
Safe Harbors, Sure Shipwrecks

Susan C. Morse*

TABLE OF CONTENTS

INTRODUCTION	1387
I. BRIGHT-LINE RULES, LEGAL STANDARDS, SAFE HARBORS, SURE SHIPWRECKS	1390
A. <i>Key Terms</i>	1390
B. <i>Accumulation of Safe Harbors and Sure Shipwrecks</i>	1392
C. <i>The Issue of Precision</i>	1394
D. <i>Safe Harbors Give “Good” Results, Sure Shipwrecks Give “Bad” Results</i>	1395
II. SAFE HARBORS AND SURE SHIPWRECKS AFFECT BEHAVIOR INCENTIVES	1397
A. <i>Convergence for Safe Harbors; Bunching for Sure Shipwrecks</i>	1397
1. <i>Drunk Driving Example</i>	1398
2. <i>Figure 1</i>	1399
3. <i>Lobbyist-Legislator Transfer Example</i>	1402
B. <i>The Discontinuity or “Notch” Feature of Safe Harbors and Sure Shipwrecks</i>	1404
C. <i>Probability of Liability Could Be Greater than 0% for Safe Harbor, Less than 100% for Sure Shipwreck</i>	1407
D. <i>Who Are the “Persons Subject to” a Safe Harbor or Sure Shipwreck?</i>	1408

* Copyright © 2016 Susan C. Morse. Assistant Professor, University of Texas School of Law. Many thanks to Ronen Avraham, Ward Farnsworth, Willy Forbath, James Spindler, and Alex Raskolnikov for their comments on multiple drafts. Thanks also for helpful comments to Lynn Baker, Lynn Blais, Oren Bracha, Emily Cauble, Mark Gergen, John Golden, David Hasen, Bert Huang, Calvin Johnson, Edward Kleinbard, Leandra Lederman, Saul Levmore, Yair Listokin, Ruth Mason, Tom McGarity, Leigh Osofsky, Philip Postlewaite, Julie Roin, Frederick Schauer, Steve Shay, Abe Wickelgren, and Deborah Widiss; and to participants at law school workshops at University of Southern California, UC Berkeley, UC Hastings College of the Law, University of Virginia, University of Texas, UCLA, Indiana University, Northwestern University, and Harvard University; and at the Harvard-Stanford-Yale Junior Faculty Forum.

III. FACTORS ENDOGENOUS TO SAFE HARBOR AND/OR SURE	
SHIPWRECK.....	1411
A. <i>Market Coordination</i>	1411
B. <i>Probability of Violation in General Standard Space May Increase, Decrease or Stay the Same</i>	1412
C. <i>Policy Maker Framing</i>	1413
D. <i>Confined to Facts?</i>	1414
E. <i>Comparison to Priors</i>	1415
F. <i>When a Space Includes Both a Safe Harbor and a Sure Shipwreck</i>	1416
IV. WHEN AND HOW TO USE SAFE HARBORS AND SURE	
SHIPWRECKS.....	1418
A. <i>Overinclusion and Underinclusion</i>	1419
B. <i>Ex Ante and Ex Post</i>	1424
C. <i>Safe Harbors and the Problem of Capture</i>	1427
CONCLUSION.....	1429

INTRODUCTION

The choice between rules and standards has received lavish attention from legal scholars. But this choice fails to capture how the law actually works. On the ground, legal regimes use combinations of rules and standards. The interaction among rule and standard building blocks, as much as the choice between rules and standards, influences the behavior of those who are subject to legal regimes.

This Article analyzes safe harbors and, as I call them, sure shipwrecks. These hybrids between rules and standards exist everywhere in the law. Existing work identifies some features of these rule/standard hybrids. Yet the literature lacks any theory that analyzes the choices and incentives presented by safe harbors and sure shipwrecks to persons subject to these ubiquitous hybrid legal regimes. This Article supplies such a theory.

To briefly define the key terms:

A bright-line rule applies a categorical legal result. It provides the result *ex ante*, before the facts have arisen in a particular case. Examples include statutes of limitations and the requirement that drivers drive on one side of the road. An environmental law that prohibits emissions of a particular substance at or above ten parts per million and permits emissions of that substance below ten parts per million is a bright-line rule.

Standards provide general considerations for reaching a decision, but leave the determination of legal results to future decision makers. Standards produce legal results on an *ex post* basis. Examples include “undue burden” tests in constitutional law and the requirement of “just and reasonable” rates in the regulation of public utility monopolies. An environmental law that only allows “safe and healthy” emissions levels is a standard.

A safe harbor combines a rule and a standard. It provides by rule that particular facts comply with the law and will result in no penalty. It leaves other facts to be judged by a standard. For example, handing a subpoena to a summoned individual is generally accepted as valid “delivery” of the subpoena. If a jurisdiction allows a court to conclude that other methods also qualify as a “delivery,” then handing a subpoena to an individual is a safe harbor from the perspective of the server. Other methods of delivery, such as showing the subpoena to the individual through the window of her home, will be judged according to a standard.

A sure shipwreck also combines a rule and a standard. It is the converse of a safe harbor. A sure shipwreck describes conduct that will definitely violate the law, while other facts remain subject to a

standard as applied by the ex post judgment of future decision makers. Automatic liability regimes generally function as sure shipwrecks. For example, a rule that provides for automatic liability when one driver rear-ends another car functions as a sure shipwreck. If a driver contributes to an accident, but does not hit another car from behind or otherwise fall within an automatic liability rule, his negligence will be judged according to a standard.

Existing literature theorizes rules and standards as a choice between two regime paradigms.¹ But this is not how it really works. Statutory drafters, administrative regulators and judges constantly provide legal answers for specific fact patterns — but do not provide all of the answers. Safe harbors and sure shipwrecks result.² These rule-standard hybrids provide boundaries on uncertain spaces in most legal regimes.³

An environmental law might permit emissions of a particular substance at or below one part per million, and simultaneously prohibit emissions of a particular substance at or above ten parts per million. Such a law would leave it for a later decision maker to

¹ E.g., Isaac Ehrlich & Richard A. Posner, *An Economic Analysis of Legal Rulemaking*, 3 J. LEGAL STUD. 257, 257 (1974) (positing a “choice along the specificity-generality continuum”); Louis Kaplow, *Rules Versus Standards: An Economic Analysis*, 42 DUKE L.J. 557, 568-69 (1992) (contrasting ex ante rules and ex post standards); Frederick Schauer, *Formalism*, 97 YALE L.J. 509, 520-21 (1988) (“Formalism in this sense is . . . the denial of choice to the judge.”); Kathleen M. Sullivan, *The Justices of Rules and Standards*, 106 HARV. L. REV. 22, 57-58 (1992) (arranging rules and standards along a “continuum of discretion”); Cass R. Sunstein, *Problems with Rules*, 83 CALIF. L. REV. 953, 969 (1995) (comparing rules and “basic principles”).

