

University of Michigan Law School
Program in Law and Economics Archive: 2003-2009

Year 2007

Paper 80

How to Repair Unconscionable Contracts

Omri Ben-Shahar*

*University of Michigan Law School, omri@uchicago.edu

This working paper is hosted by The Berkeley Electronic Press (bepress) and may not be commercially reproduced without the permission of the copyright holder.

<http://law.bepress.com/umichlwps-olin/art80>

Copyright ©2007 by the author.

How to Repair Unconscionable Contracts

Omri Ben-Shahar

Abstract

Several doctrines of contract law allow courts to strike down excessively one-sided terms. A large literature explored which terms should be viewed as excessive, but a related question is often ignored—what provision should replace the vacated excessive term? This paper begins by suggesting that there are three competing criteria for a replacement provision: (1) the most reasonable term; (2) a punitive term, strongly unfavorable to the overreaching party; and (3) the maximally tolerable term. The paper explores in depth the third criterion—the maximally tolerable term—under which the excessive term is reduced merely to the highest level that the law considers tolerable. This solution preserves the original bargain to maximal permissible extent, but brings it within the tolerable range. The paper demonstrates that this criterion, which received no prior scholarly notice, is quite prevalent in legal doctrine, and that its adoption is based on powerful conceptual and normative underpinnings.

INTRODUCTION

Imagine the following situation. As a student in a law school course, you are regularly assigned a daily reading load of 20 pages in preparation for the next day's class. Yesterday, however, you received an extraordinary assignment of 200 pages, clearly more than you can feasibly prepare in one day. What should you do? Should you resort to the normal, "reasonable" practice of preparing 20 pages? Or should you, perhaps, disregard the unreasonable command of 200 altogether and in the absence of any other affirmative instruction read 0 pages? Or, perhaps yet, should you disregard only the unreasonable increment of the command and prepare the maximal tolerable level, of say, 50 pages?

This dilemma, I argue, is similar to one that is at the core of several basic doctrines of contract law. When a party with bargaining power dictates a contract term that is excessive or invalid, the law has to set a substitute provision. Should the excessive term be replaced by the most reasonable majoritarian term (analogous to 20 pages in the above hypothetical)? Should the dictating party be "punished"—incentivized not to go too far—by replacing the bad term with something least favorable to her (analogous to 0 pages)? Or, should the excessive term be reduced merely to the highest level that the law considers tolerable (analogous to 50 pages)?

This paper explores the problem of how-to-repair-excessive-terms and illustrates its solutions in existing law. There is no single compelling approach to this problem and, indeed, as Part I of the paper shows, all three solutions can be traced across different contexts and legal traditions. Still, the analysis in this paper focuses primarily on the third regime—the one that intervenes minimally and reduces the excessive term only to the maximally tolerable level. This regime received less analytical attention than I think it deserves.

It might plausibly be conjectured, before reading this paper, that the maximally-tolerable criterion is esoteric, an academic curiosity at best. Surely, so goes the conjecture, if a court takes the trouble to correct an excessive term in the contract, it would naturally replace that term with the most reasonable alternative, not with a barely tolerable one. Why let a party who overreached get away the maximum allowable advantage? The paper demonstrates, perhaps surprisingly, that the maximally-tolerable criterion is quite prevalent. It shows how courts use it in a variety of legal contexts as a mainstream solution to the problem of excessive and unconscionable terms. One such doctrine is *partial enforcement* (and its archaic predecessor, the *Blue Pencil Rule*), and its application in the context of covenants-not-to-compete. I show that when non-compete clauses are excessive, they are

generally brought down to the maximally tolerable level. Another doctrine studied here is *unconscionability*: what do courts do when a term is struck as unconscionable? Quite often, it turns out, the vacated term is replaced with the maximally tolerable one, most favorable to the strong party. Yet another example involves the judicial supervision of liquidated damages. In many legal traditions, excessive liquidated damages are reduced to the maximally tolerable level—the measure closest to the agreed sum, such that if it were the one agreed upon in the first place, the court would have enforced it.

Before it surveys the doctrinal prevalence of the maximally-tolerable criterion, the analysis in the paper sets up the context in which the criterion operates, and provides a theoretical foundation for such regime. Part I of the paper identifies the maximally tolerable regime as one of three discrete and conceptually coherent solutions to the problem of excessive terms. Part II of the paper then analyzes the conceptual grounding of the maximally tolerable criterion—the legal principles with which it is consistent and how it ties with other practices in various areas of the law. It shows that what underlies the maximally tolerable term approach is a specific conception of severability (or divisibility) of contractual provisions, and that the principle of waiver can lead to a similar solution. Further, this Part demonstrates the prevalence of similar solutions in other areas of the law, outside contract law.

Part III of the paper offers a brief cross-doctrinal survey and invites the reader to recognize that the maximally tolerable principle has broader and more subtle application than one might expect. Recognizing the instances in which this regime applies and when it is rejected sets the stage for the normative inquiry in Part IV, which suggests several justifications for this practice. The normative defense is anything but straightforward. Admittedly, there is something objectionable about a legal rule that accords the strong party, who already overreached in exploiting its bargaining strength, the maximally tolerable term. Why not reform the contract more aggressively? If courts already step in, why not take the opportunity to undo the effects of unfettered bargaining power? Legal solutions that favor the weak party, that level the playing field, are usually more appealing. Moreover, how do courts account for the bad incentives that maximally tolerable terms generate—the incentive to draft excessive terms, knowing that at most courts will only strike down the excrescence?

The key observation made here, in justifying the use of maximally tolerable terms, is that the problem of repairing unconscionable contracts is merely a species of *gap-filling*. The court that vacated an excessive term has to decide how fill the newly created gap. If the court merely provides a gap-filler, it cannot be

too ambitious and it cannot undo (even it tries) the existence of uneven bargaining power, or else its policy might backfire.

Indeed, the theme of this paper is part of a more general thesis concerning a principle of gap-filling in contracts between parties that have unequal bargaining power. This thesis, which I develop elsewhere,¹ suggests that standard gap-filling approaches do not provide a workable prescription when the gap in the contract involves a purely distributive aspect that parties bargain over (such as price). Gaps of this sort cannot be filled with “surplus-maximizing” terms because more than one term maximizes the surplus—in fact, the choice of gap-filler is surplus-neutral. As it turns out, gaps are purely distributive quite often. Ironically, many of the cases in contracts casebooks that introduce the topic of gap-filling involve purely distributive gaps over issues such as price, for which the prescription “choose the terms that maximize the total surplus” does not provide a definitive solution.

To resolve this indeterminacy, a new principle of gap-filling is needed, of mimicking the *bargain*: the division of the surplus that would have been struck between these parties, given the allocation of bargaining power. It is not a mimic-the-parties’-will conception—there is no joint will to mimic. Rather, it is a bargain-mimicking conception of gap-filling, which requires courts to fill gaps with terms that are sensitive to the division of bargaining power, more favorable to the party with the greater bargaining power.

Thus, continues my argument, if courts have to fill a gap that arises from the elimination of an explicit (but excessive) term, the bargain-mimicking conception would dictate the term closest to the hypothetical bargain. More aggressive intervention would fail to achieve any redistributive results that might superficially underlie it.

In the end, though, the paper recognizes that any justification for the maximally tolerable rule must account for the incentive problem and for the concern that this rule would induce strong parties to draft excessive terms. The analysis concludes by showing that this concern limits the application of the maximally tolerable regime, both in theory and in practice.

This paper intends to fill a vacuum in the study of unconscionability and related doctrine. Much ink has been spilled on the questions what is (and should be) the threshold of unconscionability and when intervention is justified. In the last

¹ Omri Ben-Shahar, “A Bargaining Power Theory of Gap Filling”, Mimeo. (University of Michigan Law School, 2007).

decade this question returned to fore when various types of mandatory arbitration terms were held to be unconscionable. But no systematic discussions emerged regarding the “remedial” aspect—how to repair contracts that contain excessively one-sided terms. This paper provides a conceptual framework to consider this problem (by identifying the three competing solutions) and takes a first shot at justifying one possible solution.

I. THE PROBLEM OF EXCESSIVE TERMS

A. Excessive Terms

When bargaining power is unevenly distributed, the strong party would naturally use its bargaining leverage to draft one-sided, self serving terms. It is a basic premise of contract law that courts ought not evaluate or even inquire into the adequacy of consideration, however unequal the values exchanged.² As long as a bargain was struck without coercion or fraud, each party bears the consequences of its poor bargain.

But when the unevenness of bargaining power leads to terms that are intolerable, courts are willing to step in. This might be true even for simple, easy-to-understand terms such as price, although such instances are extremely rare.³ Intervention is more likely to occur when the excessive terms are less conspicuous than the price and are less well understood by the weak party, suggesting that flaws existed in the manner in which assent was reached. This may be the case for late payment terms and late fees, disclaimers and exclusionary clauses. Intervention is even more justified when the boilerplate terms frustrate the induced legitimate expectations of the weak party—a phenomenon that is of particular concern in insurance contracts.⁴ More recently, an increasing number of courts find arbitration terms in consumer and employment contracts, which are excessively favorable to the drafters (sellers or employers), to contain unconscionable elements.⁵

² Restatement (Second) of Contracts § 79 Comment c., § 208 Comment d.

