



Contrived Threats v. Uncontrived Warnings: A General Solution to the Puzzles of Contractual Duress, Unconstitutional Conditions, and Blackmail

Citation

Einer Elhauge, *Contrived Threats v. Uncontrived Warnings: A General Solution to the Puzzles of Contractual Duress, Unconstitutional Conditions, and Blackmail* (forthcoming 83 U. Chi. L. Rev. 2016).

Permanent link

<http://nrs.harvard.edu/urn-3:HUL.InstRepos:16073956>

Terms of Use

This article was downloaded from Harvard University's DASH repository, and is made available under the terms and conditions applicable to Open Access Policy Articles, as set forth at <http://nrs.harvard.edu/urn-3:HUL.InstRepos:dash.current.terms-of-use#OAP>

Share Your Story

The Harvard community has made this article openly available.
Please share how this access benefits you. [Submit a story](#).

[Accessibility](#)

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

HARVARD

JOHN M. OLIN CENTER FOR LAW, ECONOMICS, AND BUSINESS

CONTRIVED THREATS V. UNCONTRIVED WARNINGS: A GENERAL
SOLUTION TO THE PUZZLES OF CONTRACTUAL DURESS,
UNCONSTITUTIONAL CONDITIONS, AND BLACKMAIL

Einer Elhauge

Forthcoming in *University of Chicago Law Review*, Vol. 83, No. 2 (2016)

Discussion Paper No. 811

02/2015

Harvard Law School
Cambridge, MA 02138

This paper can be downloaded without charge from:

The Harvard John M. Olin Discussion Paper Series:
http://www.law.harvard.edu/programs/olin_center/

The Social Science Research Network Electronic Paper Collection:
<http://ssrn.com/abstract=2566053>

CONTRIVED THREATS v. UNCONTRIVED WARNINGS: A GENERAL SOLUTION TO THE PUZZLES OF CONTRACTUAL DURESS, UNCONSTITUTIONAL CONDITIONS, AND BLACKMAIL

Einer Elhauge*

83 U. CHICAGO L. REV. (forthcoming 2016)

April 3, 2015

Contractual duress, unconstitutional conditions, and blackmail have long been puzzling. The puzzle is why these doctrines sometimes condemn threatening lawful action to induce agreements, but sometimes do not. This article provides a general solution to this puzzle. Such threats are unlawfully coercive only when they are contrived, meaning the threatened action would not have occurred if no threat could be made. I show that such contrived threats can be credible because making the threat strongly influences whether the threatened action occurs. When such threats are uncontrived warnings, meaning the threatened action would have occurred even if no threat could be made, they are not coercive and can only benefit the agreeing parties. However, sometimes (as with blackmail) agreements produced by uncontrived warnings are also unlawful on the different grounds that they harm third parties. The contrived-threat test explains why the Medicaid defunding threat in Obamacare was unconstitutional and why (in the pending Supreme Court case) interpreting Obamacare as threatening to withhold tax credits from States that do not create insurance exchanges should either be rejected under the canon of avoidance or result in constitutional invalidation of that threat.

I. INTRODUCTION

The Supreme Court upheld the constitutionality of Obamacare with one notable exception: it struck down the provision that threatened to remove pre-existing federal Medicaid funding from states that refused to accept an expansion of Medicaid on the grounds that it ““crossed the line distinguishing encouragement

* Petrie Professor of Law, Harvard Law School. I am grateful for funding from the Petrie-Flom Center and Harvard Law School and for helpful comments from Michael Abramowicz, Jonathan Adler, Scott Altman, Ian Ayres, William Baude, Adam Cox, Elizabeth Emens, Richard Fallon, Joe Farrell, Charles Fried, John Goldberg, Wendy Gordon, Bert Huang, Rick Hills, Bert Huang, Daryl Levinson, John Manning, Eric Rasmussen, Chris Robertson, Mike Seidman, Steve Shavell, Sonja Starr, Matt Stephenson, Cass Sunstein, Mark Tushnet, Adrian Vermeule, Abe Wickelgren, and participants in the Harvard Law Faculty Workshop, the Harvard Law and Economics workshop, and the 2014 Yale Conference on Medicare and Medicaid.

from coercion.”¹ This constitutional standard has been condemned as amorphous and meaningless.² Because the standard explicitly incorporates contract law concepts,³ we might hope to find clear guidance in the underlying contract law. But contract law makes the legality of threats that induce contractual agreements turn on whether the threats are “improper” or in “bad faith”, conclusory labels that themselves have been deemed incoherent and meaningless.⁴ Nor could the Court find useful guidance from the general unconstitutional conditions doctrine because it has been deemed conclusory, incoherent, and “infamously inadequate.”⁵ Indeed, the Supreme Court seems to have given up even trying to define a coherent doctrine, openly admitting both in the Medicaid defunding threat case and a subsequent unconstitutional conditions case that it could not articulate the line dividing constitutional conditions from unconstitutional ones, but that it was nonetheless—somehow—confident that the conditions it was considering were over the line.⁶ This Article offers a coherent theory for solving these doctrinal puzzles that is normatively attractive and fits current legal results.

For analytical clarity, it is important to exclude some cases that lie outside these puzzles. If a threatened action by a government or private party were independently unconstitutional or illegal, there is no difficulty condemning it. Likewise, sometimes even an uncoerced agreement to a condition or contract would be independently unconstitutional or illegal, usually because it harms third parties or violates equal protection norms. In these cases, the independent doctrine of unconstitutionality or illegality is doing all the work.

Other times, whether the agreement was coerced is irrelevant because the threatener has a power to compel without any agreement, such as when the government has a power to order the relevant action because the individual or State has no constitutional right against such compulsion. Whenever direct coercion is permissible, that moots the issue of when threats of otherwise lawful action should be deemed coercive because they would be permissible either way. But in these

¹ *National Federation of Independent Business v. Sebelius*, 132 S. Ct. 2566, 2603, 2606 (2012) (Roberts, C.J., joined by Breyer & Kagan, JJ.) [hereinafter *NFIB*]; *id.* at 2661-62 (joint opinion of Scalia, Kennedy, Thomas, and Alito, JJ.).

² *Id.* at 2641 (Ginsburg, J., joined by Sotomayor, J.); Nicole Huberfeld et al., *Plunging into Endless Difficulties: Medicaid and Coercion in the Healthcare Case*, 93 B.U.L. REV. 1, 88 (2013).

³ 132 S. Ct. at 2602 (Roberts opinion); *id.* at 2659-60 (joint opinion).

⁴ Oren Bar-Gill & Omri Ben-Shahar, *Credible Coercion*, 83 TEX. L. REV. 717, 752-53, 779 (2005); Robert Hillman, *Policing Contract Modification Under the UCC: Good Faith and the Doctrine of Economic Duress*, 64 IOWA L. REV. 849, 862, 875-878 (1979).

⁵ Daryl J. Levinson, *Framing Transactions in Constitutional Law*, 111 YALE L.J. 1311, 1351 (2002); Kathleen M. Sullivan, *Unconstitutional Conditions*, 102 HARV. L. REV. 1413, 1419-1420, 1428-29 (1989).

⁶ *NFIB*, 132 S. Ct. at 2606-07 (Roberts opinion); *id.* at 2662 (joint opinion); *Agency for International Development (AID) v. Alliance for Open Society International, Inc.*, 133 S. Ct. 2321, 2330 (2013).

cases the work is done by the independent legal doctrines that create a power to coerce.

The cases of interest involve situations where a threat to engage in otherwise lawful action (like terminating funding) induces an agreement that is otherwise lawful, but the threat is nonetheless deemed too coercive to enforce the induced agreement. Those threats are the ones for which we could use some coherent principle to explain why they are sometimes being condemned and sometimes being allowed.

This article provides a simple principle for resolving this puzzle. It begins, in Part II, by resolving the underlying confusion in contract law about duress and contract modifications. This confusion can, I submit, be resolved with the following simple principle. A threat to engage in otherwise lawful action that induces contract modification is unlawfully coercive only when the threat is contrived, meaning the threatened action would not have occurred if no threat could be made.

The essence of the normative justification can easily be summarized. When a threat is not contrived, any legal prohibition that prevented the party from communicating the threat would, by definition, result in the threatened action. Moreover, because the threat, if made, would have induced the contract modification, both sides must, by definition, have thought that they were better off with the contract modification than with the threatened action. Deterring such threats would harm both parties. In such cases, communicating the planned action is less a threat than an uncontrived warning of what is coming, which gives the parties an opportunity to agree on an alternative that makes them both better off.

In contrast, when the threat is contrived, preventing the threat would not result in the threatened action and the parties would stick to the initial contract. Preventing such a contrived threat would clearly leave the threat recipient better off because she prefers the pre-existing contract. Ex post, preventing the contrived threat would leave the threatener worse off because he prefers the modification that the threat could have induced. But if the modification raises total value for both parties, a threat is unnecessary to achieve that modification because it could be achieved with a bonus that shares the joint gain. If, in contrast, the modification would reduce total value, then the prospect of such welfare-reducing threats would (if they were allowed) be priced into the contract and reduce reliance on the contract in a way that harms both parties. Moreover, sometimes contrived threats will not succeed in inducing contract modification and then be carried out, which leaves both parties worse off than if the pre-existing contract had continued. An ability to make such threats might also result in the other side making welfare-

reducing threats of its own. Thus, *ex ante* both parties would be better off preventing contrived threats.

The distinction between contrived threats and uncontrived warnings turns only on whether, in a no-threat world, the threatened action would have occurred. It does not turn on the wording of the communication. Nor does it require inquiry into the threatener's purpose. The threatener may have the benign purpose of making recipients of the threat better off, but the issue here presumes that recipients of threats begin with a legal autonomy right to make their own choices free of coercion by others, however well-intended that coercion may be. A threatening government may not even have a collective purpose, but that is unnecessary because what matters is what the government would have actually done without the threat. To be sure, the lack of any reason to take the threatened action in a no-threat world provides powerful evidence that the threatened action would not have occurred in that world. In that sense, purpose may sometimes be relevant to determining what would have happened, but the ultimate test is solely what would have happened without the threat, which need not require inquiry into purpose.

The contrived nature of a threat should also be distinguished from the credibility of that threat. A threat is credible if, *post* threat, the threatener would carry out the threatened action if the threat recipient does not agree. A threat is uncontrived if, in a *no*-threat world, the threatener would have taken that action anyway. The two differ because making a threat can strongly influence whether the threatened action occurs. As detailed below, a contrived threat to engage in unprofitable action can be credible when done in stages, as in the usual contracts case where performance is withheld day by day, because the victim's refusal to modify at each stage is itself not credible, so that the threat is expected to result in profitable modification. Making a contrived threat can also change the emotional calculus or create reputational effects that make persons carry out the threat even though they would not have taken the action in a no-threat world. Indeed, a credibility test has circularity problems where such reputational effects exist because whether a threat is credible can depend on whether the induced agreement is enforceable. Persons or governments may also make efforts to make their contrived threats more credible, such as incurring commitments, debts, or costs that make carrying out the threatened action more likely. When those efforts would not have been made absent an ability to communicate the threat, the credible threat remains contrived.

A contrived threat is thus not the same as a bluff because contrived threats are often credible. Nor are contrived threats harmless when they are bluffs. Bluffs can be very effective in coercing agreements when the victim is mistaken or even just unsure about the credibility of the threat. Indeed, if the bluff induced the victim to

accept an adverse modification, it must have been credible enough to the victim. Nor is harm avoided when the victim rejects modification. If a victim does so because he wrongly perceived the threat was not credible, the threat will be carried out, harming both parties. If a victim does so because she correctly perceived the threat was not credible, the threat will not be carried out but will still lead to unnecessary conflict and negotiation costs. It is thus always better to deter a contrived threat from being made (by eliminating any possible gain from it), regardless of how credible the threat might have been or seemed.

My distinction between contrived threats and uncontrived warnings is related to, but differs from, distinctions between threats and warnings that have been drawn by Robert Nozick, Thomas Schelling, and others.⁷ Whereas my definition turns on a pure but-for prediction, Nozick and other philosophers define threats in a way that combines moral and prediction baselines, which I show below not only muddles the results but perversely leads to more immoral conduct. Schelling and others use a prediction baseline, but a different one that is based on a pre-threat baseline rather than my no-threat baseline, and Schelling draws no normative conclusions about threats but rather focuses on analyzing their credibility.

Legally, an important benefit of using a contrived-threat test is that, as Part II shows, it explains many features and conclusions of contract law, whereas scholars who advocate a credibility test acknowledge that their test bears no relationship to existing contract law. Under contract law, modifications without mutual consideration are unenforceable if induced by a contrived threat, and modifications with mutual consideration are unenforceable if induced by a contrived threat to take action that would leave the victim significantly worse off.

I then show, in Part III, that the contrived-threat test also explains the Medicaid defunding case and provides a general solution to the problem of unconstitutional conditions. When a threat is contrived, the government benefit would have been provided in the but-for world without that condition, and thus the threat to withhold the benefit penalizes the exercise of a constitutional right. When the threat is uncontrived, the government benefit would not have been provided in that but-for world, and thus withholding the benefit imposes no penalty. Some constitutional law scholarship going back to Seth Kreimer has, following Nozick, included a similar prediction baseline but (like Nozick) combined it with other baselines in a way that I show confuse and worsen the results. My contrived-threat test also shows that (in the pending Supreme Court case) Obamacare should not be interpreted to deny federal tax credits to States that declined to create health

⁷ See *infra* at ___.

insurance exchanges because that would be a contrived threat and thus unconstitutional.

Finally, Part IV shows that the contrived-threat test also explains legal treatment of threats that induce private parties to enter into new contracts. It also solves the blackmail puzzle, showing that all blackmail contracts are necessarily produced by either (1) contrived threats, which creates inefficiency for the contracting parties without affecting information disclosure or (2) uncontrived warnings, which suppresses the disclosure of information to third parties whose interests often trump the interests of the contracting parties. The combination is what explains the full scope of blackmail law.

II. THREATS THAT INDUCE CONTRACT MODIFICATIONS

A. *Why Contract Standards Suggest a Contrived-Threat Test*

For threats of otherwise lawful action that induce contract modifications, contract law has both a special rule and a more general rule. The special rule applies only to modifications without mutual consideration, which increase the obligations of one party without providing any fresh consideration for that change. The more general duress rule applies to any threat that induces contractual agreement. Although the general legal standards for both are conclusory, the official guidance on them suggests a contrived-threat test.

1. The Special Rule for Modifications Without Mutual Consideration. A threat to breach a contract that induces a modification without mutual consideration is not binding unless the modification is (for a goods contract) in “good faith” or (for a non-goods contract) “fair and equitable in view of circumstances not anticipated by the parties when the contract was made.”⁸ Although terms like “good faith” and “fair and equitable” are conclusory, the law does provide guidance on the meaning of these terms.

For non-goods contracts, this guidance provides: “The limitation to a modification which is ‘fair and equitable’ ... requires an objectively demonstrable reason for seeking a modification.”⁹ An objectively demonstrable reason provides an independent reason to breach if no modification were possible, thus suggesting an

⁸ U.C.C §2-209 Comment 2 (1998); Restatement (Second) of Contracts §73, 89 (1981).

⁹ *Id.* §89 Comment b.

uncontrived warning. The lack of any such reason would instead suggest a likely contrived threat.

The unanticipated circumstances element is often confused with unforeseen circumstances, but that is not the test because the Restatement makes clear that such circumstances could be established by such foreseeable events as a person being offered a higher price or incurring a 17% cost increase.¹⁰ This is consistent with a contrived-threat test because such changes, though foreseeable, can often be the basis for an uncontrived warning if they make breach profitable despite contract remedies. The Restatement indicates that what it means by an unanticipated circumstance is one that “was not adequately covered” by the contract “even though it was foreseen as a remote possibility.”¹¹

“Unanticipated” thus means unaddressed by the contract. This factor is quite relevant to judging whether a threat is contrived because if any changed circumstances were already specifically addressed in the contract, then it seems unlikely that the change would have provoked a breach in a world where no modification were possible. For example, suppose costs have increased by 17%, but the contract specifically provides that if costs increase above 10%, the contract price will increase by the cost increase plus a specified markup. In that case, the parties have already determined what suffices to induce performance under those circumstances, so a claim that those circumstances would inevitably provoke breach seems contrived. Threatening to breach because of a circumstance specifically anticipated by the contract thus seems likely to be a contrived threat.

For goods contracts, the UCC comments flesh out the “good faith” test by stating:

the extortion of a “modification” without legitimate commercial reason is ineffective as a violation of the duty of good faith.... The test of “good faith” ... may in some situations require an objectively demonstrable reason for seeking a modification. But such matters as a market shift which makes performance come to involve a loss may provide such a reason even though there is no such unforeseen difficulty as would make out a legal excuse from performance...¹²

The reference to an “objectively demonstrable reason” again suggests an uncontrived warning, and the last sentence confirms that such a reason can exist even because of foreseen changes in circumstances.

¹⁰ *Id.* §89 Illustrations 3-4.

¹¹ *Id.* §89 Comment b.

¹² U.C.C §2-209 Comment 2.

2. The General Duress Rule. The general duress rule provides that any contract (including a new contract or modification with mutual consideration) is voidable when induced by (1) “an improper threat” that (2) “leaves the victim no reasonable alternative” but to agree.¹³

At one time, the no-reasonable-alternative element referred to conclusory, and somewhat mystical, notions of whether the victim’s “free will” was overborne. However, contract law has long rejected such notions and concluded that this element simply requires evidence that carrying out the threat would have made the victim significantly worse off.¹⁴

The key normative work is thus done by the improperness element, which standing alone is conclusory. But Restatement § 176 provides an extended definition of when a threat is improper. Some of the definition simply refers to threatened actions that are independently unlawful, which does not help with the current issue.¹⁵ But other parts of the definition say a threat is improper if it is in bad faith and lacks an objectively demonstrable reason, which the above shows suggests a contrived threat.¹⁶ Even more helpfully, the Restatement provides that: “A threat is improper if the resulting exchange is not on fair terms, and ... the threatened act would harm the recipient and would not significantly benefit the party making the threat.”¹⁷ If a threatened act would not significantly benefit the threatener, the party likely would not have taken the action if unable to make the threat, thus indicating a contrived threat.