² See, e.g., Saul Levmore, *Double Blind Lawmaking and Other Comments on Formalism in the Tax Law*, 66 U. CHI. L. REV. 915, 920 (1999) (“[S]afe harbors [may] minimize the problem of discontinuities.”); Peter P. Swire, *Safe Harbors and a Proposal to Improve the Community Reinvestment Act*, 79 VA. L. REV. 349, 371-75 (1993) (suggesting that safe harbor may increase certainty, reduce underinclusion and overinclusion, and/or serve transitional purpose); see also Emily Cauble, *Safe Harbors in Tax Law*, 47 CONN. L. REV. 1385, 1398-1428 (2015) (arguing that safe harbors “exacerbate distortions” compared to standards but are more “forgiving” and less likely to distort behavior compared to bright-line rules); Andrew Stumpff Morrison, *Case Law, Systemic Law, and a Very Modest Suggestion*, 35 STATUTE L. REV. 159, 174 (2013) (treating safe harbors and “unsafe harbors” as “synthetic case law” and arguing that such rules should address “easy cases” against the background of a standard); Gideon Parchomovsky & Alex Stein, *Catalogs*, 115 COLUM L. REV. 165, 168 (2015) (analyzing a type of sure shipwreck, the “catalog,” defined as “an outright ban on a detailed, but incomplete, list of specific activities and a general prohibition of all activities falling into the same category”).

³ Cf. John E. Calfee & Richard Craswell, *Some Effects of Uncertainty on Compliance with Legal Standards*, 70 VA. L. REV. 965, 966-67 (1984) (describing and analyzing uncertainty under legal standards).

determine whether emissions between one part per million and ten parts per million were acceptable under a “safe and healthy” emissions standard. The safe harbor and sure shipwreck provisions in this example limit the regime’s legal uncertainty to emissions between one and ten parts per million.

This Article shows that safe harbors and sure shipwrecks affect behavior differently, and indeed asymmetrically. Safe harbors encourage those subject to the rule to change their behavior so that their behavior converges on the boundary line drawn by the safe harbor. Sure shipwrecks do not work the same way. They encourage parties who are in the sure-shipwreck zone to change their behavior in order to avoid definite penalties, so that behavior bunches immediately on the compliant side of the line. But under a first-order analysis, sure shipwrecks do not give persons outside the penalty zone reason to flock to the boundary line.

Consider a safe harbor law that provides that someone who drives after one drink cannot be charged with a driving offense related to alcohol use. Some drivers who without the safe harbor would abstain will instead have one drink before driving. Also, some drivers who previously drove after two drinks will have only one drink. This is because the option of having one drink is more attractive than it was before the appearance of the safe harbor. Before the safe harbor, some positive chance of liability attached to driving after one drink; after the safe harbor, that chance of liability is zero.

In contrast, consider a drunk driving rule that provides for misdemeanor criminal penalties if an individual drinks four drinks before driving. This is a sure shipwreck. Drivers will, on the margin, limit their drinking to less than four drinks. But the sure shipwreck generally will not encourage drivers to increase their drinking from, say, three drinks to four. The consequences of drinking four drinks have gotten worse as a result of the appearance of the sure shipwreck, and that makes the four-drink alternative less attractive, not more attractive.

The core argument that safe harbors and sure shipwrecks affect behavior asymmetrically rests on the assumption that the appearance of a safe harbor or sure shipwreck does not affect the probability of liability for behavior that lies outside the safe harbor or sure shipwreck. The argument also assumes that participants will obtain the same expected value (leaving aside the possibility of liability) from different compliant and noncompliant behavior with and without the safe harbor or sure shipwreck. These are strong assumptions. When they are relaxed, the insights in this Article shift somewhat and

become more complex, but the core argument remains. This argument predicts convergence on a safe harbor, but it does not as a general matter predict convergence on a sure shipwreck.

This Article proceeds as follows. Part I defines safe harbors and sure shipwrecks, as well as bright-line rules and legal standards. Part II presents the core thesis. It argues that placing a safe harbor against the background of a legal standard encourages two-way convergence, while placing a sure shipwreck against a legal standard encourages bunching immediately on the compliance side of the boundary line. Part III examines the possibility that the probability of liability within the general standard space might change as a result of the appearance of a safe harbor or sure shipwreck. Part III also considers the possibility that the expected value of different compliance choices (leaving aside the possibility of liability) might shift when a safe harbor or sure shipwreck appears. Part IV raises several questions that inform how policymakers should use safe harbors and sure shipwrecks. It considers over inclusion and under inclusion features, ex ante and ex post considerations, and the vulnerability of safe harbors to interest group influence.

I. BRIGHT-LINE RULES, LEGAL STANDARDS, SAFE HARBORS, SURE SHIPWRECKS

A. Key Terms

Policy makers have different kinds of rule/standard structures in their toolbox. These include not only bright-line rules and legal standards, which are the paradigm building blocks examined in the literature. They also include combinations of rules and standards, such as safe harbors and sure shipwrecks.

A rule constrains a decision maker's discretion ex ante by requiring a particular result to follow from a particular set of facts. True, or bright-line, rules divide conduct between that which complies and that which does not. Bright-line rules include time deadlines for court filings, minimum education prerequisites for professional practice, and the requirement that liquid cosmetics must fit into three-ounce bottles at airport security checkpoints.

A legal standard provides general guidelines. It allows a decision maker discretion to apply these guidelines to a particular situation after the facts of the situation have developed.⁴ An example is the

⁴ See Kaplow, *supra* note 1, at 568-70 (contrasting ex ante costs of promulgating rules with ex post costs of applying standards); Frederick Schauer, *The Tyranny of*

“undue burden” standard developed under Constitutional law that limits state regulation of abortion clinics.⁵ Another example is the common law negligence standard.⁶

The *ex post* feature of legal standards permits answers to be specifically tailored, and presents regulated parties with *ex ante* uncertainty about what the results will be. This Article seeks to contribute to a developing literature that theorizes the behavior of persons subject to irreducible uncertainty presented by legal standards.⁷ Safe harbors and sure shipwrecks provide boundaries for this uncertainty.

A safe harbor guarantees compliance for described behavior, without foreclosing the possibility that activities outside the safe harbor are also compliant.⁸ The activity described by the safe harbor is subject to a rule; other activities are subject to a standard. Election law includes a famous example of a safe harbor, in the provision that guarantees that Congress will accept a state’s list of electors in a Presidential election if the list is submitted by an appointed day in early December.⁹ Safe harbors emerge from case law, statutory law, and regulatory and adjudicatory administrative action. They are particularly visible in administrative action, in part because they are

Choice and the Rulification of Standards, 14 J. CONTEMP. LEGAL ISSUES 803, 804 (2005) (calling “tediously familiar” the distinction between rules, “which reflect choices made by the rule-maker;” and standards, which “leave most of the important choices to be made by the subject, the enforcer, or the interpreter . . . at the moment of application”); Sullivan, *supra* note 1, at 58-59 (drawing similar distinction).