³ See WHITE AND SUMMERS, 1 UNIFORM COMMERCIAL CODE § 4.5 at 223 (4th Ed. 1995) (“reported litigation based on excessive price has dwindled to a trickle.”)

⁴ Restatement (Second) of Contracts § 211(3), cmt *f*; C&J Fertilizer v. Allied Mutual, 227 N.W.2d 169 (Ia. 1975).

⁵ Susan Randall, *Judicial Attitudes Toward Arbitration and the Resurgence of Unconscionability*, 52 Buffalo L. Rev. 185 (2004); Armendariz v. Foundation Health Psychcare Services Inc., 6 P.3d 669 (Cal.2000); Brower v. Gateway 2000, 246 A.D.2d 246 (N.Y. 1998).

The unconscionability doctrine is not the only way to limit excessive one-sidedness in contracts. In some specific areas, terms that are drafted one-sidedly may be struck without reference to the unconscionability standard. For example, liquidated damages that are clearly over-compensatory are considered punitive and unenforceable.⁶ Although the standard determining what constitutes unreasonably large liquidated damage can at times be identical to the “shock the conscience” standard of unconscionability,⁷ it is generally less strict. Or, in another area, covenants not to compete with a business or an employer may be found intolerable if the duration or the geographic scope of the non-compete obligation is too long or too broad. Finally, in some areas of contracting there are statutory caps that determine the maximal allowable stretch of the bargain. Usury laws and price gouging acts are typical examples of a maximum constraint; lemon laws and minimum wage laws are examples of minimum constraints. If the contract contains bargained-for terms that are outside these regulatory limits, the terms can be struck.

When an excessive term is invalidated, there is a gap in the contract that calls for gap-filling. It should be noted, though, that courts are not always ready to fill the gap and historically they have elected not to do so. Instead, when a contract contained an unconscionable element, the entire contract was rendered unenforceable.⁸ It was not considered the role of the courts to write the contract over, in a more reasonable fashion, for the parties. Even today, when the unconscionability of some terms is linked to flawed assent, as in the case of duress or fraud, courts may refuse to supply the more reasonable terms and instead vacate the entire agreement.⁹ Beyond the contract law consequences, overreaching can give rise to civil and criminal penalties such as disgorgement, fines and damage multipliers.¹⁰

Modern courts have less trouble reforming the contract and enforcing it. Unconscionability statutes provide clear authority to do so. For example, Section 2-302 of the Uniform Commercial Code allows courts to “limit the application of any unconscionable clause as to avoid any unconscionable result.”¹¹ Even more explicit, under the Principles of European Contract Law, a court may “adapt the contract in order to bring it into accordance with what might have been agreed

⁶ UCC §2-718(1) (“A term fixing unreasonably large liquidated damages is void as a penalty.”)

⁷ See, e.g., *Leeber v. Deltona Corp.*, 546 A.2d 452 (Maine 1988).

⁸ See *Earl of Chesterfield v. Janssen*, 28 Eng.Rep. 82 (Ch.1750).

⁹ See Richard Craswell, *Property Rules and Liability Rules in Unconscionability and Related Doctrines*, 60 U. of Chi. L. Rev. 1 (1993) (comparing the two approaches).

¹⁰ See Federal Price Gouging Prevention Act, H.R. 1252, 110th Cong. §§ 3-4 (2007); Fair Labor Standard Act, 29 U.S.C.A §16(b).

¹¹ UCC § 2-302(1). See also Uniform Consumer Credit Code § 5.108(3).

had the requirements of good faith and fair dealing been followed.”¹² There remains some debate whether courts can affirmatively replace the offending term with a different one not drafted by the parties, or whether they are restricted to crossing out existing terms and letting the legal gap-fillers set in. Section III of the paper reviews some of this debate. It demonstrates that many courts believe that they have sufficient authority to reduce an excessive term to any level they deem appropriate.

When the court decides to enforce the contract and reform the excessive term, it has to apply some principled policy in choosing the new term. True, this is not the “pure” gap-filling scenario of a contract that contains a lacuna and needs supplementation. Rather, this is an artificial gap that arises out of a legal policy that eliminates an existing express term. Nevertheless, this *is* a situation in which technically there is no longer an express provision in an otherwise enforceable contract and a new provision needs to be supplied. In the same way that courts need to turn to a gap-filling methodology when they knock out non-matching express terms in a battle of the forms,¹³ courts need to follow a systematic gap-filling pattern when repairing unconscionable terms.

B. Three Solutions

When the excessive term is struck down and a gap is created in the contract, how should this gap be filled? Courts generally follow one of three possible approaches.

1. *The Most Reasonable Term.* The standard criterion for filling gaps in contracts is to supply the most reasonable, majoritarian term. While this criterion is more often identified with gap-filling in indefinite contract or as a solution to the battle of the forms, it is also a sensible solution to the gap arising from the invalidated excessive term. Thus, if the price is unconscionable, replace it with a reasonable, intermediate level market price.¹⁴ If a liquidated damages remedy is excessive and punitive, replace it with standard expectation damages, measured by the contract-cover differential or by lost profits, excluding any uncommonly high consequential damages.

¹² Principles of European Contract Law, Article 4.109 (2002).

¹³ It is common to apply gap-fillers in situations in which the gap arises, not from indefiniteness in drafting, but from the knock-out of an express term. Compare UCC §2-204(3) (gap-fillers apply to “open terms”) and UCC §2-207(3) (gap-fillers apply as a result of the knock out of express terms.)

¹⁴ See, e.g., Aristides N. Hatzis and Eleni Zervogianni, *Judge-Made Contracts: Restructuring Unconscionable Contracts* (Mimeo., 2007).

Often, this solution of supplying the most reasonable mid-range term arises implicitly. A court might strike the offensive term and make no affirmative substitution. The default rule would then apply, effectively supplying a reasonable provision. For example, if an unconscionable arbitration term is struck, the court need not select an affirmative gap-filler; in the absence of an arbitration provision, parties resolve their dispute in court. Or, if an exclusionary clause in an insurance policy is unconscionable, it falls and the insurer's obligation is read without it. Here, even if the court does not pay explicit attention to the principle underlying the gap filling process, it often chooses an outcome that is consistent with a most-reasonable-term approach.

There is much to be said, of course, in support of this regime. The most-reasonable-terms are, by definition, the most compelling ex-post solution. Contractors are allowed to deviate from them, within limits, and a court would normally let such deviations stand. But if the court deemed the situation fit for intervention, why not take the opportunity and write the most balanced contractual term for the parties?

2. *The Most Unfavorable Term.* If the drafting party overreached by trying to secure an excessive gain, the court can "punish" this behavior by depriving this party of the entire advantage. The contract is enforced, but the excessive term is supplemented by a term least favorable (within reason) to the drafting party. Thus, for example, if a creditor bargained for excessive, usurious interest rate, a court can replace it with 0% interest. If the duration of a non-compete clause is excessive, the court can strike altogether, effectively replacing it with zero duration. In fact, in some scenarios the example mentioned above of a vacated arbitration term can be regarded as an illustration of the most-unfavorable-term approach. The most reasonable gap-filler, it can be argued, is a fair and balanced arbitration arrangement that would replace the one-sided arbitration term that was struck. Eliminating arbitration altogether and sending the dispute to court is a way to punish the overreaching party and tilting the result in favor of the other party.

This gap filling approach is punitive in the same way as the doctrine of *contra proferentum*. It is not intended to identify the most balanced outcome, ex post. Instead, it is intended to induce the drafter to make drafting choices that would not overreach and would not necessitate court intervention in the first place. It is a species of a penalty default rule, and as such it is consistent with the same policy concerns that supposedly justify penalty rules: forcing a party who enjoys a bargaining advantage (here, superior bargaining power) to forgo some of the gains that he can extract.

3. *The Maximally Tolerable Term.* Finally, if a term is considered excessive, it can be broken down to two distinct components: the maximally tolerable portion, and everything beyond it. Once the second component—the excessive increment is eliminated, the remainder is no longer unconscionable (even if still relatively one-sided), and does not necessitate further intervention. This remainder—the maximally allowable term—would be enforceable.

To compare the three solutions, consider a situation in which the reasonable price of a service is \$500. If purchased under conditions of a thick market, the price would always be \$500. But situations arise in which one of the parties may experience urgency or vulnerability, or, alternatively, enjoy bargaining leverage, such that the price for the same service may reflect those circumstances. Assume that it is not unreasonable to charge as much \$750 to a buyer-in-need, or to pay as little as \$250 to a desperate seller. Thus, any price within the range of \$250 and \$750 would be enforceable. But what if the service provider exploits her bargaining power to charge an unreasonable price of \$1000, and this price gets struck by court? Under the most reasonable term approach, the gap-filler would be \$500—this is the term that most comports with the community standards of fairness.¹⁵ Under the unfavorable term approach, the gap filler would be \$250 (or perhaps even lower, if the stated purpose is not merely to repair the contract, but to punish the offender.) And under the maximally tolerable term approach, the gap-filler would be \$750—the maximal price within the tolerable range.