B. How a Contrived-Threat Test Explains the Caselaw

1. Cases of Modifications Without Mutual Consideration. Consider first the renown case, *Alaska Packers’s Ass’n v. Dominico*.¹⁸ A group of fishermen agreed in San Francisco to take a boat to Alaska and fish in exchange for a company paying them \$50 for the season plus two cents for each salmon they caught. Once the fishermen were in Alaska, they threatened to stop work entirely and return to San Francisco unless their fixed pay was increased to \$100. Although the fishermen claimed they did so because their nets were defective, the court found

¹³ Restatement §175.

¹⁴ *Id.* §175, Comment b (rejecting test of whether “the threat must arouse such fear as precludes a party from exercising free will and judgment or that it must be such as would induce assent on the part of a brave man or a man of ordinary firmness . . . because of its vagueness and impracticability” and providing a series of examples that make clear that no reasonable alternative exists whenever the alternative would make the victim significantly worse off).

¹⁵ *Id.* §176(1)(a)-(c).

¹⁶ *Id.* §176(1)(d) & Comment e.

¹⁷ *Id.* §176(2).

¹⁸ 117 F. 99, 100 (9th Cir. 1902).

the nets were not actually defective and that the threat to stop work was “without any valid cause” and timed to take advantage of the fact that the company could not get replacement fishermen to such a remote location in time for the short fishing season.¹⁹ The company, which had also invested \$150,000 in a salmon cannery in Alaska, agreed to increase the fixed pay to \$100. The court held that, under these circumstances, this modification was unenforceable.²⁰

Given these findings,²¹ the fishermen’s threat was clearly contrived. The fishermen had no valid reason for threatening to stop work, and doing so would not only deprive them of any wages for the season, but also require them to fund their travel back to San Francisco and subject themselves to a risk of damages. There was thus no good reason to think that, if unable to make this threat, the fishermen would have stopped working. Thus, if contract law deters such threats from being made (by eliminating any gains from the modifications they induce), then the fishermen would have performed under the existing contract.

The lower court had upheld the modification because it concluded that the company’s agreement to it meant the company must have concluded the modification was better than suing for breach of contract, probably because the fishermen could not pay damages, and thus the modification was mutually beneficial.²² The problem with that lower court logic is that it wrongly assumes that, absent any communicated threat, the fishermen would have walked off the job. Instead, the findings indicate that they would not have done so, and that their threat to do so was thus contrived to extort the modification. A contract doctrine that deters such contrived threats from ever being communicated will thus result in performance of the original contract. A contract doctrine that instead made modifications produced by contrived threats enforceable would encourage such contrived threats, and thus clearly make the company worse off because it prefers original-contract performance to the modification.

¹⁹ *Id.* at 102. *See also id.* (stressing that the fishermen “willfully and arbitrarily broke” their contractual obligation and noting parallel to another case where the threat to breach was “unjustifiable”).

²⁰ Although some read this case to adopt a flat rule that modifications without mutual consideration are never enforceable, the court stressed that its holding that mutual consideration was absent applied only “under such circumstances” and depended on the conclusion that any mutual waiver of the original contract was not “voluntary” because the circumstances were like another case where the threatener “takes an unjustifiable advantage of the necessities of the other party” by threatening nonperformance and thus “coerces” the other party to increase the consideration it provides. *Id.*

²¹ Some have argued that the true facts were otherwise, *see Bar-Gill & Ben-Shahar, Credible Coercion, supra note* , at 757, but what matters for assessing the legal standard is the conclusion that the found facts made the modification unenforceable.

²² 117 F. at 102.

Moreover, a contract doctrine that made modifications produced by contrived threats enforceable would also be bad for fishermen *ex ante*. Switching to such a doctrine would encourage contrived threats and modifications that would make the fishermen's original promise to perform less valuable to the fish company, and thus make the fishing company unwilling to pay as much or to enter into the contract at all. To make itself less vulnerable to such contrived threats, the company might also be less willing to make investments in reliance on the contract (like \$150,000 for a cannery or chartering a boat), which is *ex ante* bad for the fishermen because reducing those investments reduce the productive value of their labor and thus what they can get paid.

In contrast, suppose the facts had been different. Suppose the fishermen threatened to stop work unless their pay was increased to \$100 because once they got to Alaska, another firm offered to pay them \$90. This would not be a contrived threat because, if unable to make this threat, the fishermen would walk off the job for better pay, which clearly leaves the fish company worse off than with the modification. It would instead be an uncontrived warning of what would inevitably happen if no modification were possible. Deterring the modification induced by this uncontrived warning would be harmful to both sides because without it the fishermen would get \$90 instead of \$100, and the fish company would be without fishermen.

But doesn't this uncontrived warning still diminish the willingness of the fish company to pay as much in the original contract or to make investments that rely on it? Actually, no. The uncontrived warning allows a modification that makes the fish company better off than it would be if fishermen walked off the job, which is what would happen if no warning were given. Thus, the possibility that uncontrived warnings would produce such modifications actually *increases* the value of the contract and encourages greater investment relying on it. To be sure, the prospect that the fishermen might independently want to walk off the job (in a way that a contract damages action would not fully compensate) *does* diminish the value of the contract and reliance on it. But if the warning is uncontrived, that prospect was going to arise regardless; the ability to modify the contract when it arises thus can only reduce the losses from that prospect.

Consistent with this conclusion, the Restatement indicates that a modification increasing an employee's pay for already-contracted performance would be binding if it came after an employee told his employer that he was abandoning the

contract because another firm had offered him more money.²³ The *Alaska Packers* court's own reasoning likewise suggests that its result would have been different if the threat to breach the contract were made for a "valid cause," which switching to a higher-paying firm would presumably supply. Indeed, *Alaska Packers* indicated no doubts about the enforceability of a prior modification without mutual consideration that increased the fixed pay of many of the fishermen from \$50 to \$60 *before* they left San Francisco.²⁴ The fishermen would not have been able to induce that change if the fish company could secure substitute fishermen for less than \$60, which suggests the market rate had risen to \$60 for these fishermen, and that they would thus have declined to get on the boat for Alaska if no modification were possible.

Likewise, if a party's contract-performance costs have increased above expected costs sufficiently to make performance unprofitable, then modifications that pay it more for the same performance have generally been deemed binding.²⁵ In such cases, threats not to perform are usually uncontrived warnings because the party would likely walk away from the unprofitable contract if unable to communicate any threat to do so. This is not always so because paying contract damages might be even more unprofitable than performance, but the cases finding the modifications enforceable generally indicate that the threatener was unwilling to perform at the contract price, perhaps because the threatener rationally found the performance cost worse than the risk of paying contract damages,²⁶ which would make the threat uncontrived.

2. Cases of Modifications With Mutual Consideration. Consider now the famous case of *Wolf v. Marlton Corp.*²⁷ After contracting to buy a house in a housing development, the buyers decided to divorce and wanted to escape the contract. They threatened that, unless the seller agreed to cancel the contract and return their security deposit, they would go ahead with the purchase and deliberately resell the house to an undesirable purchaser, which would reduce property values in the neighboring tracts owned by the seller. The court held that this threat of economic harm overcame the free will of the seller, but also held that was insufficient to

²³ Restatement §89, Illustration 3; *Schwartzreich v. Bauman-Basch, Inc.*, 231 N.Y. 196 (1921). Although a threat to leave for higher pay could be contrived if the pay increase were lower than expected contract damages, the employee had simply announced he was leaving for higher pay without demanding any modification, thus indicating that he would have left if no modification were possible. *Id.* at 199.

²⁴ 117 F. at 100.

²⁵ Restatement §89, Illustrations 1-2, 4.

²⁶ Restatement §89 Illustration 1; *Siebring Mfg. Co. v. Carlson Hybrid Corn Co.*, 246 Iowa 923, 926-927 (1955); *Brian Construction v. Brighenti*, 405 A.2d 72, 74 (Conn. 1978); *Goebel v. Linn*, 47 Mich. 489, 490 (1882); *Munroe v. Perkins*, 9 Pick. 305, 307 (Mass. 1830).

²⁷ 57 N.J. Super. 278 (1959).

constitute duress unless the threat was also “wrongful.”²⁸ The court acknowledged that the threatened action was perfectly lawful because the buyers had the legal right to resell to whomever they wished, regardless of the economic effect on the seller.²⁹ However, the court concluded it was nonetheless wrongful when “a party for purely malicious and unconscionable motives threatens to resell such a home to a purchaser, specially selected because he would be undesirable, for the sole purpose of injuring the builder’s business.”³⁰ The Restatement likewise makes clear that a threat to resell property to a purchaser that will have adverse effects on surrounding land is improper only if the threat is made “solely” to induce the contract modification.³¹

The court’s conclusion fits well with the theory that threats are wrongful when they are contrived. If the threat to resell to a purchaser specially selected to be undesirable was “purely malicious” and with the “sole purpose” of harming the seller’s business and inducing a modification, then there was no good reason to think that (if unable to communicate the threat) the buyers would have limited their resale options to undesirable purchasers. Limiting the set of potential purchasers would naturally tend to lower the expected resale price and thus harm the threatening buyers. Detering the threat would thus result in the original buyers making no special effort to seek out undesirable purchasers. The threat was thus contrived and, accordingly, wrongful.

The court’s reasoning indicates its holding would have been different if the threat were not made for the sole purpose of harming the seller, but was actually warning about coming action that would be independently beneficial for the threatening buyers. Suppose, for example, the buyers had told the seller they planned to resell to a registered child molester because he was willing to pay 10% over market price. Then the communication would be an uncontrived warning because, if unable to make this communication and negotiate a modification, the buyers would in fact resell to the child molester who was willing to pay the most. The seller would prefer to receive this warning because it enables the seller to negotiate a modification that makes the seller better off than it would be without the warning.

Another leading case is *Austin Instrument, Inc. v. Loral Corp.*³² Austin had a contract to supply gear parts to Loral to make radar sets for the Navy. When Loral got a second Navy contract for radar sets, Loral told Austin it would get a second

²⁸ *Id.* at 286.

²⁹ *Id.* at 287-288.

³⁰ *Id.* at 288.

³¹ Restatement §176, Illustration 10.

³² *Austin Instrument, Inc. v. Loral Corp.*, 29 N.Y.2d 124 (1971).

contract only on those gear parts for which it was lowest bidder. Austin immediately said it was opposed to supplying less than all gear parts on the second contract, and the next day Austin threatened to cease deliveries under the first contract unless Loral not only ordered all gear parts from Austin on the second contract but also gave Austin substantial price increases on gear parts under the first contract, whether or not those gear parts were already delivered.³³ Although the court focused on whether Loral had any reasonable alternative because that was the issue disputed on appeal, the court made clear that duress also required a showing that the threat was “wrongful.”³⁴

Here the threat was properly deemed wrongful because the threat to cease delivery was contrived.³⁵ True, the lower court noted that Austin “claimed” it was motivated by cost increases that made supplying gear parts unprofitable,³⁶ but the high court’s statement of facts omitted this claim in a way that implicitly rejected it. Further, such a cost-based motive seemed implausible given the facts that: (1) the timing indicated the threat was motivated by a desire to supply all gear parts on the second contract, rather than by cost increases; (2) cost increases that make supplying gear parts unprofitable is not a likely motive to insist on supplying more gear parts; and (3) a need to cover costs on future supplies cannot explain the demand for price increases on already-delivered gear parts. Thus, the facts suggested that Austin would not have naturally ceased delivery under the first contract, but threatened to do so only because making that threat would induce the desired contract modification.

When cost increases do make contract performance highly unprofitable, the Restatement concludes it is proper to threaten to breach unless the price is increased.³⁷ This is consistent with a contrived-threat test because such facts suggest the party would not perform even if there were no prospect of inducing a modification, so that the threat was an uncontrived warning.

³³ *Id.* at 129.

³⁴ *Id.* at 131.

³⁵ Dicta in the opinion suggests that any threat to breach a contract is wrongful, 29 N.Y.2d at 130-131, but the actual holding is limited to the court’s factual findings, and the Restatement makes it crystal clear that “A threat by a party to a contract not to perform his contractual duty is not, of itself, improper.” Restatement §176 Comment e.

³⁶ *Austin Instrument, Inc. v. Loral Corp.*, 316 N.Y.S.2d 528, 530 (N.Y. App. Div. 1970). The lower court did not confirm whether Austin’s costs had actually risen enough to make performance unprofitable, but simply noted that cost increases and renegotiations were common at the time. *Id.*

³⁷ Restatement §176, Illustration 8.

C. A General Proof for the Contrived-Threat Test

Call the value of a contract C_t for the threatener and C_r for the recipient of the threat. Call the value of a contract modification M_t for the threatener and M_r for the recipient. And call the value of the situation if the threatened action were taken T_t for the threatener and T_r for recipient.

We know $M_r > T_r$ or else the recipient would not have agreed to the modification. For modifications without mutual consideration, by definition $C_r > M_r$ because the modification worsens the deal for the recipient with zero offsetting benefit. Therefore, $C_r > M_r > T_r$. For modifications with mutual consideration, it may not be clear that $C_r > M_r$ at the time of modification.³⁸ But establishing the no-reasonable-alternative element of duress establishes that $C_r > T_r$. Such a threat would be unnecessary unless the recipient preferred the contract to the modification (if $M_r > C_r$, the recipient would accept the offered modification without any threat), so we can infer from such a threat that $C_r > M_r$. Alternatively, direct proof might exist that $C_r > M_r$, in which case we also know that $C_r > T_r$ because $M_r > T_r$. Either way, contract doctrine elements other than the “improperness” or “bad faith” of the threat establish that $C_r > M_r > T_r$.

We know the threatener prefers the modification both to the contract and to the threatened action, or else the threatener would not have sought and agreed to the modification, so $M_t > C_t$ and $M_t > T_t$. However, further ranking has two possibilities that differ depending on whether the threatener prefers the contract to taking the threatened action. One possibility is that $M_t > C_t > T_t$. In this case, taking the threatened action leaves the threatener worse off than it would be under the contract. The other possibility is that $M_t > T_t > C_t$. In this case, taking the threatened action would make the threatener better off than it would be under the contract.

1. Uncontrived Warnings. If $M_t > T_t > C_t$ and the law prevents the threat and thus the modification, then the threatened action will occur because it makes the threatener better off than sticking to the contract. Preventing the threat and modification would thus make both parties worse off because they would both prefer the modification to the situation with the threatened action (that is, $M_r > T_r$ and $M_t > T_t$). The threat is thus an uncontrived warning of what is inevitably coming unless the parties agree to a mutually beneficial alternative.

³⁸ Because the threat recipient opposes modification enforcement, $C_r > M_r$ at the time of litigation, but that could reflect changes in costs or market prices since the time of modification.

True, such an uncontrived threat will induce a modification that is less valuable to the recipient than the contract. Thus, relative to contract performance, the prospect of modifications might seem to reduce the contract's initial value to the recipient. But contract performance is the wrong baseline because without the uncontrived threat and modification, the victim would instead suffer the threatened action, the prospect of which would reduce the contract's initial value to the recipient *even more*. Accordingly, relative to the correct but-for baseline of the threatened action occurring, the prospect of allowing modifications when uncontrived warnings are made will raise the contract's initial value to the threat recipient, making the recipient willing to give contract terms more favorable to the threatener and to make more relationship-specific investments, both of which will also raise the contract's initial value to the threatener. Both parties are thus better off if such uncontrived warnings are allowed to induce contract modifications.

2. Contrived Threats. If $M_t > C_t > T_t$, then if the law prevents the threat and modification, the threatened action would not occur because taking the threatened action would leave the threatener worse off than under the contract. The original contract will thus be performed. Preventing such a contrived threat would thus clearly leave the recipient better off because $C_r > M_r$. That may suffice if the normative goal of duress/modification doctrine is simply to protect the recipient.

Further, *ex ante*, preventing contrived threats will also be preferable to the threatener, and thus both parties would benefit from a contract doctrine that prevents contrived threats. True, *ex post*, preventing such a contrived threat would leave the threatener worse off because $M_t > C_t$. But, *ex ante*, the prospect of contrived threats will be priced into the contract and affect the extent of reliance on that contract. To evaluate those effects, consider the following three possible consequences of a contrived threat.

a. Contrived Threats that Induce Modifications that Reduce Joint Welfare. One variation is that a contrived threat might induce a modification that lowers total combined value for the parties, which means $C_r + C_t > M_r + M_t$. In that case, the modification not only lowers joint welfare, but the prospect of such modifications will lower the contract's initial value to the recipient, leading the recipient to demand better terms and make lower relationship-specific investments, both of which lower the contract's initial value to the threatener. Further, the potential threatener would know that the recipient could make her own threats of this nature, which would leave the threatener worse off. In short, such welfare-reducing threats will be priced into the contract, deter reliance on the contract that benefits both parties, and be used against both sides, so that *ex ante* both parties would be better off preventing such threats.

b. Contrived Threats that Induce Modifications that Increase Joint Welfare. The other variation is that the modification increases total joint value because the modification's benefit to the threatener exceeds its harm to the recipient. In that case, $Mt - Ct > Cr - Mr$. But then the threatener does not need to make a threat to induce the modification. The threatener could instead obtain the same modification by offering a bonus B that exceeds the recipient's harm but is lower than the threatener's benefit from the modification, so that $Mt - Ct > B > Cr - Mr$. In short, if the modification increases joint welfare, it can be achieved without making any threat (of $Tr < Cr$) by instead offering a bonus that shares the joint gains between the two parties so that $Mr + B > Cr$. Thus, preventing contrived threats is never necessary for threats that enhance joint welfare.

c. Contrived Threats That Induce No Modification. The final possibility is that a contrived threat fails to induce a modification, perhaps because the recipient did not perceive it to be credible. If the recipient is mistaken in its perception, then the threat will be carried out, which makes both sides worse off because $Cr > Tr$ and $Ct > Tt$. If the recipient is correct in its perception, then the contract will continue, but the parties will have incurred unnecessary costs of conflict and negotiation.

d. Overall Effect of Banning All Contrived Threats. In short, preventing all contrived threats will prevent some modifications that reduce joint welfare, never prevent modifications that increase joint welfare, and prevent some contrived threats from being carried out to the detriment of both parties. Thus, banning all contrived threats leaves both parties better off *ex ante*, without any need to examine the welfare effects of the modification.