⁵ See *Planned Parenthood of Se. Pa. v. Casey*, 505 U.S. 833, 874-77 (1992).

⁶ See, e.g., 45 U.S.C. §§ 51, 53 (2012) (providing employees with negligence actions); Jones Act, 46 U.S.C. § 30104 (2012) (formerly 46 U.S.C. § 688 (2000)) (extending common law of personal injury in maritime employment context).

⁷ See, e.g., Alex Raskolnikov, *From Deterrence to Compliance: Legal Uncertainty Reexamined* 4-5 (Columbia Law & Econ., Working Paper No. 525, 2016), available at http://papers.ssrn.com/sol3/papers.cfm?abstract_id=2731356 (noting the omnipresence of legal uncertainty and the limited law and economics literature on the topic).

⁸ This definition is consistent with others’ descriptions of safe harbors. See, e.g., Cauble, *supra* note 2, at 1387-88; Morrison, *supra* note 2, at 173; Swire, *supra* note 2, at 370.

⁹ In reversing the Florida Supreme Court’s presidential election recount order, the Court majority put significant weight on the assumption that the Florida legislature meant to take advantage of this election law safe harbor. See *Bush v. Gore*, 531 U.S. 98, 110 (2000) (majority opinion) (articulating assumption based on statement by Florida Supreme Court). *But see id.* at 143 (Ginsburg, J., dissenting) (“[The] safe harbor lacks the significance the Court assigns it.”); *id.* at 149 (Breyer, J., dissenting) (emphasizing that the “safe harbor” statute is not a “mandate”).

often labeled as such. The Federal Register typically reports fifteen or more unique “safe harbor” references every month.¹⁰

Sure shipwrecks¹¹ provide that activity within the sure shipwreck is definitely noncompliant, while activity outside the space described by the sure shipwreck may or may not be compliant. Thus the activity described by the sure shipwreck is subject to a rule, while activities outside the sure shipwreck remain subject to a standard. When the European Union prohibits the sale of fish caught by vessels flagged in certain jurisdictions, it creates a sure shipwreck. The sale of fish caught by other vessels may also be prohibited, depending on the application of the EU’s legal standard applicable to illegal, unreported and unregulated fishing.¹²

B. Accumulation of Safe Harbors and Sure Shipwrecks

Court decisions and regulatory action constantly populate legal standards with networks of safe harbors and sure shipwrecks. For example, the Fourth Amendment provision prohibiting “unreasonable searches and seizures”¹³ is a standard. The series of Supreme Court decisions permitting police officers to conduct warrantless searches of areas close to the location of a stop or arrest — for example, within the “passenger compartment” of a car¹⁴ — are “safe harbor rules” from the perspective of the police officers.¹⁵ In contrast, decisions limiting officers’ warrantless searches produce sure shipwrecks.¹⁶ Outside safe-harbor and sure-shipwreck facts covered by existing case law, a defendant may raise the argument that a search violates the Fourth

¹⁰ For the twelve-month period between December 1, 2012 and November 30, 2013, the number of Federal Register entries referencing safe harbor rules equaled 208. See *The Daily Journal of the United States Government*, FED. REG., www.federalregister.gov (last visited Dec. 1, 2013) (then search for “safe harbor”).

¹¹ I use “sure shipwreck” as the descriptive mirror image of “safe harbor” rather than other terms previously used. See, e.g., Cauble, *supra* note 2, at 1395 (using “dangerous cliff”); Morrison, *supra* note 2, at 174 (using “unsafe harbor”).

¹² See Council Regulation 1005/2008, 2008 O.J. (L 286) 1 (EC).

¹³ U.S. CONST. amend. IV.

¹⁴ *New York v. Belton*, 453 U.S. 454, 462-63 (1981).

¹⁵ Susan R. Klein, *Identifying and (Re)formulating Prophylactic Rules, Safe Harbors, and Incidental Rights in Constitutional Criminal Procedure*, 99 MICH. L. REV. 1030, 1044-47 (2001); see also William J. Stuntz, *Warrants and Fourth Amendment Remedies*, 77 VA. L. REV. 881, 918 (1991) (“In a damages system, warrants can correct overdeterrence problems by providing a safe harbor for police.”).

¹⁶ See *Arizona v. Gant*, 556 U.S. 332, 344 (2009) (limiting *Belton* where police could not reasonably think that detainee could access car at time of search).

Amendment, and a court will decide the matter according to the background Constitutional standard of unreasonableness.

As another example, consider a negligence standard in tort law. A court might decide that a ship owner that employs dockworkers is negligent and liable for injuries resulting from the precarious stacking of lumber on deck.¹⁷ To the extent binding, this precedent creates a sure shipwreck from the perspective of maritime employers. In contrast, a safe harbor can result from case law holding that a ship owner is not negligent if it, say, fails to provide a light switch to illuminate a cargo space with an open hatch.¹⁸

Safe harbors and sure shipwrecks also develop as an administrative agency interprets a legal standard set forth in a statute. For example, a federal tax statute requires a taxpayer to recognize taxable gain if the taxpayer enters into a “constructive sale” transaction. The statute provides that a derivative transaction known as a “short sale against the box,” produces a constructive sale.¹⁹ This is a sure shipwreck. Internal Revenue Service guidance provides that a constructive sale will not result if an owner retains exposure to the value of an underlying asset between 100% and 125% of the asset price at the time of a derivative transaction.²⁰ This is a safe harbor.

Between the sure shipwreck and the safe harbor lie a range of transactions that do not mimic a short sale against the box, and also do not leave the asset’s owner with exposure to the value of the asset between 100% and 125% of the asset price at the time of a derivative transaction. Within this space of remaining uncertainty, a background standard that turns on whether a transaction has “substantially the same effect” as a sale²¹ helps future decision makers determine results.²²

Another federal statute charges the Occupational Safety and Health Administration (“OSHA”) with ensuring “safe and healthful” workplaces.²³ When OSHA concludes that the maximum permissible

¹⁷ Cf. *Beadle v. Spencer*, 298 U.S. 124, 131 (1936) (holding that contributory negligence did not need to be included in jury instructions as a defense to employer negligence in Jones Act-based tort case).

¹⁸ See *Henry v. Moore-McCormack Lines*, 134 F. Supp. 71, 73 (S.D.N.Y. 1955).