II. MAXIMALLY TOLERABLE TERMS: A CONCEPTUAL INTRODUCTION

Two of the three solutions discussed above are familiar to most readers. Both the most-reasonable-term and the unfavorable-term approaches have an intuitive appeal and are based on premises that are shared by more prominent theories of gap-filling and contract interpretation. They can be viewed as analogous to the two familiar criteria for default rules – majoritarian and penalty defaults. I have nothing more to say about these approaches and I will therefore focus in the remainder of this paper on the third regime—the maximally-tolerable-terms. This regime is based on a principle that many might find, upon first reflection, objectionable. It allows a party who acted poorly (by overreaching) to escape with minimum sanction. This leniency might seem both unjust and a weak deterrent. Since my goal here is to dispel some of this intuitive rejection of the maximally tolerable principle, I begin in this section with analysis of its conceptual grounding. I hope to show that that the maximally tolerable principle is consistent

¹⁵ Restatement (Second) of Contracts § 204 Comment d.

with some fundamental premises regarding contract enforcement. This will help understand, later, why this rule is quite prominent in practice.

A. Maximum Terms and the Freedom to Bargain

Parties are entitled to engage in tough negotiations, maneuver for advantages, and insist on self-serving contractual terms. Drafting a contract that contains terms other than the most reasonable ones is not illegal nor is it uncommon. It is only when these advantages are excessive—when they reach beyond a level that is regarded as tolerable—that the law steps in to invalidate them. Thus, if a court is to reform the excessive contract, it is only the illegitimate element of the one-sided term that needs to be struck. Effectively, then, the court would fill the gap with a term that is still one-sided, still favorable to the same party who dictated the original excessive term, but moderated sufficiently so that it would be tolerable—so as to fit it within the range that is considered legitimate.

This solution preserves to the maximum the bargaining advantage secured in the contract. It is therefore the one most consistent with the idea that bargaining power ought to be respected, not undone. In a companion paper, I examine the merits of a new criterion for filling gaps in purely distributive terms (such as price). Under that criterion, gaps should be filled in a way that reflects the relative bargaining power of the parties.¹⁶ Specifically, the court-supplied term ought to resemble as much as possible the term that the parties would have negotiated expressly, even if such a term clearly favors the stronger party. This idea is based on normative grounding: if there is a range in which parties are allowed to bargain, the best that default rules can do is mimic the point within this range that the parties would hypothetically choose.

The maximally tolerable term that is discussed here is consistent with this more general bargain-mimicking idea because it reflects the relative bargaining power of the parties. To be sure, it is not a pure bargain-mimicking term. The perfect mimicking term was in fact written in the contract, and yet it was found unenforceable under a policy aimed at limiting the reach of bargaining power. Obviously, the court should not reinstate the same term it has just struck down. What the court would mimic here is the hypothetical bargain that parties *negotiating over a truncated domain* would reach. This solution accords with a “general duty of the court to preserve so much of a contract as may properly

¹⁶ Omri Ben-Shahar, “A Bargaining Power Theory of Gap Filling”, Mimeo. (University of Michigan Law School, 2007).

survive its invalid and ineffective provisions”¹⁷—namely, to enforce the entire contract minus the excessive increment.

B. Grounds for Intervention

Under doctrines like unconscionability which are intended to provide relief from extreme terms, there is no clear authority or justification for courts to provide more than the minimum necessary relief. The court’s authority to intervene in the contract and to police its terms arises from the fact that an express term lies beyond society’s tolerable range. The further out this term relative to what is tolerable, the greater the justification for intervention. Once the offensive term has been pushed into the tolerable range, even if only barely so, there is no remaining justification for intervention.

One way to reinforce this idea is by considering the following synthetic illustration. Imagine again the case in which the contract contains an excessive price, \$1000, and suppose that the process of adjusting the price involves a gradual examination of successively lower prices. That is, after deciding that \$1000 is too high, the court considers one price lower, say \$990. If that price is also unconscionable, the court considers a further incremental reduction, to \$980. It continues similarly step-by-incremental-step until it reaches a price which is no longer considered intolerable. Once that price is reached, the process of adjusting the price downwards would then stop. There will be no remaining justification—at least not under direct policy grounds that give rise to the unconscionability redress—for further adjustments of the price beyond this threshold.

Put differently, if we analogize the process of judicial intervention in the contract to a force that pulls the price from its current intolerable level towards the permissible region, the force gradually weakens as the price gets closer to the tolerable level, and vanishes completely as soon as this level is hit. The point where this adjustment process runs out of justification is not the mid-range, majoritarian, most-balanced term. Rather, it is the maximally tolerable price: it is still a one-sided term, albeit not as bad as the original term.

Conceptually, this argument is consistent with the language that authorizes court intervention in unconscionable contract. If, as the UCC instructs, courts are authorized “to limit the application of any unconscionable clause as to avoid any unconscionable result,”¹⁸ this authority to further tinker with any clause expires once the term is no longer unconscionable.

¹⁷ Anders v. Hometown Mortgage Services Inc., 346 F.3d 1024, 1032 (Ala. 2003).

¹⁸ UCC § 2-302

This argument goes some distance towards justifying the maximal tolerable terms. It is based on the logic that if legal intervention in the contract is justified by a particular distributive concern, it is also limited by this very concern. There are several equivalent ways to articulate this claim. One is to compare a contract that contains the maximally tolerable price term with a contract that is all else equal, but contains an even worse price term. If there is no good reason to intervene in the former, is there a reason to intervene in the latter beyond fixing it to look like the former? We shall see later, there may be incentive-based reasons for a more aggressive intervention in the latter contract. But those are different than the distributive concerns justifying the intervention in the first place. Another equivalent way of saying this is to focus on the complaint of the weak party. This party has no reasonable grounds to demand more than the minimal redress tailored by the maximally tolerable term. Once accorded this adjustment, what basis does he have for demanding additional relief?¹⁹

C. Cross-Doctrinal Foundations

This conceptual defense of the idea of minimum-necessary-relief accords with other deep-rooted practices of the law in general, and contract law in particular, which entitle a party to *concede* a greater contested claim in order to secure a smaller uncontested claim. For example, the doctrine of *remittitur* deals with excessive jury verdicts in civil trials. The judge determines the portion of the verdict that is excessive and gives the plaintiff the option to remit—to *concede*—this increment, or face a new trial. The verdict is not entirely voided, nor is it replaced with the most reasonable amount. Rather, only the excrescence—the sum which exceeds “the highest amount which the jury properly could have awarded”—is lopped off.²⁰ Many courts have specifically rejected the more intrusive approach, which reduces the excessive judgment to the level that is most fair. Instead, they prefer the minimal intervention approach, reducing the judgment “to the maximum that would be upheld by the trial court as not excessive.”²¹

¹⁹ This rationale is recognized by Corbin: “The line [representing the enforceable term] must be drawn somewhere, and it is drawn at the point where the protection to which the buyer is justly entitled ends.” See Arthur L. Corbin, *On the Doctrine of Beit v. Beit: A Comment*, 23 Conn.B.J. 40, 46 (1949).

²⁰ *Dick v. Watonwan Co.*, 562 F. Supp. 1083, 1108 (1983); See, generally, 11 WRIGHT, MILLER, AND KANE, *FEDERAL PRACTICE AND PROCEDURE* 167 (2d ed. 1995).

²¹ *Earl v. Bouchard Transportation Co., Inc.*, 917 F.2d 1320, 1328 (2d Cir., 1990)

There is more than geometric resemblance between remittitur and the maximally tolerable terms principle. Both are based on a premise that an “outcome” can lie within a range, and as long as it is within the range there is no ground for intervention. In the remittitur, context, the outcome is the jury’s judgment regarding damages. It does not have to be a balanced mid-range compromise between the litigants opposing positions. It can lean towards one party, but not in an unreasonable, intolerable fashion. If it goes outside this region, it is pushed back in. This practice is analogous to the maximum tolerable terms principle in contract law.

Within contract law too, one can find the roots of the idea that excessive provisions can be cured by incremental, rather than total, invalidation. A party may concede a gap in the contract in favor of the other party in order to cure indefiniteness and enforce the conceded contract.²² Indeed, as we shall see in the next section, it is quite common for parties who lose their case in defense of an unconscionable term to concede the offensive element and ask the court to enforce the remainder. Many courts are receptive to such requests.²³

More generally, the idea of minimum necessary relief can be embedded in the doctrine of *waiver*. The drafting party is treated as accepting a reduction of the self-serving term, waiving her right to insist on full unlimited enforcement of that term. With the waiver in place, there is no remaining ground for intervention. As the Supreme Court of Texas recognized in this context:

“Even though the contract may be illegal and unenforceable as written, one of the parties may make it legal and enforceable by offering to take out of it the offending provision that makes it illegal...”²⁴

D. Identifying the Maximal Threshold

While the criterion underlying the maximally tolerable terms—the upper bound of the legitimate range of contracting—may be conceptually coherent in the abstract, is it implementable by courts? Is it possible to adjudicate this criterion and identify the threshold? There are aspects to this inquiry regarding the practicality of the criterion that implicate the normative discussion, and will be postponed till later. For it might be possible to identify the maximally tolerable threshold but

²² 1 Farnsworth on Contracts §3.29 (3rdEd. 1999); Ben-Shahar, *Agreeing to Disagree: Filling Gaps in Deliberately Incomplete Contracts*, 2004 Wisc. L.Rev. 389, 421.

