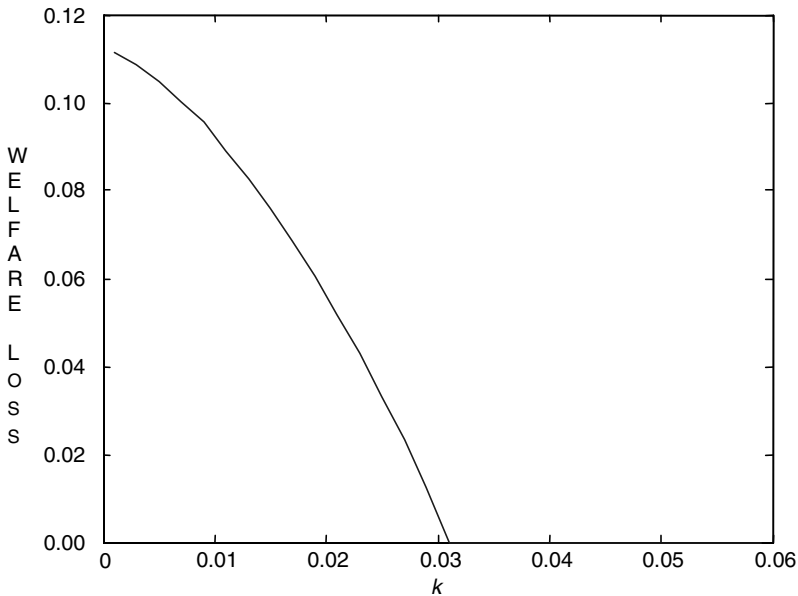
When the foreseeability threshold increases from zero, welfare initially falls as more low-value buyers exit the market. When the foreseeability threshold increases further, the marginal impact of the exclusion effect becomes weaker, since the remaining low-value buyers are not easily driven out of the market by a slight increase in the pooling price. Now welfare increases with the foreseeability threshold, as this increase economizes on communication costs. At some

point, the three-region equilibrium is replaced by the four-region equilibrium (see Figure 1, *supra*). Welfare continues to increase with the foreseeability threshold as long as the marginal communication costs saved by the high-valuation buyers outweigh the marginal communication costs incurred by the low-valuation buyers. When the toll on low-valuation buyers surpasses the benefit to high-valuation buyers, welfare starts to fall and continues to decline monotonically as the foreseeability threshold continues to rise.

Given the bimodal nature of the welfare function, when is the doctrinal constraint binding? When communication costs are large, the local maximum with a foreseeability threshold above the average buyer valuation is also the global maximum, and the doctrinal constraint is *not* binding. When communication costs are small, the local maximum with the zero-foreseeability threshold is the global maximum. Here the doctrinal constraint is binding.

How big are the welfare costs imposed by the doctrinal constraint? Figure 16 plots the welfare loss, as a percentage of the maximum welfare attainable, as a function of communication costs for $\lambda = 5$ and $A = B = 1$.

FIGURE 16
WELFARE LOSS AS A FUNCTION OF k ($\lambda = 5, A = B = 1$)