¹⁹ See I.R.C. § 1259(c)(1)(A) (2012).

²⁰ See Rev. Rul. 2003-7, 2003-1 C.B. 363 (finding no constructive sale where a shareholder entered into a derivative transaction with respect to a share of stock worth \$20 and the shareholder retained exposure to the value of the stock between \$20 and \$25).

²¹ I.R.C. § 1259(c)(1)(E).

²² See *Anschutz Co. v. Comm’r*, 135 T.C. 78, 111-13 (2010).

²³ 29 U.S.C. § 651(b) (2012).

exposure limit for formaldehyde is two parts per million for a fifteen-minute exposure, it has produced a sure shipwreck.²⁴ Other procedural, information, training and medical attention requirements may produce a violation even if an employer ensures that formaldehyde exposure stays below the legal maximum.²⁵ If an employer complies with all of OSHA's procedure, information, training and medical attention requirements, the employer will fit within a safe harbor and avoid liability for an OSHA violation.

C. The Issue of Precision

Some safe harbors and sure shipwrecks are specific²⁶ and "transparent."²⁷ For example, an individual who spends 500 hours or more in a year actively working in a business activity avoids the often undesirable "passive activity loss" tax rules.²⁸ This safe harbor leaves little doubt as to whether a person falls inside or outside it.

Sometimes the "safe harbor" label is used to describe a less specific regime that requires the application of a multifactor test. Multifactor safe harbors include the Rule 10b5-1²⁹ safe harbors under insider trading law and the Digital Millennium Copyright Act safe harbor insulating internet service providers from liability for the posting of certain online content.³⁰ Sure shipwrecks can also take a less specific, multifactor form, as does the eight-factor general anti-avoidance rule, or GAAR, applicable in Australian tax law.³¹

²⁴ See 29 C.F.R. § 1910.1048 (2015) (setting formaldehyde limits).

²⁵ See *id.* § 1910.1450 (providing standards for the operation of chemical labs).

²⁶ See Ehrlich & Posner, *supra* note 1, at 259 (describing rules and standards as existing along a "specificity-generality axis").

²⁷ See Colin S. Diver, *The Optimal Precision of Administrative Rules*, 93 YALE L.J. 65, 67-69 (1983) (stating that "transparency" means "words with well-defined and universally accepted meanings within the relevant community").

²⁸ Temp. Treas. Reg. § 1.469-5T(a) (2015) (providing 500-hour safe harbor and also "regular, continuous, and substantial" material participation standard).

²⁹ See 17 C.F.R. § 240.10b5-1 (2015) (providing, *inter alia*, an "affirmative defense[]" for the purchase or sale of a security made according to a "written plan" that exists before a person is "aware of . . . information" where the person is not allowed "to exercise any subsequent influence over . . . purchases or sales").

³⁰ See 17 U.S.C. § 512(a) (2012) (insulating service provider against liability if each of five requirements are met, including requirement that another person initiated content transmission and that service provider does not select material).

³¹ The Australian tax general anti-avoidance rule, or GAAR, voids some transactions with "the dominant purpose" of avoiding tax. See *Income Tax Assessment Act 1936* (Cth) s 177A(5) (Austl.). The statute offers an eight-part balancing test to determine the dominant purpose of the transaction. *Id.* s 177D.

Less ex ante specificity or transparency makes a safe harbor or sure shipwreck less rule-like.³² But if there is more ex ante specificity or transparency in the safe harbor or sure shipwreck compared to the adjacent standard space, the regime is precise enough for purposes of the definition of safe harbor or sure shipwreck as used here. Also, the incremental precision offered by a multifactor safe harbor or sure shipwreck may prompt the formation of more specific and transparent market rules, such as compliance plans adopted by private firms.³³ Or the availability of a multifactor safe harbor or sure shipwreck test may facilitate adjudications that begin to populate the space with more specific and transparent examples.³⁴

D. Safe Harbors Give “Good” Results, Sure Shipwrecks Give “Bad” Results

In this discussion of safe harbors and sure shipwrecks, it is assumed that persons subject to a legal regime desire a compliance result. For example, a safe harbor does not merely give assurance that a person fits into category A as opposed to category B, where both categories might provide acceptable answers.³⁵ Rather, a person that falls within

³² Some work uses the background assumption that rules and standards produce the same legal results. That assumption is too strong for purposes of this Article. See Kaplow, *supra* note 1, at 570 (“It should be emphasized that the ‘appropriate’ content is taken to be the same *ex ante* and *ex post*”); see also Ehrlich & Posner, *supra* note 1, at 281 (defining efficient rules to “exclude any rules that . . . list the ‘wrong’ set of conditions and circumstances to which a legal sanction is attached”).

³³ Both the insider trading safe harbor and the Digital Millennium Copyright Act (“DMCA”) safe harbor described above apparently have had this market rule effect. See, e.g., Christopher Glancy & Stefan Mentzer, *Second Circuit Clarifies DMCA Safe Harbor in Viacom v. YouTube*, WHITE & CASE LLP (Apr. 13, 2012), <http://www.whitecase.com/publications/alert/second-circuit-clarifies-dmca-safe-harbor-viacom-v-youtube> (advising online service providers to “examine their internal policies and procedures” following the reported case); *Client Memorandum: Rule 10b5-1 Plans: What You Need to Know*, DAVIS POLK & WARDWELL LLP (Jan. 18, 2013), http://www.davispolk.com/sites/default/files/files/Publication/c0b412f9-d08e-4abf-a327-3f215728160e/Preview/PublicationAttachment/5dbd1bac-15b1-4b37-ae75-4388773478c4/011813_10b5_1.pdf (listing “practical guidelines” for companies who develop “10b5-1 plans”).

³⁴ Many cases have been decided under the multifactor Australian tax general anti-avoidance rule. See, e.g., C. John Taylor, *Form and Substance in Tax Law: Australia*, 87 CAHIERS DE DROIT FISCAL INT’L 95, 114-17 (2002) (giving examples of decided GAAR cases).

³⁵ This distinguishes safe harbors and sure shipwrecks from elections. See, e.g., Heather M. Field, *Choosing Tax: Explicit Elections as an Element of Design in the Federal Tax System*, 47 HARV. J. ON LEGIS. 21, 23 (2010) (analyzing implicit and explicit elections, which extend to taxpayers the right to choose between different tax

a safe harbor obtains the favorable result of compliance, meaning the person will not face liability or a penalty for violating the law. Similarly, a person that fits within a sure shipwreck will face an undesirable noncompliance result.

No broad claim about overall or social welfare is intended by the label of “good” and “bad” results. Instead, these assumptions of a “good” safe harbor result and a “bad” sure shipwreck result refer to the utility of persons subject to the legal regime. So interpreted, these assumptions support the core thesis of this Article, developed below in Part II. In many cases of interest, though not in all, these assumptions are reasonable.