²³ See, *infra*, text accompanying notes 48 to 53.

²⁴ *Lewis v. Krueger, Hutchinson and Overton Clinic*, 269 S.W.2d 798, 800 (Tex. 1954).

only at an increased adjudication cost, in which case the normative argument in favor of this criterion would weaken.

I believe that there are many scenarios in which it is possible to identify a maximally tolerable level. Section III below demonstrates how courts manage fairly easily to follow this criterion. In some situations the threshold is easy to identify because it is expressly established by regulations. For example, in some States the maximal scope of non-compete clauses in employment contracts is set by statute. Likewise, the maximal interest rate to be charged for credit or for late payment is regulated in every jurisdiction. Contracts that exceed these thresholds can easily be reformed to the maximal statutorily permissible levels. We will see, however, that in situations in which the threshold is crystal clear, there is also a stronger suspicion that the overreaching occurred in bad faith. In these cases, there is a stronger argument that the exceeding party ought to be *punished* and deprived of more than just the excess.

In other situations, where the maximum threshold is not set forth in a statute, courts can follow the principle of maximally tolerable terms without pinpointing the precise maximum. As long as they do not err above the maximum, they can cautiously set a term that, while still favorable to the drafting party, is nevertheless not unreasonable. In the example above, the court does not need to know that \$750 is the uppermost tolerable level that can replace the struck-down price of \$1000. While the plaintiff is going to ask for a gap-filler equal the most reasonable term of \$500, the defendant might offer a higher price term, say \$650 (perhaps the defendant already collected \$650 and is willing to stop there.) As long as the defendant's proposed term is within the tolerable range, it is consistent with the maximally tolerable criterion.

In some situations, a party drafting a mass-market contract may "experiment" with a one-sided term (e.g., arbitration term, or exclusionary clause) that is eventually held by courts to be overreaching. In time, the drafting party will modify the boilerplate contract and offer a less extreme version of the one-sided term to new customers. But what about the old customers that accepted the original, excessive term? In these settings, the existence of a new, modified version that survives the scrutiny of courts makes a good candidate for the maximally tolerable term. This modified term would likely still be one-sided relative to the most reasonable or balanced term, but it would be tolerable.

Moreover, there is ample evidence that the parties themselves believe that the principle of maximally tolerable terms is implementable. As I will mention in the next Section, many contracts include severability or savings clauses that instruct

courts to enforce any provision that is found to be excessive *to the maximal extent permitted by law*. If an interest rate is excessive, the maximal permissible rate should apply; if a warranty disclaimer is too harsh, the maximal permissible disclaimer applies; and so on. The parties themselves, it appears, believe in this criterion.

In the end, it might still be difficult in some situations to identify a discrete threshold for the maximally tolerable term; it might be much easier to resort to market data and apply the most reasonable, majoritarian term. But the thesis of this paper does not fail even under this shortcoming. The thesis is proposed in a more tentative fashion: maximal tolerable terms ought to be considered as a way to repair unconscionable contract *to the extent that the court has information about the maximal tolerable threshold*. Such information, in other words, ought not to be ignored.

III. DOCTRINAL APPLICATIONS

This section examines some existing practices that are consistent with the maximally tolerable terms approach. It also identifies some instances in which this criterion was expressly rejected. Overall, I hope to show that in many subtle ways contract law applies an approach that is close to maximally tolerable terms regime.

A. Severability Clauses

Before turning to legal rules that determine how to repair excessive terms, it is important to note that such principles can be enacted in the contracts themselves. Maximally tolerable gap-fillers can emerge in practice as a result of contractual drafting that instructs courts to apply such a criterion. Specifically, a party who drafts a self-serving standard form contract often adds a boilerplate severability clause, stating:

If any provision in the contract is not permitted by governing law, such provision shall not be construed to be null and of no effect, but court shall construe the agreement to provide for the maximum enforceable application.²⁵

²⁵ Laurence H. Pretty, Patent Litigation, ex. 15-24, pp. 15-41 to 15-42. See also Charles Sennewald, *Security Consulting* (“If the scope of any of the provisions of the agreement is too broad in any respect whatsoever to permit enforcement to its full extent, then such provisions shall be enforced to the maximum extent permitted by law, and the parties hereto consent and agree that

Many contracts contain similar language, rendering a term enforceable “to the maximal extent permitted by law.” For example, a lending term may stipulate that the creditor is entitled to the “maximum rate” and that if the charged interest exceeds the maximal permissible rate, the creditor should refund only the amount of such excess.²⁶ Or, a warranty/liability term can disclaim all warranties or damages to the maximum extent permitted by law.²⁷ These are situations in which the parties anticipate the possibility that a one-sided term might be struck, and instruct the courts specifically how to fill gaps in the agreement. The instruction is for the court to follow a one-sided methodology, picking terms that are equivalent to the maximal tolerable provisions.

Of course, courts can disregard such severability clauses, perhaps on the basis that such terms oust the inherent jurisdiction of the court to choose a remedy for wrongful behavior. Or, courts might consider the effect of such clauses to be undesirable and contrary to public policy. But it would be difficult to strike these clauses as unconscionable, because by their very language they refer to the maximal level that is *not* unconscionable.²⁸

B. The Doctrine of Partial Enforcement

such scope may be judicially modified accordingly and that the whole of such provisions of this agreement shall not thereby fail...”)

²⁶ See, e.g., J Robert Brown and Herbert B. Max, *Raising Capital: Private Placement Forms & Techniques* (“payee shall never be deemed to have contracted for, or be entitled to receive as interest on this note, interest ...or any amount in excess of the Maximum rate. [...] if the interest received ... exceeds the maximum rate, then payee shall refund the amount of such excess.”) See also Richard T Williamson, *Selling Real Estate Without Paying Taxes: A Guide to Capital Gains Tax Alternatives* 173 (“no payment or interest shall exceed the maximum amount permitted by law. Any payment in excess of the maximum amount shall be [disbursed to the payor].”)

²⁷ See, e.g., Napster’s Terms and Conditions, available at www.napster.com/terms.html (“To the Extent that in a particular circumstance any disclaimer or limitation on damages or liability set forth herein is prohibited by applicable law, then instead of the provision hereof [...] Napster shall be entitled to the *maximum disclaimers and/or limitations on damage and liability available at law or in equity...*”). See also RealNetworks EULA (“*To the maximum extent permitted by applicable law, RealNetworks further disclaims all warranties*”); Open Source License, in Andrew M. St. Laurent, *Understanding Open Source and Free Software Licensing*; Altova term: (“Because some states do not allow the exclusion or limitation of liability, the above limitation may not apply to you. In such states, liability shall be limited *to the maximal extent permitted by law.*”) (*emphasis added*)

²⁸ Imagine, hypothetically, an arbitration clause that reads:

Any dispute arising out of this agreement will be arbitrated under the arbitration rules of the AAA and take place in the *state that is least convenient for the plaintiff, so long as this is not too inconvenient to constitute an unconscionable burden.*

Can this clause, with its built-in unconscionability constraint, be struck as unconscionable?

The doctrine of partial performance is a method that enables courts to enforce a term that is otherwise unreasonable or extreme in a partial, tolerable manner. The archaic “Blue Pencil Rule” was the origin of this method. Under this rule, when a contract contained an invalid term, the invalid portion would be literally crossed out (by the metaphoric blue pencil). If the language that remained was grammatically meaningful, it would be enforced. Otherwise, if the remainder was not coherent without some affirmative rewriting, it was entirely invalid.²⁹ Here, if the contract was selectively enforced, the remainder did not follow a principled criterion – maximally tolerable or any other criterion. The outcome had an arbitrary, inconsistent, aspect.

Yet there was an appealing feature to the Blue Pencil rule that the more modern approach to partial enforcement sought to preserve. It was a technique that allowed courts to depart from the older and even more rigid all-or-nothing regime, which simply voided any unreasonable provision in its entirety. If the provision were divisible, why not sever only the offensive part—the minimally necessary part—and enforce the remainder? While such divisibility was the policy underlying the Blue Pencil rule, the mechanical procedure of the rule significantly limited its effectiveness. A better rule for partial enforcement could implement the divisibility policy without constraining courts by a grammatical criterion.