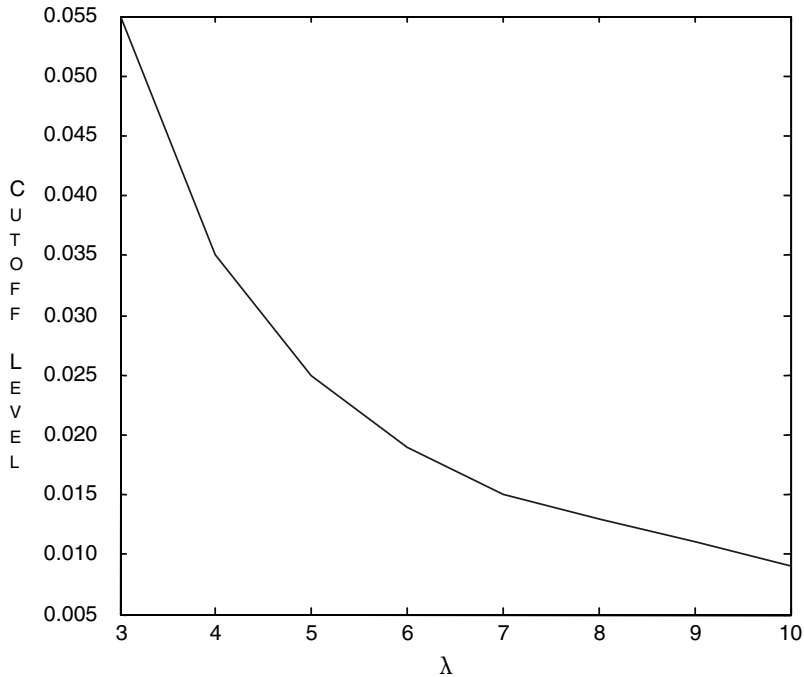
There is no welfare loss when communication costs exceed 0.031. When communication costs drop below 0.031, the doctrinal constraint

kicks in. The welfare loss increases monotonically as communication costs continue to fall. Figure 16 shows that the doctrinal constraint significantly affects welfare and can cause a welfare loss in excess of ten percent.

The welfare implications of the doctrinal constraint depend on the magnitude of the communication costs. Of particular interest is the cutoff level of communication costs beyond which the doctrinal constraint is binding. How is this cutoff level affected by changes in the performance technology and the distribution of buyer valuations?

Figure 17 depicts the cutoff level of communication costs as a function of the performance technology (represented by the λ parameter). The cutoff level is monotonically decreasing in the efficiency of the performance technology. The intuition for this result is as follows. The advantage of the zero-damages default lies in its ability to induce full separation. This, however, is also the main cost of the zero-damages default, since full separation entails high communication costs. When the performance technology is more efficient, the benefit of full separation—tailored precaution levels—is reduced. Accordingly, the zero-damages default remains socially desirable only for lower levels of communication costs.

FIGURE 17
CUTOFF LEVEL OF COMMUNICATION COSTS
AS A FUNCTION OF λ ($A = B = 1$)



The cutoff level of communication costs is less sensitive to changes in the degree of homogeneity among buyers. Still, the cutoff level decreases as the distribution of buyer valuations becomes more homogeneous.

The doctrinal constraint imposes significant welfare costs. Must we abide by this constraint? Are zero damages really unthinkable? Looking at the broad category of losses from breach of contract as a whole the answer is probably yes. But we need not look at this category as a whole. For example, Ian Ayres has argued that lost retail profits should be governed by a separate default rule: a zero-damages default rule.²³ The preceding analysis provides further support for this proposal. More generally, it may be desirable to break up the broad category of damages from breach of contract into several sub-

23. Ayres, *supra* note 10, at 891-92.

categories and to consider a zero-damages default for at least some of these categories.²⁴

VI. EXTENSIONS

The basic model studied in this Article can be extended in several directions. Three such extensions are listed below. I leave the development of these extensions for future research.

Stochastic Damages. The model developed and studied above assumes that each buyer is fully characterized by a deterministic valuation, v . Generally, however, a single buyer might suffer damages of varying magnitudes when the seller fails to perform the contract. The stochastic damages extension was first studied by Barry Adler in the standard two-type model. Stochastic damages reduce the incentive of high-value buyers to reveal their type. Consequently, efficient separation might not occur under a low damages cap and a more liberal foreseeability threshold may be warranted.²⁵ The continuous-type model developed in this Article can be extended to allow for stochastic damages. The effects of different distributions of damages levels could then be explored.

Uncertainty Regarding the Foreseeability Threshold. The preceding analysis assumed that contracting parties know with certainty what the law is, namely, at what level the court will set the foreseeability threshold \hat{v} . But even if courts try to optimally set the damages cap, they will inevitably err, producing uncertainty about the default rule. The proposed model can be extended to account for such uncertainty, at the contracting stage, about the applicable foreseeability threshold.²⁶

Foreseeability in Tort Law. Foreseeability of damages is important also in tort law. Like contract law, tort law imposes a limit on the ability of a victim to recover in damages.²⁷ Focusing on interactions

24. The “consequential damages” category—the object of the foreseeability limitation—may be one category where a zero-damages default could be applied. *See supra* note 10.