For example, a regulated party that fits within a safe harbor permitting a certain pollutant level will experience avoiding adverse administrative action as a “good” result. Perhaps someone will become ill because of the permitted level of emissions, which is an inextricably linked “bad” result. But it cannot be said that the person who becomes ill is subject to the legal regime; the provision does not target the behavior of the person who becomes ill, but rather seeks to influence directly the behavior of the polluter. So such a pollution provision functions as a safe harbor, and features the key incentive for regulated parties to seek to fit within it, as explored further below.

Sometimes, a rule/standard hybrid produces a good result for some persons subject to the regime and a bad result for other persons subject to the regime. Conclusive evidentiary presumptions,³⁶ such as the historically applicable presumption that a man who lives with his wife is the father of her child,³⁷ produce this result. Such a presumption produces an advantage for one party in litigation and an offsetting and explicit disadvantage for the opposing party in litigation. Such a presumption is a safe harbor from the perspective of one litigant and a sure shipwreck from the perspective of the other litigant.

A rule of contract drafting may also produce a good safe harbor result for some contract parties and a bad sure shipwreck result for others. Consider a rule that enforces an arbitration clause in an employment contract if the clause is worded in a certain way. If an arbitration clause lacks the magic words, it is subject to generally applicable contract drafting standards to determine its enforceability.

treatments).

³⁶ See generally W. Page Keeton, *Statutory Presumptions — Their Constitutionality and Legal Effect*, 10 TEX. L. REV. 34, 34 (1931) (“In many instances rules of law have originated by means of the adoption of the fiction of a conclusive presumption.”).

³⁷ See *Michael H. v. Gerald D.*, 491 U.S. 110, 117, 131-32 (1988), *reh’g denied*, 492 U.S. 937 (1989) (upholding conclusive presumption of paternity under California law).

This arbitration clause enforcement regime, let us assume, systematically advantages employers and systematically disadvantages employees.³⁸ It is a safe harbor from the perspective of an employer and a sure shipwreck from the perspective of an employee.

When a rule/standard hybrid produces a good result for some persons subject to the regime and a bad result for other persons subject to the regime, it may function as a safe harbor or as a sure shipwreck, depending on the relative power of the different groups.³⁹ Consider the rule enforcing an employment contract arbitration clause written in a certain way. If the employer presents contracts to hourly seasonal workers not represented by a union, the employer will likely act as if the arbitration clause enforcement provision is a safe harbor, and the employees will not resist. But if an employer negotiates with a highly-sought-after CEO, and the CEO views the rule as a sure shipwreck, the CEO may have the negotiating power to ensure that the employment contract is missing the arbitration clause with the magic words, because that would be a sure shipwreck from the CEO's perspective.

II. SAFE HARBORS AND SURE SHIPWRECKS AFFECT BEHAVIOR INCENTIVES

A. *Convergence for Safe Harbors; Bunching for Sure Shipwrecks*

The goal of this Article is to theorize the boundaries between a general standard space and the safe-harbor and sure-shipwreck building blocks that limit it. The key takeaway is that persons subject to a safe harbor converge on the line drawn by the safe harbor both from the noncompliance side of the line and from the compliance space within the safe harbor. In contrast, sure shipwrecks encourage persons who would otherwise find themselves within the sure shipwreck to bunch instead immediately on the compliant side of the sure shipwreck line. Sure shipwrecks do not necessarily influence the behavior of persons initially on the compliant side of the line.

³⁸ See Myriam Gilles & Gary Friedman, *After Class: Aggregate Litigation in the Wake of AT&T Mobility v. Concepcion*, 79 U. CHI. L. REV. 623, 631 (2012) (arguing that class action waivers embedded in arbitration provisions, if upheld, would practically foreclose aggregate litigation in contractual relationships including employment).

³⁹ Cf. Steven Shavell, *Strict Liability Versus Negligence*, 9 J. LEG. STUD. 1, 2-9 (1980) (arguing that the appropriate use of strict liability and negligence rules depends in part on what party has the capacity to avoid an accident).

This Part II.A presents a drunk driving example and a lobbyist-legislator transfer example to illustrate this point about asymmetric convergence. One key assumption underlying this initial analysis is that the expected value of different compliance choices (leaving aside the possibility of liability) will not shift when a safe harbor and/or sure shipwreck appears. Another key assumption is that the probability of liability for the behavior that lies outside the safe harbor and/or sure shipwreck will not change when a safe harbor and/or sure shipwreck appears. These assumptions of exogeneity are relaxed in Part III. Part III acknowledges that the expected value of choices and the probability of liability in the general standard space could shift as a result of the appearance of a safe harbor and/or sure shipwreck.

1. Drunk Driving Example.

Consider a misdemeanor law, drafted as a standard, that prohibits driving if the driver is “intoxicated”⁴⁰ or “under the influence of alcohol.”⁴¹ Drivers who estimate the likelihood of their liability under this standard will estimate a range of probabilities that increases with more drinking.⁴² Perhaps the range starts at 0% and increases to 100%. Perhaps the standard produces a 10% probability of liability if a person drives after having one drink and a 90% probability of liability if a person drives after having four drinks.

Assume that a safe harbor rule appears and guarantees no liability under this law if a person drives after one drink. Some individuals who used to abstain from drinking before driving⁴³ will change their behavior and have one drink before driving. This is because the reduction from a 10% probability of liability to a 0% probability of liability for one drink under the safe harbor makes the one-drink

⁴⁰ E.g., TEX. PENAL CODE ANN. § 49.01(2) (West 2015).

⁴¹ E.g., CAL. VEH. CODE § 23152 (West 2016).

⁴² In other words, drivers estimate a function that has increasing probability of liability with increasing noncompliance under a legal standard. The function could take different shapes, though it is important for purposes of this Article that the probability of liability strictly increases with increased noncompliance for the regulatory space affected by the safe harbor and/or sure shipwreck regime. Figure 1 shows a straight-line function, but this is not meant to constrain the possible shape of the function. For example, it might have an S-curve shape. See Alex Raskolnikov, *From Deterrence to Compliance: Legal Uncertainty Reexamined*, *supra* note 7, at 15-16 (proposing an “S-shaped probability of success curve”); see also Richard Craswell & John E. Calfee, *Deterrence and Uncertain Legal Standards*, 2 J.L. ECON. & ORG. 279, 285 (1986) (The “standard deviation can be interpreted as a measure of the amount of uncertainty in the legal system.”).