Under the modern partial enforcement doctrine, a court is authorized to reform an unreasonable term in a contract and enforce it to extent necessary to avoid the unreasonableness. The court does not have to use the blue pencil method. It can do more than just cross out some language and enforce the remainder. It can, in addition, substitute the offensive language with a different provision. The underlying goal is to give maximal effect to the parties’ agreement, subject to the constraint of avoiding unreasonableness.³⁰ This goal it is often implemented by amending the excessive term with the maximal tolerable term.

Perhaps the most striking (and most common) application of this partial performance technique involves covenants not to compete. When an employee enters an employment contract, he often signs a non-compete clause with the employer, applicable in the event that the employment ends. Similarly, when a business is sold with its good will, the seller often promises to not compete with the buyer. These restraints are at times too harsh, either by setting too long a duration of the non-compete period, or by defining the geographical boundaries of the non-compete region too broadly. Old decisions in the all-or-nothing tradition used to void these as unreasonable restraints altogether, and leave the parties free

²⁹ FARNSWORTH, CONTRACTS § 5.8 (4th Ed. 2004)

³⁰ Williston and Corbin, *On the Doctrine of Beit v. Beit*, 23 Conn.B.J. 40, 49-50 (1949).

of any restraint.³¹ Even today, in some European legal traditions such excessive restraints are considered void and may not be adjusted by courts.³² Blue Pencil decisions—somewhat less strict—granted relief by partially enforcing the restraints if they were grammatically meaningful without the offensive words. But most courts nowadays substitute an offensive term with a maximally tolerable one.³³

At times, the maximal tolerable level is defined explicitly by statutes. Some states have enacted bright line rules stating the maximal duration of non-compete clauses in employment contracts. In these states, when the contract contains a non-compete term that is longer than the statutory cap, it is normally truncated to be equal to that cap.³⁴ That is, only the increment of the restraint that is socially intolerable is eliminated; the rest stands. In other states there is no bright line statute. There, too, courts reduce the non-compete term, bringing it down to a level that is tolerable. The restraint “is not enforceable beyond the time or area considered reasonable by the court.”³⁵ The permissible duration varies across circumstances and jurisdictions. Case law provides numerous examples for the cap being set between 6 months to 10 years.³⁶

This approach can, and in fact applies to other contexts in which excessive terms place unreasonable restraint on a party. Thus, for example, in a German case

³¹ The earliest case is *Mitchell v. Reynolds* (1711) 1 P Wms 181, 24 Eng Reprint 347. This rule remained in force in England and in the U.S till the end of the 19th century. See C.T. Drechsler, *Enforceability of Restrictive Covenant, Ancillary to Employment Contract, as Affected by Territorial Extent of Restriction*, 43 A.L.R2d 94, §6(a).

³² See Aristides N. Hatzis and Eleni Zervogianni, *Judge Made Contracts – Restructuring Unconscionable Contracts*, p. 6 (Mimeo., 2007) (citing German Law).

³³ See, e.g., Harlan M. Blake, *Employee Agreements Not to Compete*, 73 Harv. L.Rev. 625, 646-651 (1960); *Solari Industries, Inc. v. Malady*, 264 A.2d 53 (N.J. 1970.).

³⁴ See, e.g., §542.335(1)(d)1, Fl. Stat. (1997) (“a court... shall presume unreasonable in time any restraint more than 2 years in duration”), enforced in *Flickenger v. Fitzgerald & Co.*, 732 So.2d 33 (Fl. 1999); La. Rev. Stat. Ann. §23.921C (2006); S.D. Codified Laws §53-9-11 (2007).

³⁵ See, e.g., *Justin Belt Co. v. Yost*, 502 S.W.2d 681, 685 (Tex. 1974). In other countries the unlawfulness of the restraint is a matter for the courts to determine, but again with the consequence that if found unlawful, the restraint is not struck down *in toto*, but reduced to its maximally tolerable level. For a French decision, see Cass. Soc., 21 October 1960, JCP 1960.II.11886, discussed in BEALE ET AL., *CASES, MATERIALS, AND TEXT ON CONTRACT LAW* 312-313 (2002)

³⁶ See, e.g., *Kramer v. Robec, Inc.*, 824 F.Supp. 508, 513 (E.D. Pa. 1992) (Pennsylvania court reducing a 3 year restraint to 2 years; enforcing a geographical limit of the entire U.S.); *Fullerton Lumber v. Torborg*, 70 N.W.2d 585 (Wisc. 1955) (Wisconsin court reducing a 10 year restraint to 3 years); *American Weekly v. Patterson*, 16 A.2d 912 (Md. 1940) (Maryland court reducing a 5 year restraint to 4 years); *Foltz v. Struxness*, 215 P.2d 133 (Kan. 1950) (Kansas court enforcing a 10 year restraint)

dealing with an exclusive supply arrangement, in which the buyer was obligated to purchase its entire requirements over 24 years from the seller, the court reduced the exclusivity term to the maximal tolerable duration of 16 years.³⁷

The stated aim of this jurisprudence is to protect one party's interest, as displayed in the contract, without unreasonable hardship upon the other party. While this is a cost-benefit test, it would be wrong to conclude that the courts are enforcing the most efficient or most reasonable gap-filler.³⁸ Rather, courts view their role as securing the bargain that the parties struck, recognizing the superior bargaining power of one of the parties. The cost-benefit test is used to identify, not the surplus maximizing term, but only what is excessive. A limitation that burdens an employee without according benefit to the employer is deemed unreasonable. Thus, in eliminating the unreasonable portion of the restraint, the court is setting not the most reasonable or common term, nor a majoritarian or average provision, but rather the maximally tolerable one.

It is important to recognize that the application of maximally tolerable terms in this context is limited by a safety valve which will be discussed in more detail in the normative analysis below. The concern is that the doctrine of partial enforcement and the application of maximally tolerable terms might give drafters incentives to dictate overly oppressive restraints, expecting to lose at worst only the excessive increment but to keep it anytime it is not challenged. In light of this concern, courts are ready to invalidate the entire non-compete clause (namely, replace the duration term with zero) if there is evidence of deliberate overreaching.³⁹ Maximally tolerable terms apply only if the crossing of the boundary was done without bad faith. It might be difficult at times to ascertain whether there was bad faith. A presumption of bad faith may exist when the non-compete restraint is unlimited in duration, and often courts indeed vacate such restraint completely.⁴⁰

C. Unconscionability

1. *Price and Interest Rates*

³⁷ BGH 16 and 17 September 1974, NJW 1974.2089. See also Beale et. al., *supra* note 35, at 313-314.

³⁸ In many contexts, the most common duration gap-filler is at-will termination—a zero restraint on the right of an employee to walk away. See, e.g. UCC 2-309(2) (at-will termination in sales contracts); Payne. V. Western & Atlantic R.R., 81 Tenn. 507, 519 (1884) (termination of employment contract).

³⁹ See, e.g., Central Adjustment Bureau v. Ingram, 678 S.W.2d 28, 37 (Tenn.1984).

⁴⁰ E.g., Harvest Ins. Agency v. Inter-Ocean Ins. Co. 492 N.E.2d 686, 690 (Ind. 1986).

When a term in the contract is struck under the doctrine of unconscionability, courts have broad discretion how to repair the contract. A common solution would be to strike the offensive term and replace it with nothing, enforcing the remainder of the contract. This, of course, is an implicit choice to replace the excessive term with the default rule—the gap-filler that applies in the absence of an express term. Thus, when the cross-collateral term was struck in *Williams v. Walker-Thomas*,⁴¹ it was replaced with the default rule that granted the seller no special security in the buyer’s prior purchases.

This solution—replacing the offensive term with the legally supplied default provision—would rarely end up with the maximally tolerable term. The legally supplied gap fillers are normally the majoritarian, most-reasonable provisions and do not favor the drafting party. Still, there are many circumstances in which the court chooses to affirmatively redraft the term, and to set it at the maximally tolerable terms. This is achieved by striking not the entire offensive term, but only the increment of it that is found unconscionable. The remaining component could still be one-sided, but no longer unconscionable. Corbin seems to have recognized this criterion when he explained, in the context of a loan of money, that “a contract that requires a payment of a very high interest will be enforced, *up to the point at which ‘unconscionability’ becomes an operative factor.*”⁴²

In many cases, courts have effectively replaced an unconscionable term with something that resembles the maximal tolerable term. Take, for example, the classic door-to-door sale case, *Toker v. Westerman*,⁴³ in which the buyer agreed to pay over \$1200 for a refrigerator that normally sells for \$400. After the buyer paid more than \$650, he sought relief from the oppressive price. The court indeed struck down the price as unconscionable, but allowed the seller to keep the money already paid. True, the court did not directly hold the seller is entitled to the maximal tolerable price. The \$650 figure happened to be the amount already paid when the case was initiated. The buyer only asked for the remainder to be unenforceable; he did not seek restitution of some of the money previously paid. But in another case, in which the price was not yet paid, the court specifically reversed a lower court’s stipulation of a low net-cost price, and allowed the seller

⁴¹ *Williams v. Walker-Thomas Furniture Co.*, 350 f.2d 445 (D.C. Cir. 1965).

⁴² 1 Corbin, *Contracts* § 129, p. 556 (1963) (*emphasis added*).