One might wonder whether we should also prevent uncontrived warnings because they, too, can never be necessary for modifications that enhance joint welfare relative to contract performance. The reason we should not is that although uncontrived warnings can indeed produce modifications that lower joint welfare relative to contract performance, without the uncontrived warnings we would not get contract performance. Instead, we would get the threatened action, and the induced modification does enhance welfare for both parties relative to the situation with that threatened action.

In short, to assess the effects of a threat/warning, we must compare those effects to a but-for baseline that reflects the effects that would exist in the world "but for" the making of that threat/warning. For a contrived threat, the effects are negative because the but-for baseline is contract performance, which gives higher joint value than any modification that requires such a threat. For an uncontrived

warning, the effects are positive because the but-for baseline is a threatened act that gives both parties lower value than the modification.

D. Other Legal Elements

1. No Reasonable Alternative. The above analysis establishes that whether a threat is legally “improper” should turn on whether it is contrived. For modifications with mutual consideration, duress doctrine also requires proving that the threat recipient had “no reasonable alternative,” which means that carrying out the threat would have made the recipient significantly worse off. In other words, the no-reasonable-alternative test requires evidence that Cr exceeds Tr by a significant enough amount that the threatener can use the threat to induce a modification that leaves the recipient significantly worse off than under the contract. This element thus establishes what I noted above was the requisite factual premise to show (in a mutual consideration case) that $Cr > Mr > Tr$, which was a necessary premise of the proof above, needed to show the situation involved a true threat rather than a bonus. Alternatively, one might have direct evidence that $Cr > Mr$, in which case can infer that $Cr > Tr$ because $Mr > Tr$.³⁹ Direct evidence that Cr is significantly greater than Mr should thus suffice to satisfy the no-reasonable-alternative test.

Administrative grounds explain why the law requires a significant difference even though a small difference could result in some small harm. In a world with costless, perfectly accurate adjudication, all contrived threats would be condemned even if the harm were small. But we do not live in such a world. A significance threshold assures the potential harm is large enough to exceed both litigation costs and the overdeterrence effects that can result from inevitable errors in adjudicative fact-finding. This reduces overdeterrence of uncontrived warnings that might be mistaken for contrived threats, but at a cost of underdetering contrived threats with small harm.

For modifications without mutual consideration, contract law does *not* require proving the no-reasonable-alternative test. This makes sense because, as the proof above showed, any modification without mutual consideration necessarily satisfies the condition that $Cr > Mr > Tr$. Because this automatically follows, litigation cost

³⁹ Some argue that the no-reasonable-alternative test is unnecessary because the recipient would never agree to the modification unless he had no reasonable alternative. Oren Bar-Gill & Omri Ben-Shahar, *The Law of Duress and the Economics of Credible Threats*, 33 J. LEGAL STUD. 391, 423 (2004). But that assumes the modification was adverse to the recipient when it was made, which is not necessarily true for modifications with mutual consideration.

or errors in establishing this condition are likely to be far lower than for modifications with mutual consideration.

If contract law remedies fully satisfied the expectation damages goal of putting the recipient in the same position as contract performance, a threat to breach a contract would mean $Cr = Tr$, and thus the necessary premise that $Cr > Tr$ could not be satisfied and a recipient would never agree to a modification where $Cr > Mr$. Accordingly, threats to breach a contract can cause the relevant problem only when contract remedies are undercompensatory, which alas is often the case given litigation costs and delays, uncertain adjudication, limited defendant assets, and doctrines that sometimes provide damages below expectations levels. For modifications without mutual consideration, we know that $Cr > Mr$, so we can infer that legal remedies must not have been fully compensatory or the recipient would not have agreed to the modification. For modifications with mutual consideration, courts instead often require evidence that the contract remedy is significantly undercompensatory.⁴⁰

One might think that the underlying problem of inadequate contract remedies indicates the solution should be to increase contract remedies rather than have doctrines that police modifications. However, even optimal contract remedies can only minimize the sum of underdeterrence and overdeterrence of breaches, thus making undercompensatory remedies inevitable in some cases. Nor will increasing contract remedies help in cases where the threatener lacks sufficient assets to pay the damages. In any event, to the extent that actual contract remedies leave $Cr > Tr$, it remains desirable to prevent modifications caused by contrived threats to breach a contract. The fact that different contract remedies might have deterred the threat to breach does not alter the desirability of voiding modifications induced by contrived threats that are made given actual contract remedies.

The fact that the underlying problem is inadequate contract remedies also explains why, if preventing contrived threats is good for both parties, their initial contracts do not simply prohibit contrived threats. The answer is that the same inadequate contract remedies would also fail to deter a breach of any promise not to make contrived threats. The only solution to that problem is to make modifications produced by contrived threats unenforceable, which is a solution that only the courts can provide because contract law provides that contractual prohibitions on future modifications are themselves unenforceable because the parties can always agree to modify those prohibitions.⁴¹

⁴⁰ *Austin*, 29 N.Y.2d at 130-131.

⁴¹ Restatement §311 Comment a.

2. Lack Objectively Demonstrable Reason or Significant Benefit. The legal standards indicate that a threat is improper and in bad faith unless there is an “objectively demonstrable reason” for seeking a modification and a “significant benefit” from the threatened action.⁴² While having a demonstrable reason or motive for the threatened action suggests the threat was uncontrived, one might wonder about the seeming requirement that the reason must be “objective” and produce some concrete benefit. Why shouldn’t a purely subjective reason, like anger or spite, count if that motive would have dictated the threatened action even if no threat had been made?

Suppose the *Alaska Packers* fishermen and *Marston* divorcing homebuyers were so angry about their situations that (even if no modification were possible) they would have taken the threatened action despite the fact that doing so would clearly harmed themselves financially. If so, refusing to enforce such modifications because those reasons are subjective would actually harm the recipient if we are sure both that (a) these are the facts and (b) the anger itself is not contrived.

What explains the rule requiring an objective reason or concrete benefit is that both those premises are untrue. First, claims about anger are subjective and easy to fake. Nor would it suffice to ascertain whether the anger was real; one would need to quantify its subjective effect to determine whether it would really outweigh the financial harm of taking the threatened action. This creates a proof problem not only in hard cases, but also systemically because the proof problem could be artificially generated in every case.

Second, even if anger would genuinely motivate the threatened action, anger might itself be endogenous to the desire to create a commitment or reputation that can make contrived threats more credible. Persons might work themselves into a lather because their anger makes it credible they will carry out their threats, but their anger might itself be contrived. Or they might want a reputation for acting out of anger when they are crossed so that their contrived threats will be more effective. Either way, anger itself can be contrived, and refusing to enforce modifications induced by threats that were produced by contrived anger might thus deter those threats from ever being made.

Thus, although in a world with costless, perfectly accurate adjudication, uncontrived warnings should be permissible even when based on nonpecuniary motives, in the actual world it is hard to detect anger that is faked or contrived, so admitting those motives which would lead to significant underdeterrence of contrived threats. To reduce this underdeterrence, contract law thus has sound

⁴² See *supra* at ____.

reasons to deem a threat improper (i.e., contrived) unless the threatened action is supported by some objective reason or concrete benefit other than inducing the modification, although the necessary cost is to overdeter some genuine emotional uncontrived warnings.

3. Unfair Modification Terms. Another doctrinal issue is why contract law makes the “fairness” of the induced modification relevant to whether it should be enforceable.⁴³ Given the proof above, one might think that any contrived threat necessarily induces harmful modification and any uncontrived warning induces beneficial modification, so that any separate inquiry into the fairness of the resulting modification seems unnecessary.

However, there is a sound reason for this requirement. Even though the refusal to perform a contract may be an uncontrived warning, in that the threatener would prefer nonperformance if *no* modification were possible, a threat not to perform unless a *particular* modification is granted may be contrived because a lesser modification would have sufficed to induce performance.⁴⁴ For example, a cost increase might make a party unwilling to perform at the original contract price, but a threat not to perform unless given a price increase that vastly exceeds that cost increase would still be a contrived threat because a smaller price increase would suffice. Thus, to limit the threatener to the zone of uncontrived warnings, the law must also police the “fairness” of the demanded modification. This explanation provides concrete content to what “fairness” means here—it means the modification should not exceed the minimum necessary for the threatener to be willing to perform the original contract.

For example, a Restatement illustration indicates that despite a demonstrable increase in the market price for metal needed to make a casting, it would be improper to threaten to withhold contractual supply of those castings unless the contract price were increased from 50 cents to 75 cents, in a case where rivals sold the same castings for 55 cents but with several months delay.⁴⁵ Given the demonstrable cost increase, the threat not to perform at the original contract price of 50 cents could well be uncontrived, and other Restatement provisions seem to treat such threats as proper.⁴⁶ But the problem here is that the threat was not to perform unless the price were increased to 75 cents, which does seem contrived because rivals were willing to supply the castings at 55 cents despite the same

⁴³ Restatement §89, 176.

⁴⁴ Consistent with this conclusion, Restatement provides: “Bad faith may be shown by proving that the person making the threat did not believe there was a reasonable basis for the threatened process ... *or* that he realized the demand he made was exorbitant.” *Id.* §176 Comment d (emphasis added).

⁴⁵ *Id.* §89 Illustration 5.

⁴⁶ *Id.* §89 Illustrations 1-2, 4.

increase in market costs for metal. Thus, although a threat not to perform unless the contract price were increased to 55 cents could be an uncontrived warning, a threat not to perform unless the contract price were increased to 75 cents was a contrived threat.

E. Why Not Instead Focus on Whether the Threat is Credible?

Professors Bar-Gill and Ben-Shahar argue that contract modifications should always be enforced whenever induced by credible threats or by non-credible threats that are accurately perceived, limiting potential nonenforceability to non-credible threats that are inaccurately perceived.⁴⁷ Other scholars reach a similar conclusion.⁴⁸ Such a credible-threat test differs from a contrived-threat test in important ways that make the latter more desirable.

Whether a threat is contrived turns on a *non-threat* baseline: if no threat could have been communicated, what would the threatener have done? If she would have taken the threatened action, then the threat is uncontrived; if she would not have, then the threat is contrived. Whether a threat is credible instead turns on a *post-threat* baseline: after the threat is communicated, would the threatener actually carry it out if the modification were rejected?⁴⁹ If the party would, then their threat is credible; if the party would not, then their threat is not credible.

This is a critical difference because the contrived-threat test is designed to prevent such threats from being made in the first place. A post-threat baseline instead assumes the threat would be made either way.⁵⁰ The tests reach different results in cases when, if no threat could be made, the threatener would not take the threatened action, but if a threat were made, the threatener would carry it out. In such cases, the threat is contrived but credible, and a credibility test would thus wrongly enforce any modification it induces. In contrast, my contrived-threat test would not enforce the induced modification because denying threateners any

⁴⁷ Bar-Gill & Ben-Shahar, *The Law of Duress*, *supra* note , at 392-94, 417; Bar-Gill & Ben-Shahar, *Credible Coercion*, *supra* note , at 717, 720, 727-728, 730-732, 744, 752.

⁴⁸ Daniel A. Graham & Ellen R. Peirce, *Contract Modification: an Economic Analysis of the Hold-Up Game*, 52 L. & CONTEMPORARY PROBS. 9, 11-12, 23 (1989).

⁴⁹ Bar-Gill & Ben-Shahar, *Credible Coercion*, *supra* note , at 722 (“Credibility is evaluated with an eye to the hypothetical temporal moment when the threat fails to induce the threatened party to surrender...”); *id.* at 742 (“Our analysis suggests that the correct baseline ... is not the position of the threatened party prior to the threat, but rather the position that she would be in if she were to reject the threat.”); *id.* 742 n.41 (stressing that they are using a “post-threat baseline”).

⁵⁰ *Id.* at 742 (observing that their “position takes the existence of a threat to be part of the unfortunate but relevant reality in which the dilemma has to be resolved.”)

benefits from making such threats would desirably deter such threats from ever being made.

1. Why Contrived Threats Can Be Credible. A contrived threat can be credible whenever the act of making the threat makes the party more likely to take the threatened action if no modification is given. Making a threat can have this effect for many reasons.

a. A Contrived Threat Can Be Credible When Staged. A contrived threat not to perform a contract is generally carried out in stages that make it credible because the victim's refusal to modify at each stage is itself not credible.⁵¹ Suppose, for example, that withholding performance is unprofitable standing alone, so that a one-time threat not to ever perform might not seem credible. Nonetheless, a threatener usually withholds contract performance only one day at a time. The threat to continue doing so each day is credible when a victim has no reasonable alternative because each day the victim's refusal to modify is not credible given that the victim is harmed each day he refuses to modify. Moreover, the threatener usually suffers harm only later if the victim sues or does not pay, possibilities that can be discounted because they would require the victim to follow through with a noncredible threat to refuse modification. Thus, continuing to withhold performance can, even though unprofitable standing alone, become profitable (and thus credible) because it is expected to result in a profitable modification. In a sense, the threat is credible each day because it creates a profitable reputation for credibility on following days.⁵²

Bar-Gill and Ben-Shahar acknowledge that such a staged breach would be credible even though each stage is costly to the threatener, but because they employ a credibility test they reach the mistaken conclusion that modifications induced by such breaches should be enforceable.⁵³ Their conclusion is mistaken because enforcing the modification encourages such contrived threats even though the prospect of them makes both parties worse off. A contrived-threat test instead correctly denies enforcement to the modification because no stage of the threatened

⁵¹ Lucian Bebchuk established a similar point for threats to sue. Lucian Bebchuk, *A New Theory Concerning the Credibility and Success of Threats to Sue*, 25 J. LEG. STUD. 1 (1996). Threats of contract nonperformance are even more credible than litigation threats because while in Bebchuk's model both sides incur equal costs at each litigation stage, in the contract situation usually the victims immediately incur large costs from nonperformance while the threatener incurs costs only later if the victim sues or does not pay.

⁵² The same theory works for more violent threats. The loan shark need not resort immediately to killing the nonpaying customer. Taking one finger off at a time can be more effective. Even though carrying out each threat to amputate a finger is unprofitable standing alone, it becomes profitable (and thus credible) because carrying out such a threat is likely to convince the victim that other fingers will follow and thus induce him to pay up.

⁵³ Bar-Gill & Ben-Shahar, *The Law of Duress*, *supra* note , at 417, 420-421 (concluding that any modification induced by a credible threat should be enforced and that threatening such a staged breach is credible).

action would have occurred without the prospect of such a modification, given that each stage of the threatened action is unprofitable but for its ability to induce modification.

b. A Contrived Threat Can Be Credible Because It Creates a Reputation For Carrying Out Future Threats. Once a threat is made, failing to carry it out can have reputational effects for future threats. It can thus be credible that a threatener will carry out a contrived threat (even though the threatened action is directly unprofitable) because the threatener profits from enhancing the credibility of their future threats.

Bar-Gill and Ben-Shahar acknowledge not only that this factor can make unprofitable threats credible, but also that it creates an endogeneity problem for their approach because it means that whether a threat is credible can depend on whether contract law enforces modifications induced by such threats.⁵⁴ Suppose carrying out a threat will reduce profits unless it enhances a reputation for carrying out similar threats will lead to future profitable modifications. If so, then the credibility of the threat turns entirely on whether contract law will enforce such modifications. If contract law does, then the threat is credible and the credibility test concludes that contract law correctly enforces such modifications. If contract law does not, then the threat is not credible and the credibility test concludes that contract law correctly refuses to enforce such modifications. Either conclusion is consistent with their test, which thus offers no basis for choosing among them. In other words, because their credibility test depends on whether contract law *does* enforce the threat, it cannot provide independent grounds for whether contract law *should* enforce the threat.

Although Bar-Gill and Ben-Shahar correctly conclude that analysis should take this endogeneity problem into account, their credibility test offers no answer to the question of which of the endogenous solutions one should choose. The answer instead depends on whether deterring the threat is desirable or not, which a credibility test does not answer because credible threats can be both contrived (and thus harmful) or uncontrived (and thus beneficial).

A contrived-threat test, in contrast, avoids any endogeneity problem. Contrived threats should make modifications unenforceable because in the no-threat world the threatened action would not occur. Thus, no reputation could be developed and profits from similar future modifications would be unavailable. The contrived-threat test accordingly always provides a clear conclusion that contract law should

⁵⁴ Bar-Gill & Ben-Shahar, *Credible Coercion*, *supra* note , at 725-726.

void modifications induced by such contrived threats, no matter how credible they may be.

c. A Contrived Threat Can Be Credible Because Making the Threat Creates Emotional Benefits To Carrying It Out. Given human nature, the mere act of making a threat can alter the psychic benefits of carrying it out. A person who fails to carry out a threat loses face and may feel obliged to live up to his threat. Indeed, a party may develop a character for carrying out her threats precisely because it makes her threats more credible in ways that are beneficial to her. Developing such a character can give parties psychic benefits from carrying out a threat that causes parties to do so even when there are no future reputational benefits.⁵⁵

Bar-Gill and Ben-Shahar acknowledge that emotional factors can make it credible to carry out an unprofitable threat.⁵⁶ However, they do not consider the possibility that making the threat can itself change the emotional calculus. A threat would be contrived if, absent making any threat, the psychic benefits would not suffice to take the unprofitable action. But making this contrived threat can increase the psychic benefits from taking that action in a way that offsets the lost profits and thus make the threat credible. Under the credibility test, this would make modifications induced by such threats enforceable, which would encourage such contrived threats because those modifications are profitable for the threatener. In contrast, under the contrived-threat test, the threatener has no incentive to make the threat in the first place given that any modification would be unenforceable. The credibility test is thus less desirable because the proof above shows that such contrived threats can only harm both parties.

d. A Contrived Threat Can Be Credible Because of Efforts to Improve the Threat's Credibility. Contrived threats can be made in ways that create formal commitments to carry them out and thus make them credible. Threateners might also take earlier steps that increase the credibility of their contrived threats. For example, if a threat to breach would be contrived because the prospect of damages makes breach unprofitable, the threatener can take on debt that reduces the expected damages. Alternatively, if a threat to breach would be contrived because performance would be profitable, a party could take steps to increase its costs (or fail avoid such increases) in order to make performance unprofitable and thus make a threat to breach credible.