⁴³ See, e.g., *infra* Figure 1 (data point “A”).

option more attractive than it was before the appearance of the safe harbor. In addition, some drivers who used to drink two drinks⁴⁴ will change their behavior and only have one drink, also because the one-drink alternative with a 0% probability of liability is more attractive than the one-drink alternative with a 10% probability of liability. The safe harbor thus encourages behavior convergence from both sides of the line.

Alternatively or in addition, a rule might appear that automatically provides the result of a misdemeanor conviction if a person drives after four drinks. This is a sure shipwreck. Some persons who used to drink four drinks before driving⁴⁵ will no longer do so, because the option of drinking four drinks with a 100% probability of liability is less attractive than the option of drinking four drinks with a 90% probability of liability.

But persons accustomed to drinking, say, three drinks before driving⁴⁶ will have no incentive to drink more. This is because the sure shipwreck makes the four-drink option less attractive, not more attractive, for the driver. The sure shipwreck does not, in this first order analysis, encourage convergence. Rather, it encourages the bunching of behavior on the compliant side of the line only as a result of drivers' decisions to shift out of the noncompliance space and into the possibly compliant space.

2. Figure 1

Figure 1, below, shows this insight graphically. The x axis of the graph, marked from 0 to 7, represents an index of increasing noncompliance, where 0 is maximum compliance and 7 is a high level of noncompliance. The graph illustrates the generally increasing chance of liability for increasingly noncompliant behavior under the general standard.⁴⁷ Under the liability regime presented in the figure, when only the background general standard is in place, there is a chance of liability that starts at zero for fully compliant behavior and increases to 100% for definitely noncompliant behavior.

Against the background of the dashed background general standard line, the safe harbor and sure shipwreck appear as solid lines, for

⁴⁴ See, e.g., *infra* Figure 1 (data point "B").

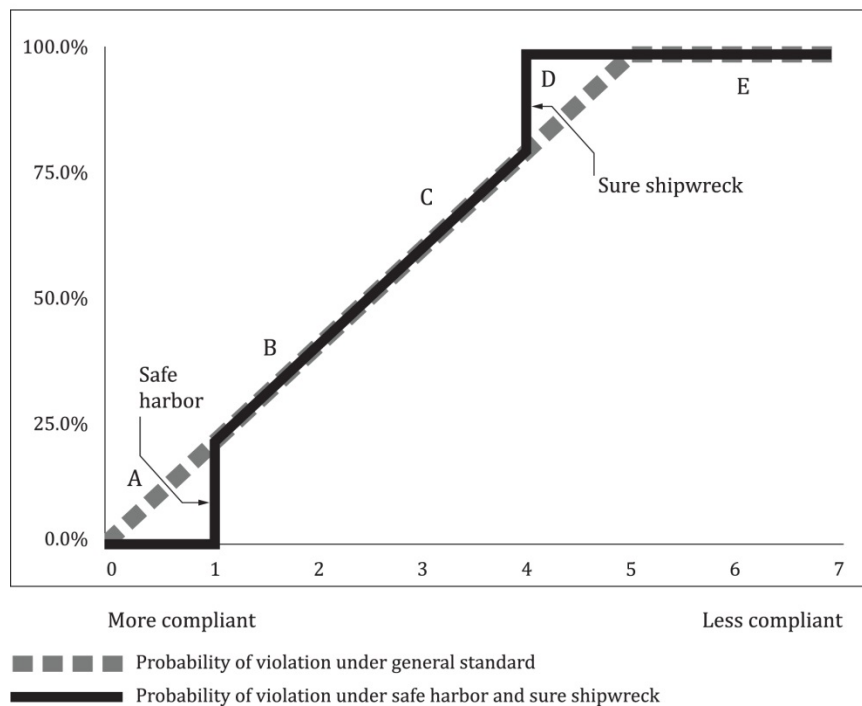
⁴⁵ See, e.g., *infra* Figure 1 (data point "D").

⁴⁶ See, e.g., *infra* Figure 1 (data point "C").

⁴⁷ This function could take several forms and is limited only by the requirement that it increases strictly, subject to the possibility that it will reach a maximum. Here it is presented as a straight-line function and shown as a dashed line. See *supra* note 42.

example at one drink and four drinks, respectively. These change the probability of liability for behavior inside the safe harbor and sure shipwreck. In this first-order presentation of the problem in Part II, the appearance of the safe harbor and sure shipwreck do not change the probability of liability outside the safe harbor and sure shipwreck, in other words in the remaining general standard space. Thus the appearance of the safe harbor and sure shipwreck result in discontinuities in the probability of liability at the boundaries drawn by the safe harbor and sure shipwreck.

FIGURE 1: PROBABILITY OF VIOLATION UNDER SAFE HARBOR AND SURE SHIPWRECK



On the Figure 1 graph, A marks a person who clearly fit within the safe harbor before the appearance of the safe harbor. B marks a person who was somewhat less compliant than required in order to fit within the safe harbor. The argument is that the safe harbor encourages both A and B to converge on the line drawn by the safe harbor.

Figure 1 also shows C, a person somewhat more compliant than the line drawn by the sure shipwreck; and D, a person clearly within the

sure shipwreck although not very far away from the line drawn by the sure shipwreck. The argument is that the sure shipwreck encourages D to move out of the sure shipwreck, but without more,⁴⁸ does not encourage a change in the behavior of C.

Finally, Figure 1 shows E, a person whose noncompliance places them very far into the sure shipwreck. E is at a place where a violation of the law clearly occurs under either the initial general standard legal regime or under the safe harbor and sure shipwreck. E's behavior is not affected by the appearance of the sure shipwreck, because E clearly violates the law both before and after the sure shipwreck becomes law.

Figure 1 does not show the expected value (aside from the probability of liability) of choices along the noncompliance spectrum for each of A, B, C, D and E. Many factors might influence these choices, because different persons have different preferences, including different personal tastes and different algorithms for profit maximization. The analysis in this Article accommodates many different expected value functions, without having to specify them.⁴⁹ The analysis takes the persons subject to the legal regime as it finds them, with the understanding that each person might have arrived at an optimal compliance point along the continuum in the absence of a safe harbor or sure shipwreck for a variety of reasons.

To state the point somewhat differently, Figure 1 raises the question of how each person subject to the legal regime (such as A, B, C, D or E) came to decide on different compliance or noncompliance behavior in the absence of the safe harbor or sure shipwreck. The answer is that each person considered various benefits and costs of different levels of compliance, such as levels of drinking, under the general standard, and chose a level of compliance that maximized that person's net benefit. The possibility of a legal violation is but one factor that would influence such a decision. Yet even though the possibility of a legal violation is but one factor that would influence a person's decision about the level of compliance, it is the only factor shown in Figure 1.

⁴⁸ See *infra* Part III.