⁴³ 274 A.2d 78 (NJ, 1970). See also *Jones v. Star Credit Corp.*, 298 N.Y.S.2d 264 (N.Y. Sup. Ct. 1969) (a \$300 freezer was purchased for a price that exceeds \$1400; the court allowed the buyers to stop payments after \$620 were paid); *Bank of Indiana Bat'l Ass'n v. Holyfield*, 476 F.Supp. 104 (Miss. 1979).

to collect a price that included all indirect costs plus a reasonable profit. For a product that has a thick market, this comes close to a maximal price.⁴⁴

Another implicit application of this approach comes from the decision of the district court in the famous *Batsakis v. Demotsis* case. Recall that this case involved a loan in Greek currency made in Greece during the war, which, in nominal terms, amounted to the equivalent of \$25. In return, the debtor promised to pay \$2000 plus interest after the war. When the time to pay the debt came, the debtor reneged. The District Court in Texas found that the promise to pay \$2000 for a loan of \$25 was not enforceable, for lack of consideration. Sympathizing with the debtor, but recognizing also that the contract cannot simply be voided and undone, the court ordered the debtor to pay \$750 to satisfy the debt. This result, we know, was overturned by the Court of Appeals, which reinstated the obligation to pay \$2000.⁴⁵ For our purpose, however, it is the lower court's decision that is of interest, because it is only this court that found grounds for intervention in the first place. This court effectively invalidated the \$2000 price and thus needed to fill a gap. It did not turn to the most reasonable and balanced term, nor to a term that comports with community standards of fairness. Rather, it used a one-sided term, just within what it perceived to be the tolerable range.

In credit transactions, the interest rate cannot exceed statutory established threshold or else it is considered unenforceable usury. But what happens when it does? What is it replaced with? In one case, the court struck the agreed upon interest rate of 200% (that was intended to apply to a short term but ended up applying over a longer duration), and replaced it with a 24% rate. It is not clear whether this was the maximal allowable rate, but it was somewhat higher than the prevailing market rates of 18-21%.⁴⁶ In Austrian law, for example, a similar maximal tolerable provision is supplied: a usurious interest rate is adjusted, brought down to equal *double* the basic interest rate. In French Law, the excessive interest rate is reduced only to the extent that they exceed the interest rate allowed by law.⁴⁷ In all these examples, a maximally tolerable criterion underlies the practice.

2. Arbitration Clauses

⁴⁴ *Frostifress v. Reynoso*, 274 NYS2d 757 (Sup. Ct. 1966), rev'd as to damages, 281 NYS2d 964 (App. 1967).

⁴⁵ *Batsakis v. Demotsis*, 226 S.W.2d 673 (Tex. 1949)

⁴⁶ *Carboni v. Arrospide*, 2 Cal.Rptr.2d 845 (1991).

⁴⁷ *Hatzis and Zervogianni*, *supra* note 14, at 7-8..

The same result of implementing the maximally tolerable term is obtained when the court decides to sever the offensive component of a term and enforce the remaining part. This practice has been widely followed in the context of unconscionable arbitration clauses. Consider one of the recent leading cases—*Brower v. Gateway 2000*.⁴⁸ There, the arbitration term in a consumer contract was unconscionable because it placed unreasonable filing cost and location burdens on the consumer. The court vacated these elements, but stopped short of eliminating the entire arbitration term (as the consumer would have liked). Instead, it remanded the case to the lower court to figure out a more reasonable arbitration forum. Specifically, it acknowledged Gateway’s proposal to use a somewhat less onerous arbitration proceeding (one that Gateway offered its new customers in the modified version of their shrinkwrap) and instructed the trial court to evaluate whether this proposed venue comes within the tolerable range.

Many courts apply a similar approach to repair other elements of unconscionable arbitration clauses.⁴⁹ In another case, the court severed only the element of the arbitration clause that required the buyer to reimburse the seller for its arbitration and attorney fees, but left everything else in tact, including the one-sided authority of the seller over the choice of arbitration.⁵⁰ Or, in another case, when the cost of arbitration was unaffordable to one party, the court allowed the other party to fix the problem by making an ad-hoc concession to pay “what we need to pay to make [arbitration] fair”—even if only so much as to make the arbitration affordable to this particular plaintiff.⁵¹ Generally, when the arbitration agreement is found to be patently unconscionable by cutting down individuals’ Federal statutory rights (such as the right to seek punitive damages), courts have reached the result that is consistent with the maximally tolerable criterion, either by giving full force to the severability clauses and vacating only the offensive exclusionary elements,⁵² or by allowing the party who is seeking arbitration to waive the elements that are unreasonable.⁵³

But courts do not always resort to this principle of striking down only the minimum necessary to bring the unconscionable term within the tolerable region. In the leading California case, *Armendariz v. Foundation Health Psychcare*, the court found some elements of an employment arbitration clause to be

⁴⁸ *Brower v. Gateway 2000*, 246 A.D.2d 246 (N.Y. 1998).

⁴⁹ See, e.g., *Swain v. Auto Services, Inc.*, 128 S.W. 3d 103 (Mo. 2003).

⁵⁰ *Vincent v. Schneider*, 194 S.W.3d 853 (Mo. 2006)

⁵¹ *Anders v. Hometown Mortgage Services, Inc.*, 346 F.3d 1024 (11th Cir. 2003).

⁵² *Id.*, at 1031-32; *Ex parte Thicklin*, 824 So.2d 723, 734 (Ala.,2002). But see *Sosa v. Paulos*, 924 P.2d 357 (Utah 1996) for the opposite result.

⁵³ *Ex Parte Celtic Life Ins. Co.*, 834 So.2d 766 (Ala. 2002) (“a party to a contract can waive a contractual provision beneficial to that party”).

unconscionable and struck down the entire clause, effectively filling the gap with a no-arbitration term.⁵⁴ The employer argued in vain for a different result, of severing the unlawful elements in the arbitration clause and eliminating only those elements. The court rejected this, explaining that it has no vested power to “reform” a contractual term. In this case, the offensive element was the asymmetry—only the employee’s claims were directed to arbitration. The court explained that it would have to add a new, non bargained-for term (“both sides must arbitrate their claims”) instead of the existing, unreasonable term (“only the employee must arbitrate its claims”). The court found this affirmative “augmentation” to be beyond its authority, and thus chose to void the entire arbitration clause.⁵⁵ It thus rejected the application of the maximally tolerable regime.

This is a puzzling justification. Surely, the contract is “augmented” by a new term even under the court’s approach. Once the term is struck, the court must supply a non bargained-for gap-filler—here a no-arbitration clause. The *Armendariz* court voiced the concern that it is not for the court to write the contract over for the parties. But surely partial enforcement of the arbitration clause involves much less of a variation from the effects intended by the parties than total non-enforcement.⁵⁶ The question, then, is not whether the court has the power to reform the contract—it clearly does.⁵⁷ The question is how much of the bargain needs to be eliminated: only the minimum, rendering the remainder tolerable, or more than the minimum, rendering the remainder fair and balanced?

It is fair to propose that what drove the court’s decision in *Armendariz* was not a formalistic minimalism a-la the Blue Pencil rule. Rather, it was the drive to reform the contract in a way that is more than the minimum necessary, to attain a result that is more balanced and fair, rather than one that comports with bargaining power. It was, in other words, a masqueraded preference for the most-reasonable-term solution to the problem of repairing unconscionable contracts, possibly justified by the perceived deliberate bad-faith drafting on part of the employer. But despite the precedent that this case set for defining what constitutes unconscionability in arbitration clauses, its approach to severability was not generally followed. Indeed, in other cases, the same California Supreme Court

⁵⁴ *Armendariz v. Foundation Health Psychcare Services Inc.*, 6 P.3d 669 (Cal.2000)

⁵⁵ *Id.*, at 698 (“Because a court is unable to cure this unconscionability through severance or restriction, and is not permitted to cure it through reformation and augmentation, it must void the entire agreement”)

⁵⁶ Williston and Corbin, *supra* note 30, at 49-50.

⁵⁷ UCC 2-302, cmt 2. (“Under this section the court, in its discretion, may [...] limit unconscionable clauses so as to avoid unconscionable results.”)

invoked the severability principle and vacated only the offensive elements of the arbitration clause.⁵⁸

D. Liquidated Damages

Another application of the maximal tolerable terms idea concerns liquidated damages. It is well-known that courts do not enforce liquidated damage terms that are clearly excessive and punitive. But what is the damage term that courts supply instead? While the text-book answer is “compensatory” damages, it is often the case that compensatory damages can be assessed with more or less accuracy, thus lie within a fairly broad range of reasonableness, from the low estimates (that rule out consequential damages and types of avoidable harm) to high estimates (that include generous measures of potential lost profit). A maximal tolerable term would replace the unenforceable liquidated measure, not with the average or most reasonable compensatory measure, but rather with the high end estimate of expectation damages.