⁵⁵ For example, a loan shark might find it advantageous to develop a character for severely punishing nonpayment, and that character may thus cause him to carry out a threat to kill someone who declines to pay, even though he hides the killing in a way that means it can have no reputational effects.

⁵⁶ *Id.* at 726-727, 756-757; Oren Bar-Gill & Omri Ben-Shahar, *Threatening an "Irrational" Breach of Contract*, 11 SUP. CT. ECON. REV. 143 (2004).

Such efforts may ultimately make a threat credible and thus make modifications induced by it enforceable under a credibility test. If those modifications are sufficiently profitable to the threatener, they can make such efforts worthwhile even though the efforts would be unprofitable if the party were unable to make the threat. Bar-Gill and Ben-Shahar acknowledge this effect is both possible and undesirable, and they recommend deviating from their credibility test when a party engages in efforts to make its threat more credible.⁵⁷ A contrived-threat test would instead directly lead to the desirable conclusion because, absent an ability to make the threat, such unprofitable efforts to make a threat more credible would not occur. Thus, without the ability to communicate such threats, neither those contrived efforts nor the threatened action would occur. The threat is thus contrived and modifications induced by it should be unenforceable.⁵⁸

e. A Contrived Threat Can Be Credible When There Are Low Odds of Great Harm. A contrived threat that is very unlikely to be carried out can still be credible enough to induce agreement given the magnitude of the threatened harm. For example, suppose someone puts a gun to your head unless you agree to modify a contract. Carrying out the threat harms the threatener because pulling the trigger gets them nothing, incurs legal risks, and costs a bullet. But suppose there are 1% odds the threatener will carry out the threat because of reputational effects or because he may be irrational, deranged, or on drugs. Even if you accurately perceive those low odds, you would likely consent to modification because a 1% risk of death exceeds the harm from the modification. The threat can thus be credible enough to induce modification even though (like here) it is clearly contrived. A contrived-threat test instead makes the modification void and thus eliminates any incentive to make the threat in the first place. Enforcing such modifications would instead encourage such threats, which clearly seems undesirable.

2. Why Contrived Threats Should Be Deterred Even If Not Credible. Bar-Gill and Ben-Shahar's theory also had problematic implications for non-credible threats. First, they conclude that modifications produced by non-credible threats should *also* be enforceable whenever victims accurately perceived the non-

⁵⁷ *Id.* at 732-734; Bar-Gill & Ben-Shahar, *The Law of Duress*, *supra* note , at 415-417; *see also* Mark Seidenfeld & Murat Mungan, *Duress as Rent-Seeking*, MINN. L. REV. (forthcoming 2015) (concluding that credible threats should be deemed wrongful when the costs of making the threat credible and carrying it out exceed the benefits of the threatened action).

⁵⁸ Consistent with my conclusion, courts have interpreted the good faith requirement of a "legitimate commercial reason" to mean a reason "outside the control of the party seeking modification." *T & S Brass & Bronze Works, Inc. v. Pic-Air, Inc.*, 790 F.2d 1098, 1105 (4th Cir. 1986). Thus, a party cannot bootstrap itself into a valid reason for breaching by contrived efforts to incur costs or debts or commitments that make performance unprofitable.

credibility of the threat.⁵⁹ They reason that in such cases the threat could not have induced the modification. Their approach thus limits the possible scope of duress doctrine to non-credible threats that the victim wrongly perceived to be credible.

Efforts to make contrived threats look credible are certainly harmful. For example, a threatener might claim its costs or debts are higher than they are, suggesting it is likely to breach if no modification is obtained, even though breach is in fact unprofitable.⁶⁰ Such claims are directly policed by contract doctrines that deem threats in bad faith when they are dishonest in fact.⁶¹ However, as the preceding section showed, a contrived threat can be credible even if the threatener is completely honest about its situation, so contract law correctly also deems threats to be bad faith whenever unsupported by objective reasons that would produce breach.

Nor is there any reason for the law get into complicated inquiries about whether a contrived threat was perceived to be credible. We are always better off if such a contrived threat were not made, rather than forcing the other side to guess about its credibility. If the contrived threat induces an adverse modification, it must have been perceived to be credible enough. As noted above, a threat can be credible enough even if the victim knows the odds are very low it will be carried out and even if fully carrying out the threat is clearly unprofitable but the threat can be made in steps. Courts inquiring about victim perceptions will often make errors, and may even suffer from hindsight bias because courts will have access to information that the victim did not have about the threat's credibility.

Moreover, even a contrived threat that fails to induce modification is harmful. If the contrived threat fails to induce modification because the victim wrongly perceived it was not credible, then the threat will be carried out, the prospect of which harms both parties *ex ante*. If the contrived threat fails to induce modification because the victim correctly perceived it was not credible, there is no still no benefit to allowing the threat, and indeed some harm because it leads to unnecessary conflict and negotiation costs. Because contrived threats never have any benefit and can create great harm, it is better for contract law to always deter them by eliminating any potential modification gains from making them.

Second, Bar-Gill and Ben-Shahar's approach produces no theory about which of the non-credible threats that are wrongly perceived to be credible should make

⁵⁹ Bar-Gill & Ben-Shahar, *The Law of Duress*, *supra* note , at 393-394, 402; Bar-Gill & Ben-Shahar, *Credible Coercion*, *supra* note , at 727-728, 730-731.

⁶⁰ Graham & Pierce, *supra* note , at 11, 23.

⁶¹ Roth Steel Products v. Sharon Steel Corp, 705 F.2d 134, 146 (6th Cir. 1983).

induced modifications voidable; instead, they conclude that this issue turns on normative considerations beyond their analysis.⁶² My analysis provides a clearer conclusion. All modifications induced by contrived threats should be void, whether or not that contrived threat was credible.

3. The Credibility Test Does Not Fit Contract Law. A final advantage of the contrived-threat test is that it fits well with existing contract law results, explaining why contrived threats are improper regardless of whether they will be carried out. In contrast, as Bar-Gill and Ben-Shahar acknowledge, their theory conflicts with current contract law.⁶³ Indeed, their insistence that credible threats should never void modifications directly conflicts with the no-reasonable-alternative element, because credible threats are the ones most likely to leave victims with no reasonable alternative. Their approach thus immunizes the worst kind of contrived threat: a credible one that will be carried out only because the threat was made. Current law also eschews any separate inquiry into whether a reasonable person would have perceived the threat to be credible, concluding (as I do) that it suffices if the threat actually induced the modification.⁶⁴

F. Comparing Other Proposed Tests

1. Nozick and Schelling on Threats v. Warning. My contrived-threat test is partly related to the distinction that Robert Nozick, Thomas Schelling, and others have drawn between impermissible “threats” and permissible “offers” and “warnings.”⁶⁵ Nozick’s is particularly important because it has influenced legal scholarship on contractual duress and unconstitutional conditions. However, their distinction and analyses differ from mine in various important respects.

First, while I use a pure prediction baseline to measure threats and warnings, Nozick and other philosophers combine a prediction baseline with a moral baseline, and thus condemn some “threats” that I would deem uncontrived warnings. Although Nozick defines an impermissible “threat” to exist when the threatened conduct would leave the recipient worse off than the “expected” course

⁶² Bar-Gill & Ben-Shahar, *Credible Coercion*, *supra* note , at 744 (“While we argue that whenever a threat is credible the deal should be enforced, we do not argue that whenever a threat is not credible, the deal should not be enforced... A normative theory ... is necessary to determine which among these noncredible threats are coercive.”); *id.* at 731-32.

⁶³ *Id.* at 719, 721, 737, 753-754; Bar-Gill & Ben-Shahar, *The Law of Duress*, *supra* note , at 392-394, 422-423.

⁶⁴ Restatement §175 Comment c (“the question is, did the threat actually induce assent on the part of the person claiming to be the victim of duress... it is not essential that a reasonable person would have believed that the maker of the threat had the ability to execute it...”).

⁶⁵ Robert Nozick, *Coercion*, in PHILOSOPHY, SCIENCE AND METHOD 440, 447-458 (Sidney Morgenbesser et al. eds., 1969); THOMAS SCHELLING, THE STRATEGY OF CONFLICT 123-24 n.5 (1960).

of events, he defines “expected” to “straddle *predicted and morally required*.”⁶⁶ Thus, his definition of an impermissible threat includes adverse deviations *not only* from what predictably would occur without any threat (which I would call a contrived threat) *but also* from what is “morally required” even though it conforms to what would predictably occur without the threat (thus including what I deem uncontrived warnings).

One problem with Nozick’s definition is that he never defines what is “morally required” (which he equates with “morally expected”), which adds a vague, conclusory element to his analysis.⁶⁷ Indeed, the typical contract modification is induced by threats to breach, which some might argue always deviates from what is morally expected,⁶⁸ which would mean such threats are always impermissible and the modifications are always voidable.⁶⁹ This would conflict with current law and be undesirable because the proof above shows that enforcing modifications induced by uncontrived “threats” is desirable.

No matter what he means by it, Nozick’s moral prong has bite only when it deems threats impermissible even when the threatened action conforms to the predicted course of events; that is, only when it condemns uncontrived “threats.”⁷⁰ The above proof shows this bite is undesirable. To be sure, such uncontrived threats by supposition threaten action that deviates from some (unspecified) moral norm, which makes them undesirable relative to compliance with that moral norm. But that moral deviation will by definition occur if the law prevents communication of uncontrived “threats” to take that deviant action. Thus, enforcing modifications induced by such uncontrived “threats” to engage in immoral action will not only reduce such immoral conduct but also leave both parties better off, given the morally unacceptable conduct that would occur if such modifications could not be enforced.

The fact that the threatened immoral conduct would occur may suggest that the underlying problem is that the law does not sufficiently penalize that conduct. But

⁶⁶ Nozick, *supra* note , at 447, 450 (emphasis in original). Other scholars likewise define threats by combining prediction and moral baselines. *See, e.g.*, ALAN WERTHEIMER, COERCION 206-221 (1988); Scott Altman, *Divorcing Threats and Offers*, 15 LAW & PHILOSOPHY 209 (1996).

⁶⁷ Nozick, *supra* note , at 450 (“In some such situations it will be unclear what P is morally expected to do, and hence unclear whether his statement is a threat or an offer.”).

⁶⁸ Seana Shiffrin, *The Divergence of Contract and Promise*, 120 HARV. L. REV. 708 (2007).

⁶⁹ If a threat adversely deviates from what is morally expected but not from what is predicted, Nozick condemns the threat whenever the victim would prefer what is morally expected to what is predicted. Nozick, *supra* note , at 451. Thus, if performance were morally expected, Nozick would always condemn threats to breach because victims prefer contract performance to breach.

⁷⁰ If a threat adversely deviates from what is predicted, Nozick condemns the threat even if it does not deviate from what is morally required. *Id.* at 453. Thus, the only effect his morally-required prong can have is to condemn some threats that the predicted prong would not.

the remedy for that is increasing that penalty. If penalties are not sufficiently large to deter the immoral conduct, banning uncontrived warnings about that conduct will merely assure that the immoral conduct occurs instead of preferable modifications. Moreover, there are usually good reasons why the law does not impose greater penalties on immoral conduct, such as the fear that imprecision in application of the law will also deter desirable conduct. Thus, even optimal legal regulation will leave undeterred some immoral conduct, the harm of which could be reduced by enforcing modifications induced by uncontrived warnings about that conduct.

Second, these scholars often use a different prediction baseline than I do. Although in parts Nozick uses the same no-threat baseline as me,⁷¹ in other parts he and other philosophers use a *pre*-threat baseline that turns on what the victim would have expected to occur before the threat was made.⁷² Given that just about every victim expects contract performance before they hear about a threat to breach, this pre-threat expectation test would condemn all contract modifications induced by a threat of breach. This would not only deviate from actual contract law, but would condemn desirable uncontrived warnings whenever the warning was unexpected by the victim. Deterring such unexpected warnings would thus increase unexpected harm to the victim and prevent mutually desirable modifications to avoid it. Schelling uses a different pre-threat baseline to define a “warning,” which turns on whether, immediately before making the threat, the threatener would have found it in its interests to take the threatened action.⁷³ His definition of a “warning” would thus include cases where parties make contrived commitments that give them independent incentives to take the threatened action, which I noted above should be condemned as contrived threats.⁷⁴ Schelling also draws no normative conclusions about what he calls threats because his work is instead about analyzing (quite brilliantly) the credibility and effectiveness of threats.

Third, Nozick concludes that contrived threats should be deemed permissible “offers” when they induce agreements that the recipient itself feels leaves him better off than the predicted course of events.⁷⁵ That is, even if a contrived threat results in $Cr > Tr$, he concludes it should count as a permissible offer if $Mr > Cr$.

⁷¹ *Id.* at 453-457.

⁷² *Id.* at 447-453; WERTHEIMER, *supra* note , at 207.

⁷³ SCHELLING, *supra* note , at 123-124 n.5.

⁷⁴ *See supra* at ____.

⁷⁵ Nozick, *supra* note , *Id.* at 448-449.

However, there is no benefit from creating a legal exception for contrived threats that induce modifications that leave the recipient better off. If it were really true that $Mr > Cr$, then the threatener could always induce the modification without any contrived threat. Indeed, the fact that the threatener felt a threat was needed suggests that she really believed $Cr > Mr$. The best way to test whether the recipient really believes $Mr > Cr$ would be to prevent contrived threats that leave the recipient worse off than Cr . This preserves the recipient's autonomy right to choose for himself whether he prefers the modification, rather than having that right turn on whether a court concludes the recipient really believes the modification made him worse off.

Fourth, and related to all the above points, Nozick believes that the distinction between impermissible threats and permissible offers or warnings is inescapably normative. Many other scholars have shared this view.⁷⁶ In my account, however, the distinction is purely empirical: it turns solely on what, in a no-threat world, the threatener would have done. This empirical distinction has, as proven above, clear normative implications given the consequences that flow from it. But the distinction is a factual one that does not require independent normative criteria. That is critical because we are dealing with threatened actions that are otherwise lawful and thus by definition within bounds where contract law and constitutional law allow parties and governments to pursue their own normative goals.

2. Changed Circumstances. Other leading scholars conclude that the enforceability of modifications should depend on the existence of various types of changed circumstances. Although such changed circumstances often correctly indicate an uncontrived warning that should lead to enforceability, their tests are under- and over-inclusive.

Alan Schwartz concludes that modifications should be enforceable when prompted by any changed circumstances.⁷⁷ However, many changed circumstances will not even make performance unprofitable, let alone so unprofitable that threateners would not perform even if modification were impossible. Unless it does so, threats not to perform are contrived threats despite changed circumstances, and enforcing such modifications under Schwartz's over-inclusive test will thus encourage contrived threats with undesirable consequences.

⁷⁶ CHARLES FRIED, CONTRACT AS PROMISE 95-99 (1981); Sullivan, *supra* note __, at 1446 n.133 (collecting sources).

⁷⁷ Alan Schwartz, *Relational Contracts and the Courts: An Analysis of Incomplete Agreements and Judicial Strategies*, 21 J. Legal Stud. 271, 312 (1992).

Richard Posner concludes modifications should be enforceable when changed circumstances make performance by a willing party impossible without a modification.⁷⁸ However, even if performance is possible, it will not actually occur if changed circumstances make performance less profitable than breach. In such cases, the threat to breach is an uncontrived warning and failing to enforce modifications under Posner's under-inclusive test will thus lead to breaches that both parties find worse than modification.

Aivazin, Trebilcock, and Penny conclude that modifications should be enforceable only when changed circumstances affect risks that are both remote and for which the threat recipients are superior risk bearers.⁷⁹ They reason that other modifications reallocate contractual risks away from superior risk bearers. However, even when their test is not met, threateners would (if modification is impossible) still breach when that is more profitable than performance. When that is so, threats to breach are uncontrived warnings and their underinclusive test will prevent desirable modifications. True, modifications induced by such contrived warnings may, as they argue, reallocate risks in a way that is less efficient than the original contracts. But if the modifications are prevented, we will not get performance of the original contracts. We will get breaches that both parties regard as worse than the modifications. Aivazin, Trebilcock, and Penny worry that the prospect of risk reallocation through modification will reduce the threateners' incentives to avoid risks that make them want to breach.⁸⁰ But the contrived-threat test eliminates that concern by deeming threats contrived if nonperformance became advantageous only because of changed circumstances that resulted from contrived decisions that the threatener would have found unprofitable but for the ability to make a threat that induces modification.⁸¹

Jason Johnston concludes modifications should be enforceable only when sought because changed circumstances made performance of the original contract unprofitable.⁸² Although such changes in circumstances often indicate an uncontrived warning, his test is both under- and over-inclusive.

First, even if changed circumstances make performance of the original contract unprofitable, sometimes the prospect of paying expectation damages is even less profitable. In such situations, threats not to perform are contrived despite those

⁷⁸ Richard A. Posner, *Gratuitous Promises in Economics and Law*, 6 J. LEGAL STUD. 411, 421-422 (1977).

⁷⁹ Varouj A. Aivazian, et al., *The Law of Contract Modification: The Uncertain Quest for a Bench Mark of Enforceability*, 22 OSGOODE HALL L.J. 173, 193-212 (1984).

⁸⁰ *Id.* at 194.

⁸¹ See *supra* at ____.

⁸² Jason Scott Johnston, *Default Rules/Mandatory Principles: A Game Theoretic Analysis of Good Faith and the Contract Modification Problem*, 3 S. CAL. INTERDISC. L.J. 335, 339-341, 366, 375 (1993).

changed circumstances, and enforcing modifications under Johnston's test will encourage such contrived threats, with undesirable consequences. Thus, courts are right to require further evidence that unprofitability really would have led to nonperformance even if no modification were possible.⁸³

Second, even if changed circumstances do not make performance of the original contract unprofitable, threats not to perform may be uncontrived warnings. For example, suppose performing the contract remains profitable, but a third party has offered a higher price that would be even more profitable. Suppose further that the threatener prefers to take the higher price (and risk contract damages) over performing the original contract. If so, threats not to perform are uncontrived warnings and refusing to enforce modifications will thus undesirably lead to breaches that make both parties worse off. Thus, contract law is right to enforce modifications when a party would otherwise take an even more profitable offer.⁸⁴

Third, even if changed circumstances make performance of the original contract unprofitable, that may not justify modifications induced by contrived threats unrelated to performance. For example, in *Wolf* the buyers' divorce was a changed circumstance that made performance unprofitable because now they had to resell.⁸⁵ But their threat was not to breach, but rather to complete the house purchase and then re-sell to an undesirable purchaser, which was even more unprofitable. That threat remains contrived and thus the court was correct not to enforce the modification.