⁴⁹ The expected value or utility functions of different persons may be not only different, but also differently shaped. These functions need not strictly increase with noncompliance; indeed, alcohol consumption is a good example of an activity for which utility does not strictly increase with more of the activity. Finally, these utility functions might be smooth, or might be discontinuous and lumpy. This is a more textured view of utility functions than is sometimes used in other literature. See generally Benjamin E. Hermalin, Avery W. Katz & Richard Craswell, *Contract Law*, in 1 HANDBOOK OF LAW AND ECONOMICS 3, 22 (A. Mitchell Polinsky & Steven Shavell eds., 2007) (describing individual utility curves and the social welfare function that results from aggregating such individual curves).

This is consistent with the first-order argument presented in this Part, which assumes that the only decision factor that changes when a safe harbor and/or sure shipwreck appears is the possibility that conduct described by the safe harbor and/or sure shipwreck will violate the law. Thus Figure 1 supports the argument's initial focus on how the change in the possibility of liability alone influences behavior.

3. Lobbyist-Legislator Transfer Example

To further illustrate the core argument that a safe harbor encourages behavior convergence but a sure shipwreck does not, consider a law, drafted as a standard, that prohibits a transfer from a lobbyist to a legislator if the transfer carries an inference of intent to influence.⁵⁰ A "cup-of-coffee" safe harbor under this law might provide that transfers of food or drink worth \$10 or less will definitely not violate the underlying standard. The cup-of-coffee safe harbor leaves open for determination under the standard the question of whether food or drink transfers worth more than \$10 violate the law.

A "fancy dinner" sure shipwreck under the background standard might provide that transfers of food or drink worth more than \$100 definitely violate the law. Under this sure shipwreck, transfers of food or drink worth \$100 or less may or may not violate the law. Future decision makers will determine results for food or drink transfers worth \$100 or less by applying the background standard.

Consider the impact of the \$10 cup-of-coffee safe-harbor rule on two lobbyists, A and B. A spent \$5 per legislator cup of coffee before the imposition of the safe harbor; B spent \$15. Before the safe harbor, the \$5 cup of coffee was the optimal point for Lobbyist A and the \$15 cup was the optimal point for Lobbyist B. The safe harbor provides an incentive for both lobbyists to choose \$10 cups of coffee instead.

Imagine that A chose a \$5 cup of coffee before the appearance of the safe harbor because, although a \$10 cup would produce slightly more value for A without considering possible penalties, A was concerned that a \$10 cup would produce the noncompliance result of inference of intent to influence. After a \$10 safe harbor is put into place, A faces the same possible penalty for a \$5 or \$10 cup — \$0. As a result, the slightly higher value of a \$10 cup of coffee encourages A to switch to the \$10 cup.

Note that not every \$5-per-cup lobbyist will switch to \$10 cups. Some prefer — aside from any possible penalty — cheap coffee at run-down diners. It is those lobbyists that prefer \$10 cups to \$5 cups in

⁵⁰ Cf. 2011 N.Y. Sess. Laws § 14-126 (McKinney).

the absence of any penalty considerations who drive this “going up to the line” tendency when a safe harbor is imposed. The analysis is marginal.

Lobbyist B illustrates the convergence tendency from the other side of the safe harbor line. B paid \$15 per cup of coffee before the imposition of the safe harbor. In other words, the difference between the value of the coffee for B and the probability-adjusted fine for noncompliance was at a maximum at \$15. The appearance of the safe harbor changes B’s calculus. Because the possibility of a fine for noncompliance for a \$10 cup of coffee is now zero, B must now compare her previous optimal value for coffee with a legislator — equal to the difference between the value of \$15 coffee and the probability-adjusted fine for noncompliance at \$15 — to the value of \$10 coffee, unadjusted for any fine, since a fine will definitely not be imposed at \$10. Lobbyist B will switch if the benefit of definitely avoiding the possibility of liability exceeds the loss in value from having a \$10 cup of coffee rather than a \$15 cup of coffee.

Next, consider the results of the imposition of a fancy-dinner sure shipwreck providing that transfers of food or drink costing over \$100 will violate the underlying standard. Take two more lobbyists, C and D. C spent \$90 on dinner before the sure shipwreck; D spent \$110. Before the sure shipwreck, the \$90 dinner was the optimal point for Lobbyist C and the \$110 dinner was the optimal point for Lobbyist D. The sure shipwreck encourages Lobbyist D, but not Lobbyist C, to choose a \$100 dinner instead.

Lobbyist C will not face an incentive to increase the price of the \$90 meal he provides to a legislator merely because of the appearance of the sure shipwreck.⁵¹ This follows from the assumption that the sure shipwreck does not change the attractiveness of different compliance choices outside its bounds. If the same legal standard applies to the \$90 meal before and after the appearance of the safe harbor, Lobbyist C will continue to conclude that purchasing the \$90 meal is the optimal course of action. The sure shipwreck only makes choices costing over \$100 less attractive.

In contrast, consider Lobbyist D. Assume that, prior to the imposition of the \$100 sure shipwreck, D spent \$110 on dinner and that the probability of violating the standard with a \$110 dinner was less than 100%. After the sure shipwreck, the probability of the

⁵¹ Note however, that Lobbyist C’s behavior could change as a result of second-order effects. For example, the bunching of other lobbyists’ per-plate spending at \$100 may facilitate a market coordination effect, which could induce Lobbyist C to also spend \$100. *See infra* Part III.A.

imposition of a fine for a \$110 dinner is equal to 100%. This changes D's calculus. D faces a greater incentive to choose a \$100 meal, because the probability of the imposition of a fine for a \$100 dinner is less than 100%.

Finally, consider Lobbyist E, who typically spends \$500 on dinner with a legislator. Lobbyist E, let us assume, was certain in the absence of any sure shipwreck that \$500 dinners violated the inference of intent to influence standard. E accepted that penalties would apply; and even considering the penalties that would apply, E concluded that \$500 dinners produced the maximum value for E. The appearance of a sure shipwreck for dinners costing more than \$100 will not change E's behavior. Lobbyist E illustrates that persons at noncompliance extremes who face the same penalties with or without a sure shipwreck face no incentive to change their behavior as a result of the sure shipwreck.⁵²

B. *The Discontinuity or "Notch" Feature of Safe Harbors and Sure Shipwrecks*

The results of behavior convergence for safe harbors and behavior bunching for sure shipwrecks follow from a key premise about the structure of the relevant legal regime. This premise is that, as shown in Figure 1, a discontinuity exists at the boundary of the safe harbor or sure shipwreck. The discontinuity means that the chance of a "good" (or "bad") result will all of a sudden jump, and become greater in a non-smooth fashion, if a person subject to the rule is just inside the boundary of the safe harbor (or sure shipwreck) rather than just outside that boundary.