There are some statements in American case law that reject this notion. When an excessive liquidated damages clause is held unenforceable, it is wholly invalidated. In these situations, the most that courts are willing to award is the ex-post proven expectation damages. Courts refuse to apply a method of reducing the liquidated damages to bring them within the reasonable range.⁵⁹ Effectively, then, courts reject the maximally tolerable regime.

But other legal traditions deal differently with penalty clauses. Under Israeli contract law, for example, courts are instructed merely to reduce excessive damages to the level reflecting the magnitude of loss reasonably expected at the time of contracting.⁶⁰ In one case, a liquidated damages clause required the seller of a business to pay \$700 per day of delay in fulfilling his obligations. The seller was late by 100 days. The court found the liquidated amount to be excessive, and held that actual damages were probably zero or close to it, because the business was running at a loss. Still, the court decided to reduce the damages, not to the actual harm of \$0, but instead to \$200 per day, explaining that “\$200 per day is *the maximal amount that the parties could have anticipated* as possible harm from delay.”⁶¹ A leading commentary states that excessive liquidated damages should

⁵⁸ See, e.g., *Little v. Auto Stiegler*, 63 P.3d 979 (Cal. 2002).

⁵⁹ For an explicit rejection of the reduce-and-enforce methodology in penalty clauses, see *Cad Cam, Inc. v. Underwood*, 521 N.E.2d 498 (Oh. 1987)

⁶⁰ *Contract Law (Remedies for Breach of Contract) 1970*, Sec. 15(a), *Sefer Ha-Chukkim* No. 609. p. 13.

⁶¹ *Zaken v. Ziva*, Civil Appeal No. 539/92 (Unpublished), p. 4.

be reduced “to the *highest level* that the court regards as reasonably related the harm anticipated at the time of contracting...; that is, reduced to the measure *closest to the agreed sum, such that if that measure were the one agreed upon in the first place, the court would not have been justified in reducing it.*”⁶² This, in other words, is the maximally tolerable level.

IV. WHY ARE MAXIMALLY TOLERABLE TERMS USED?

The discussion so far had little to say about the normative justifications for maximally tolerable terms. It is time to address this issue, if only because it is now clear that this regime is not merely an intellectual curiosity, but rather a fairly prevalent solution in a variety of contexts. In light of the fact that the maximally tolerable term is often selected over the most reasonable term, one wonders what makes this solution so surprisingly prevalent. This section develops two insights. First, there is a good reason to repair contracts in a way that maintains bargaining advantages. Second, there is a built-in doctrinal limitation to the maximally tolerable approach, which takes care of the problem of incentives-to-overreach.

A. Maintaining the Bargaining Advantage

When bargaining power is unevenly distributed, the strong party will naturally seek ways to secure advantages in the contract by drafting one-sided terms. As long as he does not cross the boundary of reasonableness, contract law will respect and enforce the outcome of the bargain. Given this tolerance by the law to contracts that are (not unreasonably) one-sided, and in light of the various ways that the strong party can secure advantage, an aggressive intervention beyond the minimal necessary might backfire.

The legal standards that determine the boundaries of permissible terms and what constitutes unconscionability are not always obvious to the drafting party. For example, arbitration clauses in employment contract that were traditionally enforceable and still are in many jurisdictions are occasionally held to be unconscionable by some courts.⁶³ These terms set out many characteristics of the

⁶² U. Yadin, CONTRACT LAW (REMEDIES FOR BREACH OF CONTRACT) 1970, p. 132 (2d Ed. 1979) (in Hebrew). *See also* Eyal Zamir et. al, BRIEF COMMENTARY ON LAW RELATING TO PRIVATE LAW 302 (1996) (in Hebrew) (“the measure of reduction of liquidated damages ought to be to the level for which the element of excessiveness no longer applies...[such that] if that level was set in the first place, it would not have been reduced by the court.”)

⁶³ Compare *Szetela v. Discover Bank*, 118 Cal.Rptr.2d 862 (2002) (California court holding a credit card arbitration clause unconscionable because it forbids class actions) with *Hutcherson v.*

mandatory arbitration procedure, and there is a degree of uncertainty as to which would run afoul when challenged in court. Even when the law is clear, parties who draft contracts are not always informed about the legal standards. For example, if it is permissible to write a non-compete clause in an employment contract, is a 25-mile radius marking the non-compete territory excessive?⁶⁴ And even if parties are informed, it is not always clear whether the effect of a particular provision in the contract will be unconscionable. It depends on the subsequent circumstances.⁶⁵

If the drafting party was not fully aware of the threshold of unconscionability and, crossing this threshold was not done deliberately (the following section will discuss the deliberate case.) If the drafting party is punished for overreaching by, say, replacing the excessive term with one significantly below the maximal permissible level, she faces a familiar dilemma. Any additional sliver of the surplus she tries to appropriate through an incrementally more self-favorable term has a small upside equal to this increment if the term is enforced, but a downside of potentially significant magnitude if the term is held to be excessive and replaced with a mid-range term.

This asymmetry between the benefit and the risk of one-sided drafting could have two effects on the drafting party. First, it could lead her to draft more cautiously to assure that she does not bump against the maximal permissible boundary.⁶⁶ She will maintain a safety cushion against bumping into the ceiling by forgoing part of the surplus that she would otherwise extract. This effect alone is not a social cost, but it may be inefficient once we examine the effect of such extra caution on other terms of the deal. Second, the drafting party would have added incentive to invest in information that would enable her to assess the exact location of the boundary. Such investment has a private value because it can help the drafting party avoid the need for an overly cautious safety cushion, but it has a low (or zero) social value.⁶⁷

Sears Roebuck & Co., 793 N.E.2d 886 (Ill. 2003) (Illinois court holding same clause not unconscionable).

⁶⁴ See *Brecher v. Brown*, 17 N.W.2d 377 (Ia. 1945)

⁶⁵ For example, short term credit can have a very high interest rate that reflects not only the time-value of money, but also the “closing” fees. If the debt is to be paid off as scheduled, the high interest rate would be tolerable. But if it is not paid off in the short term, the interest rate can become unconscionable. See, e.g., *Carboni v. Arrospide*, 2 Cal.Rptr.2d 845 (1991).

⁶⁶ Craswell & Calfee, *Deterrence and Uncertain Legal Standards*, 2 *J. L., Econ. & Org.* 279 (1986).

⁶⁷ The argument that uncertainty over legal standards can lead parties to invest excessively in acquiring information has been developed in the literature on the negligence rule. See, e.g., Kaplow and Shavell, *Accuracy in the Assessment of Damages*, 39 *J.L. & Econ.* 191 (1996);

It is the first effect that requires attention. The incentive to moderate the drafting of a term to avoid bumping against the upper permissible limit may induce the drafting party to shift her bargaining leverage to other terms of the contract. In extreme cases, the party with the bargaining power might not enter the contract in the first place unless she can secure a very favorable term (e.g., a high interest payment in the face of high risk). In less extreme cases, she might force the other party to surrender to bad terms that are permissible, but no less costly to that party. For example, if the law were to apply the maximally tolerable terms, an employer who is interested in a mandatory arbitration clause or a non-compete restraint would experiment with drafting such terms, knowing that at worst, if a court views the term as excessive, it would be corrected only incrementally. But if the law replaces an excessive term with one that is unfavorable to the employer (say, no arbitration at all; or eliminating the non-compete restraint), the employer might be reluctant to draft such terms in the first place, and would use her bargaining power to insist on lower wages, bonds, or some other costly burden on the employee. Or, in rent-to-own cases, a seller can replace the option-to-own at a price that reflects unconscionable finance charges with a perpetual rental scheme that, while expensive, is less obscure and thus less likely to shock the conscience.⁶⁸ Generally, an aggressive legal intervention in one area of the contract can shift bargain dominance to other areas, affording no true relief to the weak party, *ex ante*.

Some parties might be better off by the “substitution” effect—those that have more to lose from the inclusion of the excessive term. For example, some employees benefit from access to court rather than arbitration, if they are likely to go to court. Some employers benefit from elimination of non-compete restraints, if they are likely to want to compete against their employer. But for many other employees, these benefits are less relevant, and the substitution effect on other terms of the contract (wage, benefits) is more relevant. From their perspective, the aggressive legal intervention backfires.

It is hard to weight the benefit to some against the disadvantage to others. But several observations lead me to be skeptical whether the more aggressive repair strategy would be overall beneficial. First, as long as the law does not resolve the underlying imbalance of bargaining power, the employer would still be able to dictate one sided terms. But he may be induced to do so in ways that take away some of the total surplus or which require a more costly drafting and transaction

Kaplow and Shavell, Private Versus Socially Optimal Provision of Ex Ante Legal Advice, 8 *J. L., Econ. & Org.* 306 (1992).

⁶⁸ *Murphy v. McNamara*, 416 A.2d 170 (Conn.1979).

procedure. Second, from an ex-ante perspective, even if some employees enjoy the added protection, it is funded to a significant extent by other employees who care less about the excessive term. The more aggressive the legal intervention in the excessive term, the greater this cross-subsidy.