III. UNCONSTITUTIONAL CONDITIONS

A. *Unconstitutional Conditions Affecting States*

The Medicaid defunding decision follows straightforwardly from the above contractual analysis. As the Court noted, its Spending Clause precedent incorporate contract law standards on duress, viewing the relationship between the federal government and the States as contractual in nature.⁸⁶ As Part I showed, the

⁸³ See *supra* at ____.

⁸⁴ See *supra* at ____ . Posner would go even further and enforce any modification if the threat to terminate was prompted by a third party offering a higher price. Posner, *supra* note , at 424. His position is overbroad because such threats are uncontrived warnings only if the higher price would have led to nonperformance (despite the risk of contract damages) even if no threat could be communicated.

⁸⁵ See *supra* at ____.

⁸⁶ See *supra* at ____ . A key premise for the contractual characterization was that the federal government could not directly command States to accept the Medicaid Expansion. *NFIB*, 132 S. Ct. at 2601 (Roberts opinion). If the

threat to defund Medicaid unless States accepted the Medicaid Expansion violated those contract law standards if (1) the threat was contrived (and thus “improper”) and (2) carrying out the threat would have left the States significantly worse off (thus meaning they had “no reasonable alternative” but to agree).

1. Findings Showing a Contrived Threat. Much of Chief Justice Roberts’ opinion focused on the second element because it was important to distinguish some key precedent. But before doing so, he offered an analysis that established the first element for the threat to take away pre-existing Medicaid funds, stating:

[1] The States claim that this threat serves no purpose other than to force unwilling States to sign up for the dramatic expansion in health care coverage effected by the Act. Given the nature of the threat and the programs at issue here, we must agree. [2] We have upheld Congress’s authority to condition the receipt of funds on the States’ complying with restrictions on the use of those funds, because that is the means by which Congress ensures that the funds are spent according to its view of the “general Welfare.” [3] Conditions that do not here govern the use of the funds, however, cannot be justified on that basis. When, for example, such conditions take the form of threats to terminate other significant independent grants, the conditions are properly viewed as a means of pressuring the States to accept policy changes.⁸⁷

Proposition [1] was a finding that the threat had “no purpose” but to force States to accept the modification.⁸⁸ This finding means the threat was contrived; i.e., that absent an ability to make the threat, Congress would not have terminated pre-existing Medicaid. Proposition [2] further indicates that threats are acceptable when they reflect Congress’s view that use of its funds advances the general welfare only if certain conditions are met. Such a view would indicate that Congress would find cutting off funding preferable to continuing it without conditions, which would make a threat to cut off funding an uncontrived warning of what Congress would do if attaching the condition were impossible. Proposition [3] concludes that this is not the case when Congress threatens to terminate significant independent grants, which Congress would have no reason to discontinue but for the fact that the threat pressures States to agree to a change, thus indicating a contrived threat.

federal government could, it would not matter whether the defunding threat was coercive because coercion would be permissible.

⁸⁷ *Id.* at 2603-04 (brackets added).

⁸⁸ For a similar conclusion, see Mitchell N. Berman, *Coercion, Compulsion, and the Medicaid Expansion: A Study in the Doctrine of Unconstitutional Conditions*, 91 TEX. L. REV. 1283, 1323, 1334-35 (2013).

The conclusion that the threat to eliminate pre-existing Medicaid funds was contrived seems entirely correct. There was no evidence that any significant faction in Congress (let alone a Congressional majority) preferred eliminating Medicaid over having pre-existing Medicaid continue. The threat to defund pre-existing Medicaid was thus a contrived threat to engage in action that Congress would not have engaged in if unable to make the threat. Indeed, now that the Supreme Court has prevented the Secretary of Health and Human Services from threatening to eliminate Medicaid to induce states to accept the Medicaid Expansion, we have not seen any Congressional movement to eliminate Medicaid, even though Congress has only gotten more conservative since the Congress that enacted Obamacare. To put it another way, Congress' clear preference ranking was Medicaid Expansion > Pre-Existing Medicaid > No Medicaid. If unable to impose Medicaid Expansion, then Congress would continue pre-existing Medicaid. Thus the threat to instead eliminate pre-existing Medicaid was a contrived threat that penalized State choices not to accept the Medicaid Expansion. Indeed, Robert explicitly confirmed he thought the threat was contrived when he concluded that (absent the condition) "practical constraints would plainly inhibit, if not preclude, the Federal Government from repealing the existing program."⁸⁹

Conversely, suppose we lived in a different world where Congress really thought that pre-existing Medicaid was working so poorly that the nation was better off eliminating it rather than allowing it to continue, but also thought that the Medicaid Expansion would be better than both options. That is, suppose its preference ranking were Medicaid Expansion > No Medicaid > Pre-Existing Medicaid. In that imaginary world, the threat to eliminate Medicaid would be an uncontrived warning and thus constitutional. This is desirable because if the law prevented Congress from communicating such an uncontrived warning, Congress would have eliminated Medicaid. That by definition would leave the States worse off, because they preferred agreeing to the Medicaid Expansion to having Medicaid eliminated.

The case also involved an uncontrived warning that the Court did sustain. Namely, the Court held Congress could threaten to withhold new Medicaid funds from States that would not agree to Medicaid Expansion.⁹⁰ That threat was an uncontrived warning because the only reason Congress was willing to offer the new funds was to get the Medicaid Expansion. Thus, if Congress could not condition those new funds on the States expanding Medicaid, then Congress would not provide the new funds at all. In other words, Congress' preference ranking was Medicaid Expansion > Pre-Existing Medicaid > Additional Funding Without

⁸⁹ 132 S. Ct. at 2606 n.14 (Roberts opinion).

⁹⁰ *Id.* at 2607.

Medicaid Expansion. If threats to induce Medicaid Expansion were impossible, this ranking means Congress would just provide States with pre-existing Medicaid funding, which was the uncontrived threat the Court allowed. Because some States preferred Medicaid Expansion with the new funds over Pre-Existing Medicaid, denying them the option of the former would just make those States (and Congress) worse off.

Roberts again confirmed this analysis because he concluded that while the threat to withhold *new* Medicaid funds was “offering” funds, the threat to deny *existing* Medicaid funds would “penalize” States.⁹¹ This distinction would not make sense if he was determining whether a threat was an offer or penalty relative to a baseline of what Congress legally *could* do because Congress could legally withhold either source of funds. But it does make sense if we understand Roberts to be using a baseline of what Congress *would* have done absent the condition.

Further confirmation was provided because Roberts’ analysis of *Steward Machine Co. v. Davis*⁹² concluded that the threat there was constitutional for reasons that showed it was an uncontrived warning. In *Steward*, Congress assessed a tax to fund federal unemployment insurance, but gave a tax credit to employers who paid into state unemployment plans. The Supreme Court held that the threat to withhold tax credits from employers in States without unemployment plans did not coerce States to adopt unemployment plans.⁹³ As Roberts correctly observed, that holding rested on the premise that “Congress was willing to direct businesses to instead pay the money into state programs only on the condition that the money be used for the same purposes.”⁹⁴ This premise meant that the threat to withhold tax credits was an uncontrived warning because, if Congress had no ability to obtain the condition, Congress would have provided no tax credit rather than give unconditioned tax credits.

To put it another way, in *Steward* Congress’ preference ranking was Conditioned Tax Credits > No Tax Credits > Unconditioned Tax Credits. Thus, the threat to deny tax credits was just an uncontrived warning of what would happen if tax credits could not be conditioned on states providing unemployment insurance. Prohibiting the conditioned tax credits would thus leave the States who accepted them worse off because their acceptance showed that they preferred the tax credits with conditions to no tax credits at all. In other words, offering the conditioned tax credits to States added an option they might prefer to the zero tax credits they

⁹¹ *Id.* at 2607.

⁹² 301 U.S. 548 (1937).

⁹³ *Id.* at 587.

⁹⁴ *NFIB*, 132 S. Ct. at 2603 (Roberts opinion); *Steward*, 301 U.S. at 587-591.

would otherwise get, rather than penalizing a choice not to adopt state unemployment insurance by withholding unconditioned tax credits that would otherwise be forthcoming. Accordingly, *Steward* correctly concluded that there Congress was using “temptation” (adding options preferable to the but-for world) rather than “coercion” (imposing penalties worse than the but-for world).⁹⁵

The joint opinion wrongly suggested in dicta that the sole test of coercion was whether the no-reasonable-choice test was met.⁹⁶ That test could condemn uncontrived warnings about denials of funding or tax credits, which would leave States worse off because the States would then suffer those denials even when they prefer to agree to the conditions necessary to avoid them. It would also deviate from the underlying contract law principles (which the joint opinion agreed applied) because those principles also require a contrived (i.e., improper) threat. However, this suggestion was dicta because other parts of the joint opinion indicate those justices believed the actual Congressional threat to eliminate pre-existing Medicaid was contrived, stating that “Congress surely” did not want to leave our “most vulnerable groups ... out in the cold” by eliminating Medicaid in States that did not accept the Expansion.⁹⁷ Because the actual threat was contrived, eliminating the ability to make it would not (and did not) result in an elimination of Medicaid that would harm the States. In any event, because the votes in the Robert’s opinion were necessary for the judgment, they limit the holding of unconstitutionality to threats that not only leave States with no reasonable choice but are also contrived.

2. The Lack of a Reasonable Alternative. The rest of Roberts’ analysis established the no-reasonable-alternative element, finding that States had “no real option but to acquiesce” because eliminating federal Medicaid funding threatened a “loss of over 10 percent of a State’s overall budget.”⁹⁸ This finding confirms that (as under contract law) the Spending Clause version of this element merely requires showing that carrying out the threat would have left the victim significantly worse off. Roberts distinguished *South Dakota v. Dole*⁹⁹ on the ground that there the threat was to eliminate funding that was less than 0.5% of the State’s budget.¹⁰⁰ The opinion thus suggests that losing less than 0.5% of a State’s budget does not leave it “significantly” worse off, but that losing 10% of a State’s budget does.

⁹⁵ 301 U.S. at 590-591.

⁹⁶ *NFIB*, 132 S. Ct. at 2661 (joint opinion).

⁹⁷ *Id.* at 2665.

⁹⁸ *Id.* at 2604-05.

⁹⁹ 483 U.S. 203 (1987).

¹⁰⁰ 132 S. Ct. at 2604 (Roberts opinion).

As noted above, where to draw the line on what counts as a significant difference turns on administrative grounds related to litigation costs and concerns that inevitable inaccuracies in adjudication (such as incorrectly finding a threat contrived) might create overdeterrence problems.¹⁰¹ It thus makes sense to limit the doctrine to cases with significant underdeterrence concerns, and if carrying out a threat would leave States less than 0.5% worse off, then modifications induced by that threat cannot harm States any more than that.

3. The Critiques. The Government argued that the threat was uncoercive because the Medicaid Expansion was so generous that it left the States better off.¹⁰² This parallels the Nozick argument that I rejected above: that a contrived threat should be deemed a permissible offer if it compels a party to accept an agreement that makes it better off.¹⁰³ As the joint opinion pointed, if the States really believed that the Expansion left them better off, then a threat to take away pre-existing Medicaid funds would have been unnecessary.¹⁰⁴ The fact that (with the threat now removed) many States have not accepted the Medicaid Expansion confirms they must not actually prefer the Expansion to pre-existing Medicaid. To be sure, many argue this choice is financially irrational given the huge share of federal funding for that expansion. But the legal premise for coercion analysis is that States have a constitutional autonomy right to choose for themselves whether to accept the Medicaid Expansion,¹⁰⁵ and in making autonomous choices, States are free to either assess the financial implications differently or make decisions based on reasons unrelated to finances. Sometimes we admire such decisions as standing on principle. To allow such contrived threats when the federal government and courts believe the end result makes the States better off, but the States do not, would violate the posited constitutional autonomy right.

Justice Ginsburg argued that the threat could not be coercive because Congress has the legal right to eliminate Medicaid funding.¹⁰⁶ But Spending Clause cases incorporate contract law principles, and her argument conflicts with the contract law principle that a threat to engage in otherwise lawful action can be coercive when it is contrived and leaves the victim no reasonable choice but to agree.¹⁰⁷ Her argument also conflicts with unconstitutional conditions doctrine, which invalidates many threats to take action (like cutting off funding) that is itself

¹⁰¹ See *supra* at ____.

¹⁰² 132 S. Ct. at 2665 (joint opinion).

¹⁰³ See *supra* at ____.

¹⁰⁴ 132 S. Ct. at 2665 (joint opinion).

¹⁰⁵ See *supra* note ____.

¹⁰⁶ 132 S. Ct. at 2630 (Ginsburg opinion).

¹⁰⁷ See *supra* at __; *infra* at __.

perfectly legal.¹⁰⁸ These doctrines are correct because no matter what Congress legally could do, a threat that deviates from what Congress would do without the challenged condition imposes a penalty on the State's constitutional choice.

Ginsburg also argued that the threat was not coercive because “Congress has not threatened to withhold funds earmarked for any other program.”¹⁰⁹ Roberts disputed her premise, concluding that pre-existing Medicaid and the Expansion were so different that they were not “all one program.”¹¹⁰ How can we tell whether programs are separate? The answer cannot turn, as Roberts correctly concluded, on how Congress “styled” the programs because the whole point is to impose some limit on Congress’ power.¹¹¹ The answer must instead depend on why it functionally matters whether programs are separate. The reason it matters is that if the threatened cutoff concerns an independent program, there is generally no reason to think that (absent linkage to the threat) Congress would have wanted to discontinue that independent program, which means the threat to do so is likely contrived. Thus, whether pre-existing Medicaid should count as an independent program from the Expansion turns not on formalisms, but on whether we think Congress would have preferred to continue pre-existing Medicaid in a world where it could not be linked to the Expansion. Because Congress clearly would have, the two should count as separate programs.

Ginsburg further claimed that the Court’s holding made no sense because Congress could instead have first repealed Medicaid and then offered a new Medicaid with the Expansion.¹¹² However, her premise again conflicts with contract law principles, which have long treated rescission and replacement as the same as modification.¹¹³ Nor is her premise accurate because repealing Medicaid would still be a contrived threat if it would not have occurred absent the prospect of inducing States to agree to the Expansion.¹¹⁴

Finally, Ginsburg argued that the threat was constitutional because Congress and the States had already agreed that Congress could amend Medicaid and the Expansion was a foreseeable use of that power.¹¹⁵ This argument is quite different because it claims that coercion does not matter because voluntary State consent to

¹⁰⁸ See *infra* at _.

¹⁰⁹ 132 S. Ct. at 2634 (Ginsburg opinion).

¹¹⁰ *Id.* at 2605 (Roberts opinion).

¹¹¹ *Id.*

¹¹² *Id.* at 2629, 2636 (Ginsburg opinion).

¹¹³ Restatement §89 Comment b.

¹¹⁴ After all, “a threat need not be verbally expressed; it may be perfectly clear from actions performed what the threat is.” Nozick, *supra* note , at 444.

¹¹⁵ 132 S. Ct. at 2630, 2638-39 (Ginsburg opinion).

the Expansion was unnecessary given that States had already agreed to it in their earlier Medicaid contracts. This claim raises a different contract issue: when a contract gives one side unilateral power to amend a contract, what are the limits on how that power can be exercised without the other side's consent? Roberts' opinion that such a unilateral power can be exercised to make only minor modifications is consistent with how contract law treats such provisions because without that implied limit the initial contract would contain illusory promises and be too indefinite to enforce.¹¹⁶ This explains why past minor Medicaid modifications were permissible even when unilaterally imposed by Congress.

4. The Pending Obamacare Tax Credit Case. In *King v. Burwell*,¹¹⁷ the challengers argue that Obamacare should be interpreted to deny federal tax credits to States that do not create their own health insurance exchanges. My own view is that this interpretation conflicts with most of the relevant text, as well as with the statutory structure, the statutory purpose, and the canon that conditions on federal grants to States must be unambiguous.¹¹⁸ But the challengers' interpretation should also be rejected because it interprets the statute as threatening to withhold those tax credits in order to coerce States to create insurance exchanges. This threat is clearly contrived because (if making this threat were impossible) the enacting Congress surely would have preferred giving tax credits without this condition over denying taxing credits. After all, as *NFIB* emphasized, Congress' overriding goal was universal insurance coverage.¹¹⁹ Obamacare imposes community rating and guaranteed issue on all States that would drive healthier persons out of individual insurance markets unless coupled with tax credits and mandates that keep them in.¹²⁰ Because Obamacare conditions mandates to buy insurance on the federal tax credit received, withholding tax credits would also lift the mandates for all employers and for 83% of individuals in States that refused to create exchanges.¹²¹ Withholding federal tax credits would thus cause a collapse of State insurance markets that is precisely opposite to the Congressional goal of universal coverage. Indeed, it would trigger a collapse of individual insurance markets below pre-Obamacare levels that would also be contrary to the aims of even Congressional opponents to Obamacare.

¹¹⁶ *Id.* at 2605-06 (Roberts opinion); Restatement §2 Comment e & §33; 1 Samuel Williston, *Contracts*, §§ 4:24 (4th Ed. 1990); 2 Arthur L. Corbin, *Corbin on Contracts* §§ 5.28 (2003).

¹¹⁷ 135 S.Ct. 475 (2014)(granting certiorari).

¹¹⁸ *NFIB*, 132 S. Ct. at 2605 (Roberts opinion); *id.* at 2659 (joint opinion). However, I do not think the interpretation conflicts with *Chevron* deference because such deference applies only when the agency action is a good proxy for current enactable preferences, and the IRS interpretation is clearly contrary to the views of the current Congress. See ELHAUGE, *STATUTORY DEFAULT RULES* 101-107 (2008).