25. Adler, *supra* note 2.

26. Uncertainty about the legal rule is important also within the confines of the two-type model. In the two-type model, limited liability is assumed to allow for full recovery by the low-valuation type and only partial recovery by the high-valuation type. This assumption implies that $v_l \leq \hat{v} \leq v_h$, where v_l and v_h represent the low and high valuations, respectively. If the level of uncertainty, ε , is sufficiently small, such that $v_l \leq \hat{v} - \varepsilon \leq \hat{v} + \varepsilon \leq v_h$, then such uncertainty will not affect the parties’ equilibrium behavior. But if uncertainty is larger or if \hat{v} is closer to either v_l or v_h , then uncertainty about the legal rule will affect the results in the two-type model.

27. *See* RESTATEMENT (THIRD) OF TORTS: LIABILITY FOR PHYSICAL HARM § 3 (Tentative Draft No. 1, 2001); *see also* Omri Ben-Shahar, *Causation and Foreseeability*, in 2 ENCYCLOPEDIA OF LAW AND ECONOMICS: CIVIL LAW AND ECONOMICS 644, 662-63 (Boudewijn Bouckaert & Gerrit De Geest eds., 2000).

among strangers, the classic tort scenario can be described as a special, polar case within the contracts framework developed in this Article—the case of prohibitively high communication costs. Accordingly, the proposed model can be readily used to calculate the optimal cap on tort damages.

VII. CONCLUSION

I conclude with a reality check. There are two main challenges. The first concerns the institutional limits of common law adjudication. The preceding analysis shows how different parameters, specifically the magnitude of communication costs, the performance technology, and the distribution of buyer valuations, affect the optimal foreseeability threshold. Clearly courts do not estimate the magnitudes of the parameters identified in the theoretical model, and courts obviously do not apply a numeric algorithm to solve for the optimal damages cap before they decide whether to allow recovery in a specific case. Moreover, it is probably unwise to ask courts to engage in such an exercise.

Nevertheless, it is useful when evaluating the foreseeability doctrine to have a better understanding of the theoretical optimum. Consider, for example, the result that the optimal damages cap is significantly above average damages. At first blush, this result runs contrary to common doctrinal statements that “foreseeable” damages are average damages. On deeper reflection, however, these “average damages” statements hide a powerful tendency to narrowly apply the foreseeability requirement²⁸—a tendency that sits well with the theoretical result.

The optimality of a zero-damages default when communication costs are sufficiently small is another result that seems inconsistent with contract doctrine and especially with the dominance of the compensation principle. This result can be interpreted as reinforcing recent challenges to the compensation principle.²⁹ Viewed differently, this result may suggest that opting out of the default rule is not as easy as it seems. While communication costs narrowly defined are most likely small, there may be other factors contributing to the stickiness of default rules.³⁰ And if the default is sticky, then a zero-damages default becomes less attractive.

Even apart from its potential value for understanding or criticizing contract doctrine, the analysis in this Article can be useful for private contracting parties. Private parties have better information

28. See, e.g., 3 FARNSWORTH, *supra* note 8, at 259.

29. See, e.g., Scott & Triantis, *supra* note 9.

30. See sources cited *supra* note 22.

than courts. They also have powerful incentives to optimally design their contracts.

The second challenge to the relevance of the preceding analysis returns to the argument that communication costs are trivial, and thus the default rule does not matter. This is a powerful challenge.³¹ A main purpose of this Article was to define the contours of the triviality argument by quantifying the magnitude of communication costs necessary to render the default rule relevant. Even when communication costs are narrowly defined, as in this Article, there is a class of cases where getting the default right is of nontrivial importance. This class of cases will likely grow when broader costs of opting out of a default rule are considered.

31. In fact, it is a challenge to much of the default rules literature.