This discussion helps explain why, in the current context, it is immaterial whether maximally tolerable terms are in some other normative sense inferior to the “most reasonable” terms. Terms that best satisfy the criterion of reasonableness are, by definition, more consistent with the norms of reasonableness than terms that just barely make it. The question, though, is not about the morally superior bargain, the fairest and most balanced one. Rather, it is about repairing or supplementing a contract where it is clear that one party has all the bargaining power. The argument for maximally tolerable terms is not that these terms are ideal. They are probably not. The reason why they may be desirable is that, *very much like standard gap fillers, they do not force upon the parties a costly circumvention*. Put differently, to fully address the underlying problem of uneven bargaining power and impose fair bargains the law needs to do more than the occasional tinkering with excessive terms. It is competition policy, not gap-filling rules, which can make a difference.⁶⁹

Granted, these are familiar concerns with the unconscionability doctrine—whether it can really help weak parties or merely backfire. My argument is that the policy choice courts face is not all-or-nothing—either provide full protection or no protection at all. Rather, in calibrating the degree of protection, courts deciding whether to apply the maximally tolerable regime are exercising a choice between minimal protection, such that raises less of the familiar “backfire” concerns, versus greater protection with its suggested costs.

B. Incentive to Overreach

If a party with the bargaining power who is drafting an excessive term expects that the court would only strike the excessive increment, what incentive does she have to avoid overreaching?⁷⁰ At best, the express term will stand; at worst, it will be replaced with the most favorable term permitted by law. Why draft a term that reflects this maximal permissible standard if you can get away with—in fact, benefit from—drafting a more excessive term? In contrast, if the law were to

⁶⁹ An example for ‘competition policy’ in the area of excessive terms is the Federal Price Gouging Prevention Act which punishes a whole class of retailers for raising prices unconscionably in times of emergency. Federal Price Gouging Prevention Act, H.R. 1252, 110th Cong. §2 (2007).

⁷⁰ See *Central Adjustment Bureau v. Ingram*, 678 S.W.2d 28, 37 (Tenn. 1984) (noting the adverse incentive problem)

replace the excessive term with a mid-range balanced term, or perhaps even with a *contra-proferentum* provision, the drafter has more to lose from overreaching and would have an incentive to draft less extreme terms.

This is a powerful objection, recognized by many commentators and courts.⁷¹ Ironically, it is almost too powerful to succeed under its own terms, namely, when it is invoked to support the most-reasonable term regime. A wrongdoer who deliberately drafts unreasonable terms can withstand even the more aggressive reformation of the contract in those few cases in which he is challenged, continuing his illicit business otherwise. In the few challenged cases he will lose more than the minimal increment, when the contract is reformed to the most reasonable term. But if there is a pattern of unconscionable behavior, this occasional loss will likely be dwarfed by the upside of the illicit gain. To deter such calculated violations, something more is needed, perhaps even stronger than the most-unfavorable term. More severe sanctions are appropriate, such as punitive damages or anti-fraud measures.

The key to reconciling this deterrence objection with the prevalence in practice of maximally tolerable terms is the distinction between deliberate and inadvertent overreaching. As argued above, the line between what is permissible and what would be considered intolerable is not clearly drawn and not always recognized by the drafting party. At times, the line is crossed knowingly and in bad faith, exploiting the high likelihood that such overreaching will go unchallenged. Often, though, the boundaries are not deliberately and maliciously crossed. Rather, the strong party is choosing one way to exploit her bargaining power—and could have easily chosen other ways that would have been deemed legitimate.

To account for this distinction between the type of offense, the scope of the maximally tolerable terms regime needs to be restricted. The regime would apply only when the strong party did not knowingly and deliberately overreach. If the boundary of permissible terms is known and nevertheless crossed, counting on the majority of parties to capitulate, the term should be replaced in a way that provides deterrence—something significantly less than the maximally tolerable term (and, I would argue, significantly less than the most reasonable terms as

⁷¹ White and Summers, *supra* note 3, at 234-35; Craswell, *supra* note 9, at 16-17; Walker v. Sheldon, 179 N.E.2d 497 (1961); Machtinger v. HOJ Industries Ltd., [1992] 1 S.C.R. 986 (“if the only sanction that employers potentially face for failure to comply with the minimum notice periods prescribed the Act is an order that they minimally comply with the Act, employers will have little incentive to make contracts with their employees that comply with the Act.”)

well.⁷²) A good example is the minimum wage law, which sets a clear and bright line between tolerable and intolerable wages. An employer that pays less than minimum wage cannot plead ignorance. Accordingly, the statute awards the aggrieved employee more than the minimally tolerable wages—it doubles the unpaid wages.⁷³ But if the boundary is fuzzy and was violated without bad faith, the law would only reduce the excessive term back to the boundary—to the maximally tolerable level. This regime would not lead to overly cautious drafting; and at the same time, it would give the drafting party something to lose in cases in which the unreasonable term was knowingly inserted.

This conditional application of the maximally tolerable terms regime is consistent with the observed practice.⁷⁴ When the drafter is a repeat transactor dealing in matters for which it is easy to know what the maximal standards are, a term that violates a standard is presumptively bad faith, replaced with the most reasonable term.⁷⁵ Or when the crossing of the boundary was egregious, demonstrating a deliberate disregard of the threshold, an aggressive intervention is again more justified as a deterrent.⁷⁶ With these safeguards in place, we can now better understand why maximally tolerable terms are quite so prevalent.

⁷² One possible solution is punitive damages; another is government enforcement. Both solutions are used in the case of deliberate price gouging. See Federal Price Gouging Prevention Act, H.R. 1252, 110th Cong. §3 (2007) (Enforcement by FTC with fines of up to \$3 million) and §4 (stiff criminal penalties – up to \$150 million and 10 years of imprisonment).

⁷³ Fair Labor Standard Act, 29 U.S.C.A §16(b).

⁷⁴ See FARNSWORTH, *CONTRACTS* 347 (4th Ed. 2004) (“absent a showing that [the excessive clause] was drafted in good faith..., the court may fix a lesser restraint that it would allowed the parties themselves to fix in their agreement.”)

⁷⁵ See *Machtinger v. HOJ Industries Ltd.*, [1992] 1 S.C.R 986, where it was presumed that an employer who drafted a term in violation of statutory minimum protection standards knew that this was a violation.

⁷⁶ See, e.g., *Central Adjustment Bureau v. Ingram*, 678 S.W.2d 28, 37 (Tenn.1984), where the court notes:

“We recognize the force of the objection that judicial modification could permit an employer to insert oppressive and unnecessary restrictions into a contract knowing that the courts can modify and enforce the covenant on reasonable terms. [...] the employer may have nothing to lose by going to court, thereby provoking needless litigation. If there is credible evidence to sustain a finding that a contract is deliberately unreasonable and oppressive, then the covenant is invalid.”

See also *Jenkins v. Jenkins Irrigation*, 259 S.E.2d 47, 51 (Ga. 1979). The same practice is followed in England. Compare *Mason v. Provident Clothing & Supply Co. Ltd.*, [1913] AC 724 (the entire non-compete clause is severed when drafted in deliberately unreasonable fashion) with *Goldson v. Goldman*, [1915] 1 Ch. 292 (with no evidence of deliberate overreaching, only the unreasonable increment is severed.)

CONCLUSION

The purpose of this article was not to advocate for the general use of maximal tolerable gap-fillers, but to identify it as a conceptual and practical possibility and discuss some arguments in support of such a regime. Upon first encounter, I imagine, readers were likely to be skeptical. If a court already goes into the trouble of reforming a contract, why not provide the most reasonable repair? And indeed, the argument in favor of the most-reasonable-terms is compelling and intuitive, quite easy to make. Yet despite this inclination, I set out in this article to explore an alternative. I found surprising pervasiveness in the use of maximally tolerable terms, in scattered areas of American contract law, as well as in comparative law. That this principle managed to permeate the law so broadly suggested to me that in some more subtle ways it can be justified, or at least grounded in broader existing practices, and it can trump even the hard-wired predisposition in favor of the most-reasonable-terms criterion.

In the end, I find the incentive problem to be the most troubling one. Would this regime of maximally tolerable terms encourage parties to draft excessive and unconscionable terms? If this concern is the crucial one, and I think it is, it marks the limits of this approach. If the boundary of permissible contracting is easily known and nevertheless crossed, the term should be replaced in a way that provides deterrence—something significantly less than the maximally tolerable term. In fact, true deterrence may require some kind of punitive response. But if the threshold is fuzzy and was violated without bad faith, the law would only reduce the excessive term back to the threshold—to the maximally tolerable level.

This principle of maximally tolerable terms is part of my more ambitious thesis, that bargaining power matters for contract doctrine. In a companion paper, I argue that bargaining power should affect the way courts supplement true gaps in contracts.⁷⁷ The gap-fillers should mimic the bargain that the parties would have struck, even if that bargain favors one of the parties. Maximally tolerable terms, in my view, are a species of bargain-mimicking terms. They are favorable to the party with the bargaining power, mimicking to the maximally permissible level the bargain of the parties.

⁷⁷ Ben-Shahar, *supra* note 1.