¹¹⁹ 132 S. Ct. at 2606 (Roberts opinion); *id.* 2613 (Ginsburg opinion); *id.* at 2664-65, 2668-70 (joint opinion).

¹²⁰ Amicus Brief of Jewish Alliance for Law & Social Action at 14-15, 30.

¹²¹ *Id.* at 15-17, 32-33.

The challengers' interpretation would also meet the no-reasonable-choice element because carrying out the posited threat would leave States significantly worse off. The direct financial harm would alone be 6% of the average state budget,¹²² 12 times the 0.5% in *Dole* and closer to the 10% that *NFIB* found way past the constitutional line.¹²³ Moreover, this financial harm would be multiplied because the threatened action would also cause individual insurance markets to collapse in the States that refused to establish their own exchanges. Not only would this be harmful to their citizens and state economies (and thus to state tax revenue), but to deal with all the newly uninsured these States "would almost certainly find it necessary to increase its own health-care expenditures substantially" and indeed would be legally obligated to fund their emergency care in State hospitals.¹²⁴ The total harm to States would surely exceed the 10% that *NFIB* held was obviously unconstitutional.

Accordingly, the same principles that explain why the Court was correct to hold the Medicaid defunding threat unconstitutional show that interpreting Obamacare to threaten to withhold tax credits would also result in unconstitutionality. One might try to argue that the two cases differ because in *King* the threat pressures States indirectly by withholding new tax credits from its citizens, but that was precisely the form of threat at issue in *Steward*, which held that the applicable legal standard to judge such threats was the same legal standard that *NFIB* applied to the Medicaid defunding threat.¹²⁵ A holding that threats to deny new benefits cannot be coercive would also conflict with unconstitutional conditions doctrine, which often condemns withholding new benefits unless individuals waive their constitutional rights, and with underlying contractual principles, which often condemn threats to withhold new benefits to induce new contracts.¹²⁶ In any event, the threat in *King* does not merely withhold new benefits: given the associated regulation, it threatens a collapse of individual insurance markets below pre-Obamacare levels.

Interpreting Obamacare as threatening to withhold tax credits must thus be rejected under the canon requiring that statutes be interpreted to avoid constitutional invalidity. Indeed, as Justice Kennedy correctly stated at oral argument, the modern canon of constitutional avoidance requires avoiding that interpretation as long as it raises significant constitutional doubt or complexities, which would

¹²² *Id.* at 9.

¹²³ 132 S. Ct. at 2606-07 (Roberts opinion); *id.* at 2662 (joint opinion).

¹²⁴ *Id.* at 2657, 2664 (Joint Opinion).

¹²⁵ *Id.* at 2602-04 (Roberts opinion) (applying the *Steward* standard); *id.* at 2659 (joint opinion)(same).

¹²⁶ See *infra* at ___. The correct baseline is a predictive baseline, not a historical baseline, although history is sometimes relevant to prediction. *Id.* at ___.

certainly be the case.¹²⁷ Further, even if one (in my view, incorrectly) concluded that the statute threatened to withhold tax credits so clearly that the canon of avoidance was inapplicable, that threat must be invalidated as unconstitutional. Some have argued that the remedy for such unconstitutionality is unclear, but the routine remedy for unconstitutional conditions has always been voiding the condition, not denying the government benefit to everyone.¹²⁸ The contrived threat test explains why: such conditions are unconstitutional only when the threat was contrived, which by definition means that (if unable to impose the condition) the enacting legislature would have preferred granting the benefit over withholding it. This same legislative preference ranking means that statutory default rules on severability, which require severing the statute in a way that most likely to advance the enacting legislature's preferences, require voiding the threat to withhold benefits rather than voiding the grant of benefits for everyone.¹²⁹

B. Unconstitutional Conditions Affecting Private Persons

1. A General Proof. Unconstitutional conditions cases involve situations where the government has the power to deny a benefit and persons have a constitutional autonomy right not to engage in certain conduct, and the government conditions the benefit on the person not exercising that right. Sometimes courts declare such conditions unconstitutional, sometimes not. The essential critique is there is no coherent theory for explaining when courts should declare such conditions unconstitutional. But the following proves that the contrived-threat test does provide a coherent theory of which conditions violate the constitutional right not to engage in that conduct.

¹²⁷ Transcript at 16-18, 49; *NLRB v. Catholic Bishop*, 440 U.S. 490 (1979).

¹²⁸ That was true both in *NFIB* and all the cases considered next in Section III.B. When, for example, the Court holds it unconstitutional to withhold a tax exemption from veterans who refuse to give a loyalty oath, the remedy is to give the tax exemption to all veterans, not to deny it to all veterans. *Speiser v. Randall*, 357 U.S. 513, 531 (1958).

¹²⁹ See ELHAUGE, *STATUTORY DEFAULT RULES* 133-134 (2008). Justice Scalia suggested at oral argument that no prior case has ever held the Court can “rewrite” a statute when a clear provision is held unconstitutional. Transcript at 17. However, no such rewrite seems required: the condition “by the State” need simply be voided in the provisions that calculate tax credits based on premiums and coverage months in an “Exchange established by the State.” 26 U.S.C. § 36B. In any event, Justice Scalia is mistaken in his claim that no prior case has ever held a rewrite was the appropriate approach. In *United States v. Booker*, 543 U.S. 220, 246-249 (2005), which held mandatory sentencing guidelines without jury factfinding unconstitutional, the Court stressed that severability analysis required not simply grammatically separating the invalid terms, but rather choosing the option that Congress would have most preferred, and that this sometimes required “modification of a statutory provision,” which the Court did by making the guidelines advisory. Justice Scalia himself agreed that this decision required “rewriting” the statute. *Id.* at 284-285 (Stevens, J., dissenting, joined by Souter & Scalia, JJ). Moreover, Justice Scalia favored instead a different rewrite that would have added jury factfinding to the sentencing process on the grounds that Congress would have preferred that instead. *Id.* at 246 (pointing out that the dissent approach required the “addition” of provisions requiring jury factfinding); *id.* at 303-304 (Scalia, J).

All persons necessarily prefer unconditioned benefits over both conditioned benefits and no benefits. Persons who take conditioned benefits must also prefer them to no benefits, and thus for them unconditioned benefits > conditioned benefits > no benefits. Persons who reject conditioned benefits must find them worse than no benefits, and thus for them unconditioned benefits > no benefits > conditioned benefits.

The government threatens no benefits unless persons accept the conditioned benefits. There are two possibilities depending on whether, if conditioned benefits could not be offered, the government would prefer to provide the benefits or not.

a. *Uncontrived Warning.* If the government prefers conditioned benefits > no benefits > unconditioned benefits, then threats to provide no benefits are uncontrived warnings because if constitutional law does not allow the conditioned benefits, then the government will provide no benefits. Uncontrived warnings that benefits will be denied to those who reject the condition thus impose no penalty relative to the but-for world; they instead add a conditioned benefits option that some persons prefer to the but-for option of no benefit. To be sure, persons would be even happier with unconditioned benefits, but that is not what they will get if conditioned benefits are unconstitutional. Prohibiting conditioned benefits would be bad for both the government and persons who would accept the condition because they both must prefer conditioned benefits to no benefits. Prohibiting conditioned benefits would not affect persons who would reject the condition because they would get no benefits either way. Allowing conditioned benefits would thus make the government and some persons better off without making any persons worse off.

b. *Contrived Threat.* If the government instead prefers conditioned benefits > unconditioned benefits > no benefits, then threats to deny benefits to persons who do not accept conditioned benefits are contrived because if constitutional law does not allow the conditioned benefits, then the government would provide unconditioned benefits. Such contrived threats penalize decisions to reject the condition by denying unconditioned benefits that would have been given in the but-for world. Making a contrived threat thus has precisely the same effect as imposing a fine on those who exercise the constitutional right contrary to how the government wishes.

For persons who reject the condition, the contrived threat means they will get no benefits, rather than the unconditioned benefits they would have received without the threat. This effect makes these persons and the government worse off because they both prefer unconditioned benefits to no benefits.

For persons who accept the condition, the contrived threat means they will get conditioned benefits rather than unconditioned benefits. This makes these persons worse off because they prefer unconditioned benefits. But it makes the government better off because it prefers conditioned benefits to the unconditioned benefits it otherwise would have given. The benefits are the same either way. The only difference is that with the threat the constitutional right is exercised how the government wishes, whereas without the threat the constitutional right would be exercised how the person wishes. But the existence of the constitutional autonomy right by definition means that the Constitution has normatively concluded that the right should be exercised how the person prefers, unless the person receives benefits they are willing to exchange for those rights. Adding a contrived threat converts the situation so that, without any difference in benefit received, the right is now exercised how the government prefers instead. The government has basically taken over the autonomy right without providing any offsetting benefit. The effect is identical to the effect that imposing a fine has on those persons who comply and thus do not pay the fine. Given the premise of the constitutional right, the harm to these persons by definition exceeds the gain to the government.

In other words, contrived threats force persons who accept the condition to give up their constitutional right in exchange for nothing because they get the same benefit they would have gotten in the but-for world if conditioned benefits were not possible. That necessarily violates the posited constitutional right. Indeed, the contrived threat harms all persons whether or not they accept the condition. Because the government conditioned the benefits, it must find the policy gain from individuals who accept the condition greater than the policy harm on individuals who reject the condition, but that policy gain is unconstitutional by premise and thus illegitimate.

The above argument depends on the existence of a constitutional autonomy right, like the State right not to administer federal programs that was at issue in *NFIB* or the individual constitutional rights at issue in unconstitutional conditions caselaw. The same analysis would not apply if there is no constitutional autonomy right, but rather only a claim that the federal government is exceeding its Commerce Clause power. For example, suppose Congress attaches to federal spending or tax exemptions a contrived threat to deny that benefit to individuals that do X, that individuals have no constitutional right not to be compelled to do X, but that compelling X lies beyond Congress's Commerce Clause power. In that case, the contrived nature of the threat means it is coercive – i.e., regulatory – but there is no constitutional autonomy right to be free of such regulation. In these cases, the question is instead whether regulatory uses of the federal spending and tax powers should be limited to the scope of the Commerce Clause power. In a system of

multiple overlapping enumerated powers, it is not at all clear why it should be. Indeed, *NFIB* itself held the contrary, sustaining a regulatory use of the taxing power to penalize the nonpurchase of health insurance even though it struck down a mandate to buy health insurance as beyond the Commerce Clause.¹³⁰ Consistent with this limitation, courts have not struck down even contrived threats to deny benefits when the affected conduct is not protected by any constitutional autonomy right but simply lies beyond the scope of the Commerce Clause power.

c. Conclusion. Accordingly, conditioning benefits on the waiver of constitutional rights violates those rights whenever the threat to deny benefits is contrived. Such conditions do not violate those rights whenever the threat to deny benefits is uncontrived. Conditioning benefits reflects a contrived threat only when, if unable to make such threat, the government would still provide the benefits.

To put it another way, the contrived-threat test asks: in a but-for world where the condition could not be adopted, what would the government have done? If the government would have offered the benefit in the but-for world, then denying the benefit imposes a penalty relative to that but-for world, and thus the condition unconstitutionally penalizes the exercise of a constitutional right. If the government would not have offered the benefit in the but-for world, then denying the benefit imposes no such penalty and thus does not violate the recipient's constitutional right.

This analysis shows the error in the general claim that the greater power includes the lesser. When threats are contrived, the greater power to withhold benefits entirely would not be exercised if the supposedly lesser power to condition benefits did not exist. In such cases, the power to condition benefits is not lesser at all because it enables the government to achieve effects it would not achieve with the supposedly greater power.

Nothing in the above proof requires relying on a contract analogy or on whether benefits were provided in the past. Indeed, as we shall see, unconstitutional conditions doctrine fully applies to contrived threats to deny new benefits. The relevant baseline entitlements are set by the Constitution, not by the past provision of benefits. However, one could properly deem those constitutional rights to be part of the initial contract between the Government and its States and citizens. Further, regardless of whether one characterizes the initial set of constitutional entitlements as contractual, courts rely on contract principles to explain why that initial set of constitutional entitlements can ever be waived (i.e., modified) in exchange for government benefits. It thus makes sense to conclude that the same

¹³⁰ 132 S. Ct. at 2594-2600.

contract principles that allow some such modifications also limit others, so that the limits on contrived threats that induce agreements to modify initial contractual rights also apply to contrived threats that induce agreements to modify initial constitutional rights. Accordingly, while not required for the proof, the contract analogy provides legal grounds for the same conclusions as that proof.

To illustrate the above points, consider the classic case of *Frost & Frost Trucking Co. v. Railroad Comm'n*.¹³¹ California could legally ban truckers from its highways but truckers had (at the time) a substantive due process right not to be ordered to be common carriers. The Supreme Court held it was unconstitutional for a California statute to condition highway access on truckers agreeing to be common carriers. This holding has been critiqued because the truckers preferred being common carriers to lacking highway access.¹³² But this critique wrongly assumes that, if unable to condition the benefit, California would deny highway access to truckers. That seems implausible because trucking is highly valuable to the state economy. California's preference ranking was surely instead Common Carrier Condition > Unconditioned Trucker Access > No Trucker Access. Therefore, if California could not impose the common carrier condition, California would instead give truckers highway access without that condition. The threat to deny truckers access was thus a contrived threat that harmed truckers because they preferred unconditioned access to conditioned access. The Court thus correctly held the statute unconstitutional.¹³³

The threat in *Frost* was also entirely credible because the statute committed to carry out the threat and enough truckers complied that legislators had no incentive to repeal the statute. This is generally the case with unconstitutional conditions, and illustrates why contrived threats can be highly credible and why immunizing credible threats would produce the wrong result.

d. Avoiding Analytical Confusion. An important analytical benefit of the above analysis is that it carefully separates the issue of whether withholding benefits coercively violates a constitutional autonomy right not to engage in certain conduct from claims that (1) the threatened withholding is illegal, (2) the induced agreement is independently illegal even if uncoerced, or (3) the induced agreement

¹³¹ 271 U.S. 583 (1926).

¹³² Adam Cox & Adam Samaha, *Unconstitutional Conditions Questions Everywhere: The Implications of Exit and Sorting for Constitutional Law and Theory*, 5 J. LEGAL ANALYSIS 61, 65 (2013)

¹³³ Likewise, it is unconstitutional to condition the ability of corporations to do business in a State on waiving their access to federal court. *Terral v. Burke Construction Co.*, 257 U.S. 529 (1922). This holding is correct because, if unable to adopt that condition, a State would surely prefer having corporations operate in the State without the condition over having no corporations operate in the State. Thus the threat to withhold the right to operate was contrived.

is legal even if coerced because there is no relevant constitutional right to be free of government coercion. Past works have found unconstitutional conditions intractable in part because they often mixed up these analytically separate issues.

For example, suppose the government provides a \$100 prize to the person who writes the best essay arguing that the next President should be Hillary Clinton. The threat to withhold \$100 seems uncontrived because the government would not provide the \$100 if it could not attach the condition. Such a prize certainly would (and should) be unconstitutional. Does that mean the unconstitutional conditions doctrine does properly condemn uncontrived threats? No, because the reason this prize is unconstitutional is *not* that it coerces the prize recipient into giving up her constitutional right to free speech. Instead, it is because, even without coercion, such agreements violate equal protection because they fund speech favoring one political party over another, and thus harm third parties even if they benefit the agreeing parties. To test this, suppose the government instead equally funds political speech for both parties, as it has done with Presidential campaigns. Then there is no constitutional violation because the funding is equal. For other examples, whether an uncoerced agreement is independently unconstitutional may be more difficult, but analytical clarity comes from seeing that such conclusions must rest on grounds other than coercion. Just as some private contracts are illegal as a matter of public policy even if not produced by duress, likewise some government-person agreements are independently unconstitutional even if not produced by coercion. Analytical clarity requires separating the issues.

e. Difficulties of Proof. One might be concerned that it is hard to determine whether a threat is contrived. However, often, as in *Frost* and the Obamacare cases, it is not that hard, and this has not proven an insuperable problem for contract law, which has to resolve the same issue. Even when it is hard, a difficult empirical question beats a vacuous standard. Difficulties of proof are also a common issue in law. The possibility of adjudicative error creates overdeterrence and underdeterrence, and we have our usual legal tools for dealing with it. Here, because determining whether a threat is contrived will produce litigation costs and adjudicative error that deters uncontrived warnings to some extent, it makes sense to limit this doctrine (like the contract doctrine) to cases where underdeterrence concerns are significant by also requiring evidence that the carrying out the threat would have left the victim significantly worse off.

No matter how difficult it may be for courts to figure out the true facts, we cannot avoid the reality is that some threats are actually contrived and some threats are actually uncontrived, and that the above proof shows the effects for agreeing persons are always harmful for the contrived threats and beneficial for the

uncontrived threats. This is important because it means that any legal test one uses must be judged by the extent to which it does successfully deter contrived threats while avoiding deterring uncontrived warnings. As we shall see below, the legal tests used by the Supreme Court do fairly well on this score.

2. The Fit With Caselaw. Supreme Court standards on unconstitutional conditions correlate, albeit imperfectly, with a contrived-threat test. The application of those standards can be improved by clarifying that in ambiguous cases they should be interpreted to implement a contrived-threat test.

a. Threats to Deny Funds Used for Government-Desired Objectives to Those Who Exercise Constitutional Rights Without Using Government Funds. A series of cases draws the following distinction. Congress can deny funds to persons that use federal funds for activities Congress does not want to subsidize. But Congress cannot deny funds to persons that use federal funds only for “activities Congress wants to subsidize” based on how those persons exercise constitutional rights that are not federally funded.¹³⁴ Although the Court itself admits this “line is hardly clear” or “self-evident,”¹³⁵ a contrived-threat test helps define the line.

Denying funds used for activities Congress does not want to subsidize is an uncontrived warning because, if unable to impose that condition, Congress would not provide the funds given that by definition it does not want to subsidize the activity. Allowing this condition will thus make the persons that want to accept it better off without harming anyone.

Denying funds used for activities Congress does want to subsidize is likely a contrived threat because, if unable to impose the condition, Congress would likely provide the funds because the funds are by definition being used for activities Congress wants to subsidize. The Court thus correctly prohibits such threats, rejecting the contrary view that such threats are constitutional whenever the condition advances Congressional objectives.¹³⁶ The problem with this contrary view is that all it means is that Congress prefers Conditioned Funds > Funding Activities Congress Wants > No Funding. Thus, if unable to use the challenged condition, Congress would still fund the activities it wants. Accordingly, its threat to instead cut off all funding is a contrived threat that should be prohibited.

This precedent further holds that, if the activities Congress wants to subsidize can be separated from the exercise of constitutional rights without those federal funds,

¹³⁴ *AID*, 133 S. Ct. 2321, 2328-2332 (2013) (collecting cases).

¹³⁵ *Id.* at 2328.

¹³⁶ *Id.* at 2328, 2330.

then Congress must allow such separation.¹³⁷ Thus, Congress may not, by commingling funds, use an uncontrived warning to provide funds only for activities it wants as an excuse for a contrived threat to withhold those funds because persons engaged in protected unfunded activities. That is, when separated funding is possible and legislatively preferable to denying all funding, the refusal to provide separated funding would be a contrived threat.

Consider *Harris v. McRae*,¹³⁸ which held that Congress could not threaten to withhold non-abortion Medicaid funding from women who had abortions using nonfederal funds. Critics have argued that this conclusion offers no sound theory of whether the benefit of getting non-abortion Medicaid should be included in the transaction.¹³⁹ But that critique assumes non-abortion Medicaid would have been withheld in the but-for world without the condition. In fact, the Court was right that in that but-for world, Congress would (and demonstrably does) prefer funding non-abortion Medicaid to eliminating Medicaid. A threat to deny all Medicaid funding would thus be contrived. Accordingly, no “benefit” of non-abortion Medicaid would be caused by the posited condition. Rather, such a condition would coercively threaten to take a benefit that would exist without any condition.

b. Penalties v. Subsidies. The Supreme Court has held that conditioning government benefits on persons giving up their constitutional right to travel, speak, or observe their Sabbath is an unconstitutional penalty,¹⁴⁰ but that conditioning benefits on persons giving up their right to strike or avoid mandatory home visits is a constitutional subsidy.¹⁴¹ Critics have argued that these two lines of cases are inconsistent.¹⁴² But whether they are consistent depends on empirical predictions of what the government would likely do if unable to adopt these conditions.

In the former line of cases, it is clear that, if left with a choice between providing the benefits without the challenged condition and not providing the benefits at all, the government would choose to provide the benefits. Indeed, that is what has happened. The threats to deny those benefits were thus contrived threats to impose penalties relative to the but-for world.

The validity of the latter line of cases turns on whether one thinks that, if unable to adopt the challenged conditions, the government would have preferred to withhold

¹³⁷ *Id.* at 2328-2332

¹³⁸ *Harris v. McRae*, 448 U.S. 297, 317 n.19 (1980).

¹³⁹ Levinson, *supra* note , at 1345-47.

¹⁴⁰ *Hobbie v. Unemployment Appeals Comm’n*, 480 U.S. 136 (1987); *Shapiro v. Thompson*, 394 U.S. 618 (1969); *Speiser v. Randall*, 357 U.S. 513 (1958).

¹⁴¹ *Lyng v. International Union*, 485 U.S. 360 (1988); *Wyman v. James*, 400 U.S. 309 (1971).

¹⁴² Sullivan, *supra* note , at 1433-39.

the benefits rather than provide them without those conditions. The Court found that Congress adopted the no-strike condition because it wanted to make sure food stamps were not used to fund strikes, and that New York adopted the home-visit condition because it wanted to make sure welfare benefits were being used to advance the best interests of children.¹⁴³ If the desire to avoid such unwanted uses of federal funds were strong enough, then (if the conditions were unavailable) Congress might well prefer withholding these benefits to providing benefits without conditions that would prevent unwanted uses. If so, then the threats to deny those benefits were uncontrived warnings, and the Court was right that they did not impose penalties. The Court's implicit empirical predictions are not self-evidently wrong, and if they are correct the two lines of cases are normatively consistent. But courts should clarify the penalty/subsidy distinction by explicitly interpreting those terms to be measured relative to what the government would have done without the challenged condition.

c. Nexus or Germaneness. Another line of cases concerns situations where the government can deny a development permit outright, but cannot directly take some of the owner's property rights without just compensation. The government threatens to deny the permit unless the owner agrees to a condition that exacts some of its property rights, such as giving an easement to the public. The Supreme Court has held that such threats are constitutional only if the condition's purpose has an "essential nexus" to whatever purpose would be served by withholding the permit.¹⁴⁴ More generally, the pattern of Court results on unconstitutional conditions indicates it is more willing to uphold a condition that is "germane" to any government reasons for withholding the benefit.¹⁴⁵

If a nexus or germaneness exists between the condition and withholding the benefit, then a government that cannot use the condition is more likely to advance the same purpose by withholding the benefit. Offering an unconditioned benefit is less likely in the no-condition world because that does not serve the same purpose. Accordingly, nexus or germaneness indicates that threats to withhold benefits are likely uncontrived warnings that should make the condition constitutional.

If no nexus or germaneness exists between the condition and withholding the benefit, then inability to use the condition is less likely to cause the government to withhold the benefit because the two serve different purposes. Offering an unconditioned benefit is more likely in the no-condition world because giving the benefit does serve the same purpose as the benefit that would be given in a

¹⁴³ Lyng, 485 U.S. at 370-371; *Wyman*, 400 U.S. at 315-316, 318-319.

¹⁴⁴ *Dolan v. City of Tigard*, 512 U.S. 374, 386 (1994); *Nollan v. Cal. Coastal Comm'n*, 483 U.S. 825, 837 (1987).

¹⁴⁵ Sullivan, *supra* note , at 1457-76 (summarizing cases).

conditioned benefit. (For example, providing a conditioned permit and an unconditioned permit both provide whatever policy gains flow from real estate development.) Thus, a lack of nexus or germaneness indicates that threats to withhold benefits are likely contrived threats that should make the condition unconstitutional.¹⁴⁶

However, even if the condition and denying a permit further very different purposes, the government may value those different purposes equally, in which case different purposes need not suggest a contrived threat. For example, *Nollan* found no nexus between the purpose of the condition (providing public access to the beach) and the purpose of denying the permit (protecting visual access to the ocean).¹⁴⁷ Although those purposes are different, the government might value them the same and thus value denying the permit similarly to a conditioned permit. If so, then holding such conditioned permits unconstitutional will prevent uncontrived warnings and result in the permits being denied outright. This is by definition worse for the property owner and the government, because both prefer the conditioned permit to no permit. We can thus improve doctrine by interpreting nexus and germaneness to mean a similar government valuation of the condition and denial that is likely to make the threat uncontrived.

d. Plea Bargaining. In a plea bargaining agreement, a prosecutor threatens to bring criminal charges unless the defendant agrees to a certain sentence. The prosecutor thus offers a benefit (non-prosecution) in exchange for the defendant waiving constitutional rights. Some argue that plea bargains are coerced by the threat of imprisonment.¹⁴⁸ But plea bargains are usually enforced. The reason is that although the threat of prosecution may leave the defendant with no reasonable alternative to agreeing to the plea bargain, duress only exists if the threat is also improper, which the threat of prosecution ordinarily is not.

How can we determine which threats of prosecution are improper? By determining whether or not the threat is contrived. If the prosecutor is threatening to bring the same charges she would have brought if no plea bargain were possible, then she is simply offering an uncontrived warning. This does not penalize the criminal for exercising his constitutional rights because he would have faced the

¹⁴⁶ Some argue that any ban on conditioned permits without the needed nexus could easily be evaded by using eminent domain to take the easement and giving the permit as just compensation. Levinson, *supra* note , at 1349. But if the threat to deny the permit is contrived, then the permit would have been given in the but-for world and thus giving the permit provides no compensation at all. Giving the permit provides compensation only if the threat is uncontrived, in which case the conditioned permit would be constitutional anyway.

¹⁴⁷ 483 U.S. at 386-87.

¹⁴⁸ Albert W. Alschuler, *The Changing Plea Bargaining Debate*, 69 CAL. L. REV. 652, 695-703 (1981); John H. Langbein, *Torture and Plea Bargaining*, 46 U. CHI. L. REV. 3, 12-19 (1978).

same charges even in the no-threat world. In this case, offering a plea bargain can only make the defendant better off, by giving him the option of agreeing to a sentence he finds less harmful than the expected harm from prosecution in the no-threat world.

In contrast, suppose the prosecutor is overcharging: threatening to bring charges she would not have brought in the no-threat world in order to coerce a plea bargain. Then the prosecutor is making a contrived threat that does penalize the defendant for exercising his constitutional rights because he would not have faced the same charges in the no-threat world. Accordingly, plea bargains produced by such contrived threats should be unenforceable.

Usually, it is difficult to determine whether prosecutors have overcharged to induce plea bargains because there is no observable non-plea-bargain baseline. However, sometimes there is relevant evidence on this. For example, suppose a prosecutor initially brought certain charges but then, after the defendant refused to plead guilty, added more charges. The Supreme Court has indicated that such a clearly contrived threat likely violates due process.¹⁴⁹ But when prosecutors indicate their intention to bring certain charges before the plea negotiations begin, the Court found insufficient evidence to show a contrived threat that violated due process.¹⁵⁰ Still, if concrete evidence did exist that the prosecutor's initial charges were deliberately excessive to coerce plea bargains, that should suffice to show a due process violation.

3. Comparing Other Proposed Tests. Illustrious scholars have advocated many other tests for dealing with unconstitutional conditions. However, their tests all either underinclusively allow some contrived threats or overinclusively condemn some uncontrived warnings. Where their tests produce results that overlap the conclusions of a contrived-threat test, my analysis contributes a consequentialist proof in favor of those conclusions that has heretofore been lacking.

Kathleen Sullivan argues that courts should focus on the distributive effect of conditioned benefits.¹⁵¹ Such distributive effects may well show that the selective denial or grant of benefits violates equal protection or the constitutional rights of third parties. But if so, then either the threatened denial of benefits or agreements granting benefits are themselves unconstitutional regardless of any coercion. Cass Sunstein makes a similar argument, but recognizes that it amounts to abandoning

¹⁴⁹ *Bordenkircher v. Hayes*, 434 U.S. 357, 360-61 (1978)

¹⁵⁰ *Id.*

¹⁵¹ Sullivan, *supra* note at 1421.

the doctrine of unconstitutional conditions doctrine.¹⁵² Their analysis thus cannot explain the puzzle of why conditions are sometimes held unconstitutionally coercive even when the threatened denial of benefits and agreement to grant benefits are both lawful. Further, when both are lawful, their tests would immunize all conditioned benefits, which would incorrectly allow contrived threats that penalize the exercise of constitutional rights

Philip Hamburger argues that all conditioned benefits should be unconstitutional whenever the condition waives a constitutional right.¹⁵³ His test cannot explain the puzzle of why courts often hold precisely the opposite. His test would also incorrectly condemn all uncontrived warnings, which would harm both the government and the persons with the constitutional rights in question.

Richard Epstein argues that current constitutional law deviates from ideal constitutional law by weakly protecting contract and property rights and allowing government redistribution of wealth. To move us closer to his ideal, he argues that conditioning benefits should be: (a) unconstitutional when his ideal constitution would not allow denial of the benefit (like preventing businesses from operating); but (b) constitutional when his ideal constitution would not allow providing the benefit (like providing food stamps).¹⁵⁴ One obvious problem is that one might disagree with his ideal constitution. But the problem is deeper: his approach would actually move us further from his own ideal. When the threat is to deny one of his ideally-required benefits, he and I would both condemn contrived threats but he would also condemn uncontrived warnings, thus resulting in precisely the denials that he fears, as well as preventing a government-person bargain that would make them both better off. When the threat is to deny one of his ideally-prohibited benefits, he and I would both allow uncontrived warnings but he would also allow contrived threats, thus allowing the government to use their ideally-prohibited power to penalize the exercise of constitutional rights that Epstein himself deems ideal.

¹⁵² Cass R. Sunstein, *Why the Unconstitutional Conditions Doctrine Is an Anachronism (With Particular Reference to Religion, Speech, and Abortion)*, 70 B.U. L. REV. 593 (1990). Sunstein recommends abandoning the doctrine for two reasons. First, he rejects what he assumes are the only possible baselines to determine when conditions are a penalty or subsidy: (a) what government could do; (b) the common law; or (c) a historical status quo baseline. *Id.* at 597-604. He is right to reject those tests, but he does not consider a predictive standard that asks what government would do if it could not condition the benefit. Second, he notes that penalizing constitutional rights is not always unconstitutional. *Id.* at 603. But when that is so, direct coercion is permissible, and thus a key premise for the doctrine is not met.

¹⁵³ Philip Hamburger, *Unconstitutional Conditions: The Irrelevance of Consent*, 98 VA. LAW REV. 479 (2012).

¹⁵⁴ Richard A. Epstein, *Unconstitutional Conditions, State Power, and the Limits of Consent*, 102 HARV. L. REV. 4, 27-104 (1988).

Mitchel Berman argues that threatening to deny benefits should be presumptively unconstitutional if the purpose of carrying out the threat would be to make it more costly to exercise a constitutional choice.¹⁵⁵ Unfortunately, purpose is a murky test, especially because different legislators have different purposes, and Berman stresses that his purpose test deviates from a pure predictive test like mine.¹⁵⁶ Nor does he define the grounds for rebutting his presumption. Wherever his test produces conclusions that deviate from a contrived-threat test, his test will either undesirably allow contrived threats or undesirably prohibit uncontrived warnings.

Kenneth Simons would deviate from a predictive baseline when the government plausibly believes the condition is relevant to the legitimate purposes of the government program.¹⁵⁷ But such deviations would undesirably allow contrived threats that penalize the exercise of constitutional rights whenever the condition has some relevance but that relevance is not important enough to mean the benefit would be denied without that condition. As Simons acknowledges, his conclusion conflicts with caselaw prohibiting the government from conditioning benefits on working the Sabbath.¹⁵⁸ His argues that without his related-purpose exception, the government could not limit political campaigning by public employees.¹⁵⁹ But while such limits do seem permissible, the reason is not that they are uncoercive but because the constitutional interest of undistorted politics justifies compelling public employees not to campaign, making coercion irrelevant.

Seth Kreimer proposes a mix of historical, equality, and predictive baselines to judge whether threats to withhold benefits are unconstitutionally coercive.¹⁶⁰ However, whenever his historical or equality baselines allow conditions that a predictive baseline would condemn, his test wrongly allows contrived threats that penalize the exercise of constitutional rights. Whenever his historical or normative baselines condemn conditions that a predictive baseline would allow, his test wrongly condemns uncontrived warnings that would benefit both the government and persons with the constitutional right.

True, Kreimer's historical baseline is relevant to the right test because the past government conferral of unconditioned benefits helps suggest it would have likely

¹⁵⁵ Mitchell N. Berman, *Coercion Without Baselines: Unconstitutional Conditions in Three Dimensions*, 90 GEO. L.J. 1, 7, 46 (2001).

¹⁵⁶ *Id.* at 46.

¹⁵⁷ Kenneth W. Simons, *Offers, Threats, and Unconstitutional Conditions*, 26 SAN DIEGO L. REV. 289, 291-92, 318-20, 323 (1989).

¹⁵⁸ *Id.* at 323.

¹⁵⁹ *Id.* at 319.

¹⁶⁰ Seth Kreimer, *Allocational Sanctions: The Problem of Negative Rights in a Positive State*, 132 U. PA. L. REV. 1293, 1359-78 (1984).

continued to do if it could not add the condition. However, government policy does change over time, and the historical baseline should be rejected whenever there is evidence that it does not accurately predict what the government would have done without the condition. Kreimer argues otherwise because persons rely on prior benefits.¹⁶¹ But when the historical test deviates from the predictive test, it will be either when it (1) allows undesirable contrived threats to withhold a new benefit, for which reliance is irrelevant, or (2) condemns desirable uncontrived warnings about the coming denial of old benefits, in which case his test would produce a denial of old benefits that affirmatively harms reliance interests.

Likewise, Kreimer's equality baseline is relevant to the right test because government conferral of a benefit on all the persons who accept the condition may suggest a purpose for doing so that would likely extend to everyone if it could not adopt the condition. However, to accurately judge whether the condition unlawfully coerces the threat recipient, this equity baseline should be rejected when it deviates from the predictive baseline. After all, when there is such a deviation, the equity baseline will either (1) allow contrived threats to withhold benefits that others do not get, which penalizes the constitutional autonomy right, or (2) condemn uncontrived warnings about the coming denial of benefits that others do get, in which case his test would produce a denial of benefits that harms everyone.

IV. THREATS THAT INDUCE NEW PRIVATE CONTRACTS

A contract is obviously unenforceable if the terms are independently illegal or if induced by threatening illegal conduct. But threats of otherwise lawful action to induce an otherwise lawful new contract can also constitute duress when the threat is improper and the victim has no reasonable choice.¹⁶² Such threats are deemed "improper" only when they are contrived.

A. Threats to Exercise Unrelated Rights

Even if a party has every legal right to terminate a contract, a threat to do so "for some purpose unrelated to the contract, such as to induce the recipient to make an entirely separate contract, is ordinarily improper."¹⁶³ Such facts indicate the threatener had no reason not to continue performing the first contract, and thus

¹⁶¹ *Id.* at 1362.

¹⁶² Restatement §175-176.

¹⁶³ *Id.* §176 comment e & Illustrations 9, 11.

would have performed it if unable to communicate a threat designed to induce the separate contract. Such a threat is thus contrived.

The same can be true when a party threatens not to enter one contract to induce another. In one case, a lessor said it would rent to the purchaser of a leaving tenant's business, but then threatened not to do so unless the leaving tenant forked over 45% of his business sale proceeds. This threat was held improper and thus invalidated the new contract to pay 45% of the business sales proceeds.¹⁶⁴ This fits the contrived-threat test because the lessor had clearly indicated it was happy to rent to the new tenant and threatened otherwise only to extort part of the business sale proceeds.

In short, when a party threatens to exercise unrelated rights to induce a new contract, such threats are improper when the threat is contrived; that is, when the party would not exercise the unrelated rights in that way but for the ability to communicate the threat. Such contrived threats impose penalties the victim would not have incurred in the no-threat world. Because the threat of such penalties can induce the victim to agree to terms on the new contract that are not in its interest, they eliminate the normal assurance that the new contract allocates resources to the highest-value user. Thus, such contracts are properly held unenforceable.

If, instead, the threat were to exercise unrelated rights in a way that would be in the interests of the threatener even if no new contract could be induced, then the threat would be an uncontrived warning. Because the unrelated rights would be exercised in an adverse way in the no-threat world, the new contract must make the threat recipient better off because he prefers the new contract terms to having the unrelated rights exercised adversely. Thus, in such cases, the new contracts are properly held enforceable.

B. Threats Not to Enter Into the Challenged New Contract

1. Threats to Withhold Desired Goods Unless the Market Price Is Paid. The typical threat that induces a new contract is a threat to withhold goods unless the market price is paid. Some argue that all such threats are coercive because individuals have no reasonable choice but to pay to get the goods they need to live.¹⁶⁵ However, if the seller could not sell to the individual at the market price, then the seller would withhold the good because it would prefer to sell to someone else for the market price. The threat to withhold the good is thus an uncontrived

¹⁶⁴ *Wolf*, 57 N.J. Super. at 287-288.

¹⁶⁵ Sullivan, *supra* note , at 1446-47 (collecting sources).

warning. If duress law made the contract unenforceable, this individual would thus not get the good that they need. (Of course, if price regulations generally required a lower price, then some individuals would get the good, but recall our premise is that the terms of the contract are not independently illegal.) This is true even if the seller is a monopolist because then the market price is the monopoly price, and the seller would prefer to sell to someone else at the monopoly price rather than sell to any particular individual at a sub-monopoly price. Thus, refusals to sell for less than the market price are enforceable no matter how much the buyer needs the good.¹⁶⁶

*Batsakis v. Demotsis*¹⁶⁷ shows that this point holds even in extreme cases. In war-torn Greece in 1942, Demotsis borrowed 500,000 drachmae (worth about \$25) in exchange for a promise to later pay \$2,000 after the war. The court held this agreement was enforceable. Because the war-torn circumstances doubtless made Demotsis desperate for the money, critics have argued that her promise was extracted under duress.¹⁶⁸ But those same war-torn circumstances also likely made having cash highly valuable and lending money highly risky, and no evidence was offered that the high rate charged differed from the going market price. Thus, if Batsakis did not believe the contract would be enforceable, he likely would not have lent the money to Demotsis, which would have made her even worse off given her apparent desperate need for money. Because the threat to withhold the money unless promised \$2000 was thus an uncontrived warning, the contract was properly held enforceable. Moreover, assuming the high market rate reflected the desperate need of persons in war-torn Greece for cash, we want such high rates to be enforceable in order to encourage others to increase the supply of cash to war-torn Greece to meet that need.

2. Threats to Withhold Desired Goods When Parties Have Been Cut Off From the Market. Sometimes parties might be cut off from the market, so that it does not provide an effective constraint. If the buyer has been cut off, a seller might threaten not to sell a product unless it receives an above-market price. If the seller has been cut off, the buyer might threaten not to buy unless it receives a below-market price. When should such threats be deemed contrived?

Consider first cases where a seller has taken steps to cut the buyer off from the market in order to try and charge the buyer an above-market price. Then a threat not to sell at less than an above-market price is clearly contrived because if that

¹⁶⁶ Restatement §176 Comment a & Illustration 14; *Standard Box Co. v. Mutual Biscuit Co.*, 10 Cal. App. 746, 758, 761-62 (1909).

¹⁶⁷ 226 S.W.2d 673 (Tex. Civ. App. 1949).

¹⁶⁸ FRIED, *supra* note, at 109-111.

threat could not be made, then the seller would be willing to sell to the buyer at the market price and would not have taken steps to cut off the buyer to begin with. For example, if a seller misleads a buyer into thinking he will supply a good at the market price and thus causes the buyer to forego other options until it is too late, then a threat not to sell unless an above-market price is paid is contrived and a contract at that above-market price is unenforceable.¹⁶⁹

Now suppose a seller, through no productive effort, stumbles upon a buyer who has been cut off from the market for adventitious reasons. For example, suppose Bill Gates shows up at the emergency room needing immediate lifesaving treatment, and the doctor says she will not treat him unless he promises to pay her \$1 billion. Then the doctor's threat is contrived because if that threat were not possible, the doctor would be willing to treat Bill Gates at the doctor's usual market rate. Thus, Bill Gate's agreement to pay \$1 billion is unenforceable.¹⁷⁰

If, in the adventitious case, there is no clear market rate, an alternative price baseline might still show that the seller made a contrived threat. This point is illustrated by *Post v Jones*.¹⁷¹ In the 1849, three whaling ships happened to find another whaling ship stranded on rocks in the Arctic right before winter and agreed to buy the stranded ship's whale oil for a very low price. Although the whale oil was nominally "auctioned," the stranded ship had more oil than the three ships could take, so they had no need to compete with each other on price.¹⁷² The Supreme Court recognized that even when salvage is an "absolute necessity," parties can set salvage prices by contract when the stranded ship has access to "a market and competition."¹⁷³ However, the Court held that this contract was unenforceable for two reasons. First, the stranded ship had no reasonable choice but to take whatever the three ships offered because "there was no market [and] no competition."¹⁷⁴ Second, taking the whale oil at the default price (which is set by salvage law when no binding contract exists) was still highly profitable because salvaging the oil required little risk or effort and the opportunity costs of filling the cargo space were low given that the whaling season was ending.¹⁷⁵

¹⁶⁹ Restatement §176(2)(b) & Comment f & Illustration 13.

¹⁷⁰ Steven Shavell, *Contract Holdup and Intervention*, 36 J. LEGAL STUD. 325, 338 (2007) (noting that courts hold adventitious holdup contracts unenforceable when they are substantially above market prices).

¹⁷¹ 60 U.S. 150 (1856).

¹⁷² *Id.* at 159.

¹⁷³ *Id.* at 157-158; *The Elfrida*, 172 U.S. 186, 196-197, 204 (1898) (stressing that the law does not generally regulate the reasonableness of salvage prices and allowing a high contract price when the stranded ship had competitive salvage bids).

¹⁷⁴ 60 U.S. at 159.

¹⁷⁵ *Id.* at 160-161.

The second finding in *Post v. Jones* meant that the three ships' threat not to take the oil at the default rate was a contrived threat because, if they could not make that threat, they would have happily salvaged the oil at the still highly-profitable default rate. No salvaging was the wrong but-for baseline because salvaging would have occurred in the no-threat world. Indeed, the Court explicitly rejected the three ships' claim that the stranded ship was not harmed because it was better off than if no salvaging had been provided.¹⁷⁶ The effects of allowing salvagers to price below the default salvage price in such accidental monopoly cases is harmful because if ships know that rescues will cost them almost all the value of their cargo, they will take excessive precautions to avoid the need for rescue.¹⁷⁷ In contrast, because the default price was high enough to make salvage profitable and finding the stranded ship was adventitious, using it as the baseline for judging whether a threat is contrived will not reduce any salvage effort. But this result depends critically on having some known desirable price baseline to determine that a contrived threat was made.

Now suppose a situation where a seller earns being in a situation where the buyer is cut off from the market. For example, suppose the rescue ships were actually in the business of searching the seas for stranded ships. If the default price provided by salvage law suffices to compensate those search costs, then a threat to withhold salvage unless paid more than that default price would be contrived because absent an ability to make that threat, the rescue ship would have incurred the search costs and provided salvage. If the default price provided by salvage law does not suffice to compensate for those search costs, then a threat to withhold salvage unless paid more than that default price would be an uncontrived warning because absent an ability to make that threat, the rescue ship would not have incurred the search costs and thus not been there to provide salvage. This result could also be achieved by adjusting the default price to compensate for efficient search costs.¹⁷⁸

3. Bilateral Negotiation to Split Extra-Market Surplus. Suppose neither party has been cut off from the market but the seller's goods are worth more to the buyer than to any other buyer. This is a frequent situation, often because of relationship-specific investments. An employee who works at a firm may develop firm-specific skills that make the employee more valuable to the firm than to any other firm. Or a business that rents a location may find at renewal that the location is more valuable to it than any other location is. In these cases, we have a bilateral negotiation over how to split an extra-market surplus because although the buyer

¹⁷⁶ *Id.* at 160.

¹⁷⁷ STEVEN SHAVELL, FOUNDATIONS OF ECONOMIC ANALYSIS OF LAW 336 (2004).

¹⁷⁸ *Id.* at 335-36.

cannot find any other seller whose goods it values as much, the seller also cannot find any other buyer who values its goods as much. There will be a bargaining range, with the minimum price the seller would take being the market price other buyers would pay, and the maximum price the buyer would pay being the higher value it puts on the good. The results of such bargaining are indeterminate.

In these cases of bilateral negotiations about the split of extra-market surplus, the contrived-threat test is not helpful because we have no baseline price to assess whether a contrived threat has been made. One might think that a contract at any price above the seller's reservation price must reflect a contrived threat because the seller would take the reservation price if it were unable to insist on a higher price. However, it is also the case that any price below the buyer's reservation price must reflect a contrived threat by the buyer because the buyer would take its reservation price if it were unable to insist on a lower price. Such negotiations thus cannot help but reflect contrived threats by *both* sides, and a contrived-threat test thus cannot give either side a unique right to void the contract. Moreover, both sides would be worse off if contracts between them were unenforceable because any contract between them leaves them better off than a contract with others. Thus, contracts produced by such bargaining are enforceable.¹⁷⁹

C. Blackmail

In the case of blackmail, the threat is to divulge information to third parties unless the victim agrees to pay money. Blackmail law had long been regarded as “one of the most elusive intellectual puzzles in all of law.”¹⁸⁰ Contrived-threat analysis helps solve that puzzle by showing that all blackmail contracts are necessarily produced by either contrived threats or uncontrived warnings and that each has distinctive harmful effects. Contrived blackmail threats are coercive and cause inefficiency between the contracting parties, without affecting the information revealed to third parties. Uncontrived blackmail threats are noncoercive and benefit the contracting parties, but prevent information disclosures to third parties whose interests often trump the interests of the contracting parties. To explain blackmail law, we need to understand the combination of these effects.

¹⁷⁹ Restatement §176 Comment a. One might think that any price above the seller's reservation price is undesirable because it discourages the buyer (say a tenant) from making relationship-specific investments that make the seller's goods more valuable to it. However, the seller might also make relationship-specific investments that make its goods more valuable to the buyer. Moreover, if buyer investments of this sort are efficient, the seller can help pay for them. For example, in commercial leases, sellers often give a tenant improvement allowance to induce tenants to begin a lease.

¹⁸⁰ James Lindgren, *Blackmail: An Afterword*, 141 U. Pa. L. Rev. 1975, 1975 (1993).

1. Contrived Blackmail Threats Are Always Coercive and Invalid. Consider first cases when divulging the relevant information “would not significantly benefit the party making the threat.”¹⁸¹ Given that the act of revealing this information would not significantly benefit the threatener, it is likely a contrived threat. True, if divulging information is very cheap, one might wonder whether the threatener would divulge the information in the no-threat world for even insignificant benefits. But effort is necessary to collect information and (even when acquired fortuitously) to document the information in a way that can be divulged in a way that is sufficiently persuasive to be effective and overcome the risk of suits for libel or slander. A blackmailer would not likely incur the costs of such effort for insignificant benefits without the prospect of making blackmail threats.

Thus, any threat to divulge information is likely to be a contrived threat if divulging the information confers no significant benefit to the threatener. Allowing such contrived blackmail threats would harm the victim because in the no-threat world he would not pay money and the information would still not be divulged. Because such contrived threats do not affect the revelation of information, preventing such blackmail would not affect third parties. Finally, preventing such contrived blackmail threats would reduce blackmailer efforts that inefficiently incur costs to induce nothing but an unproductive transfer of wealth. Because these effects are always negative, contract law correctly condemns all blackmail contracts induced by such contrived threats, regardless of what sort of information is at issue.

2. Uncontrived Blackmail Threats Benefit Contracting Parties But Harm Third Party Interests. Consider now cases where divulging information would significantly benefit the threatener. For example, suppose the information was collected by an investigative reporter who would, in the no-threat world, publish the information in an article. In that case, the threat to divulge the information is clearly an uncontrived warning.¹⁸² Accordingly, an agreement to suppress the information is in the interests of both contracting parties because the victim prefers paying the money to having the information divulged and the blackmailer prefers getting the money to revealing the information. But this blackmail contract does prevent a disclosure of information to third parties that otherwise would have occurred. Whether to invalidate blackmail contracts induced by uncontrived

¹⁸¹ Restatement (Second) of Contracts §176(2)(a) & Comment f (“Clause (a) is concerned with cases in which a party threatens to do an act that would not significantly benefit him but would harm the other party.... A typical example is a threat to make public embarrassing information concerning the recipient unless he makes a proposed contract.”)

¹⁸² Some believe that the costs of collecting and revealing information are sometimes so trivial that it would occur even without any significant benefit. To the extent they are right, such threats are also uncontrived warnings and have the effects described in this section.

warnings thus turns on whether third party interests in the information trump the interests of the contracting parties, which varies with the information at issue.

In particular, when the information concerns crimes, the public obviously has a powerful interest in knowing the information because that helps deter crimes that harm the public. Because the public interest in discovering crimes exceeds the interests of the contracting parties in concealing crimes, such blackmail contracts are invalid even when uncoercive. Accordingly, contract law correctly condemns all blackmail about crimes, whether or not divulging the information would confer a significant benefit that makes the threat uncontrived.¹⁸³ The reason is that all such blackmail about crimes necessarily involves either (1) a contrived threat that creates inefficiency for the contracting parties without affecting information disclosure or (2) an uncontrived warning that suppresses the disclosure of criminal information to third parties whose interests trump the interests of the contracting parties.

When the information involves embarrassing noncriminal information, then third party interests are weaker and it is less clear whether they outweigh the interests of the contracting parties. In the investigative reporter example, one might worry about the harm to the reporter's newspaper or its readership from suppressing the information. On the other hand, one might regard their interests in salacious information as weak. Such judgments vary by jurisdiction and court and with the specific information at issue. General contract law thus does not condemn agreements induced by threats to divulge embarrassing noncriminal information when divulging would significantly benefit the threatener (i.e., when it is an uncontrived warning),¹⁸⁴ leaving it to other bodies of law to decide when such noncoercive contracts should be void on public policy grounds.

Why, however, do some criminal blackmail statutes seem to overinclusively condemn even uncontrived warnings about information with little legitimate third party interest? Perhaps partly because adjudicative errors in distinguishing between contrived and uncontrived blackmail threats would create strong underdeterrence concerns if such an exception were recognized. Nor are blackmail statutes literally applied in certain cases. Consider the fact that litigation usually ends in a settlement where someone gets money and agrees not to disclose information about the other party. Because such confidentiality has value, something must have been paid for it in the settlement, so literally this exchange fits broad definitions of blackmail. But no one treats it that way. The analysis here

¹⁸³ *Id.* §176(1)(b)&(2)(a) & Comments c, f.

¹⁸⁴ *Id.* §176(1)(b) & Comment c.

explains why. Litigation already provides copious incentives to collect and document the information, so disclosure is costless. Further, litigants generally dislike each other, so disclosure likely produces serious psychic benefits. Thus, threats to divulge the information collected in litigation are particularly likely to be uncontrived. Adjudicative errors and under-deterrence are thus less of a concern and an implicit exception is recognized. The analysis here thus helps explain not only why blackmail prohibitions are so broad, but also why nonliteral exceptions are sometimes recognized.

3. Prior Literature. Prior works on blackmail have tended to find it perplexing because they either focus on the legality of the threatened act rather than on the practical effects of the threats,¹⁸⁵ or because they focus on some of those practical effects to the exclusion of others. Some articles focus on protecting the interests of third parties in information,¹⁸⁶ but they cannot explain the universal ban on contrived blackmail threats, which do not affect third party information. Other works focus on the harm to victims and the inefficiency of efforts to collect information with no productive gain,¹⁸⁷ but they cannot explain the ban on many uncontrived blackmail warnings, where the victim benefits, the information would be collected anyway, and the blackmail contract does alter the information revealed to third parties. Scott Altman has offered what he calls a “patchwork” theory that combines many of these effects.¹⁸⁸ What contrived-threat analysis clarifies is the deeper structure that all blackmail must have one or the other of these sets of effects and that only the combination can explain the full scope and limits of blackmail law.

V. CONCLUSION

Contractual duress, unconstitutional conditions, and blackmail have been longstanding puzzles because the focus has been on the wrong baseline: what the threatener should do or legally could do. One must instead focus on what the threatener *would* do absent the threat to resolve these puzzles. Only contrived

¹⁸⁵ Glanville Williams, *Blackmail*, 1954 CRIM. L. REV. 79, 163; Jeffrie G. Murphy, *Blackmail: A Preliminary Inquiry*, 63 MONIST 156, 156-58 (1980).

¹⁸⁶ Lindgren, *Unraveling the Paradox of Blackmail*, 84 COLUM. L. REV. 670 (1984); Epstein, *Blackmail, Inc.*, 50 U. CHI. L. REV. 553 (1983).

¹⁸⁷ Douglas H. Ginsburg & Paul Shechtman, *Blackmail: An Economic Analysis of the Law*, 141 U. PA. L. REV. 1849, 1857-60 (1993); Leo Katz, *Blackmail and Other Forms of Arm-Twisting*, 141 U. PA. L. REV. 1567, 1598-99, 1615 (1993); Wendy J. Gordon, *Truth And Consequences: The Force Of Blackmail's Central Case*, 141 U. PA. L. REV. 1741 (1993); Coase, *Blackmail*, 74 VA. L. REV. 655, 673-74 (1988); ROBERT NOZICK, ANARCHY, STATE, AND UTOPIA 84-87 (1974).

¹⁸⁸ Scott Altman, *A Patchwork Theory of Blackmail*, 141 U. PENN L. REV. 1639 (1993).

threats to take action that differs from this but-for action can harm the recipient of the threat and be coercive. Uncontrived warnings can only benefit the agreeing parties and thus should be unlawful only when even an uncoerced agreement would be unlawful, usually because the agreement harms third parties.