Discontinuities in legal regimes have been previously studied in at least three respects. First, paradigm bright-line rules have been studied. Second, "notches" have been examined. Third, the question

⁵² Lobbyist E's incentives also depend on whether there is a maximum penalty or an increasing penalty for increasing noncompliant behavior. An increasing penalty presents a greater likelihood that Lobbyist E will be appropriately deterred from noncompliant behavior. Similarly, a sure shipwreck that increases the probability of the imposition of an increasing penalty is more likely to prompt Lobbyist E to move out of the noncompliant sure shipwreck space and into the adjacent standard space. See generally Mark P. Gergen, *Uncertainty and Tax Enforcement: A Case for Moderate Fault-Based Penalties*, 64 TAX L. REV. 453, 457-59 (2011) (arguing that a penalty rate equal to the inverse of the probability of detection is not enough to deter a risk-neutral taxpayer when the correct result is uncertain); A. Mitchell Polinsky & Steven Shavell, *The Theory of Public Enforcement of Law*, in 1 HANDBOOK OF LAW AND ECONOMICS 403, 408 (A. Mitchell Polinsky & Steven Shavell eds., 2007) (explaining that fines should equal harm caused to produce appropriate cost internalization and deterrence).

of whether a discontinuity exists at the boundary of an uncertain space — which is the core animating idea of this paper — is also engaged in contemporaneous work.

A paradigm bright-line rule that gives a compliant result on one side of a line and a noncompliant result on the other side of a line provides a clear and important example of a discontinuity. Others have observed that such a bright-line rule produces the effect of convergence.⁵³ To illustrate the convergence effect of a bright-line rule, consider a continuation of the lobbying expense example. Assume a bright-line rule that permits transfers of \$100 or less, and prohibits transfers in excess of \$100.⁵⁴ This rule encourages both Lobbyist C (who previously bought \$90 dinners) and Lobbyist D (who previously bought \$110 dinners) to buy \$100 dinners instead. This result contrasts with the sure shipwreck, which provided no incentive for the \$90 dinner lobbyists to increase their spending, under the assumption that the sure shipwreck did not affect the probability of liability for dinners costing up to \$100. In other words, a “going up to the line” incentive exists for bright-line rules (as well as safe harbors) but not for sure shipwrecks.⁵⁵

The existence of “notches,” or “discontinuous jumps” “in the choice set of individuals and firms” has also been examined. An example is a tax, such as an excise tax on cigarettes that is higher in one state than in another. Such a notch is predicted to produce a bunching effect, in the form of “a mass of purchases just on the low-tax side of the border.”⁵⁶ A similar analysis motivates the description here of the behavioral incentives presented by sure shipwrecks.

The motivating assumption and intuition for the core analysis in this Article is that discontinuities, or notches, generally exist at the boundaries of uncertain legal spaces, where safe harbors and sure shipwrecks cordon off compliant and noncompliant behavior. It is also

⁵³ See Craswell & Calfee, *supra* note 42, at 285 (“The standard deviation can be interpreted as a measure of the amount of uncertainty in the legal system. . . . [T]he effect of a shift in the mean is much greater when the standard deviation is small than when it is large.”); David A. Weisbach, *An Efficiency Analysis of Line Drawing in the Tax Law*, 29 J. LEGAL. STUD. 71, 76-77 (2000) (analyzing the deadweight loss produced by the shifting of behavior to the more favorable side of a line).

⁵⁴ This rule draws the same lines as the fancy-dinner sure shipwreck discussed above, but does so in a bright-line fashion. See *supra* Part II.A.

⁵⁵ Second-order effects, such as market coordination, may cause a sure shipwreck to behave more like a bright-line rule. See *infra* Part III.A.

⁵⁶ Joel Slemrod, *Buenas Notches: Lines and Notches in Tax System Design*, 11 J. TAX RES. 259, 264 (2013) (noting that the consumer’s decision depends in part on “whether the transportation costs exceed the savings from the tax differential”).

possible to reason that, in general, discontinuities do not exist at the boundaries of uncertain legal spaces.⁵⁷ Part III details some of the considerations that are relevant to the question of whether the appearance of a safe harbor or sure shipwreck will result in a discontinuity.

If there is no discontinuity, then one might conclude that a safe harbor provision that cordons off certain behavior and labels it compliant only affects the cordoned off, compliant behavior.⁵⁸ But available examples suggest the acceptance of the idea that a discontinuity exists at a line drawn by a safe harbor, even in unlikely cases. One famous example that illustrates this point is the *Bush v. Gore* litigation that decided the 2000 presidential election. It turned in part on an election law safe harbor that provides that Congress will accept a state's list of electors if the state turns that list in by an appointed date in early December. The law is silent about the result if elector names are provided after the safe harbor date, and there is little reason to think that Congress would have refused to accept a late list from Florida that reflected a more complete recount in that election. Yet the Supreme Court majority put great weight on an assumption that the Florida legislature meant to take advantage of the safe harbor and therefore that the Florida legislature had properly declined to extend the recount effort. Whether or not that assumption was correct or wise is certainly open to debate. Here, I mean only to suggest that the very plausibility of the idea that the Florida legislature desired the protection of the safe harbor is evidence that safe harbors can have influence beyond the boundaries of the compliant behavior they describe, through a discontinuity at the safe harbor boundary that encourages behavior convergence.⁵⁹

Table 1 summarizes the core analysis of Parts II.A and II.B. It compares the incentives of persons subject to a legal regime when a safe harbor, sure shipwreck or bright-line rule appears against a background legal standard.

⁵⁷ Contemporaneous work describes this possibility. See Raskolnikov, *From Deterrence to Compliance: Legal Uncertainty Reexamined*, *supra* note 7, at 15.

⁵⁸ Considerations such as uncertainty aversion provide a separate reason, in addition to a discontinuity in the probability of liability, for why behavior might converge at a safe harbor line. See *infra* note 79.

⁵⁹ See *Bush v. Palm Beach Cnty. Canvassing Bd.*, 531 U.S. 70, 78 (2000) (labeling a "safe harbor" the statute that guaranteed the acceptance of Florida's elector list by December 12, 2000) (discussing 3 U.S.C. § 5 (2012)); see also *supra* note 9 (outlining the Supreme Court's treatment of the safe harbor provision in *Bush v. Gore*, 531 U.S. 98 (2000)).

Table 1: Safe Harbor, Sure Shipwreck, Bright-Line Rule Incentives

	Safe harbor	Sure shipwreck	Bright-line rule
Incentive to decrease compliance for persons initially on compliant side of line	Yes “Going up to the line” incentive	No Assuming unchanged calculus in adjacent standard space	Yes “Going up to the line” incentive
Incentive to increase compliance for persons initially on noncompliant side of line	Yes Incentive to comply if initially outside safe harbor because safe harbor offers lower chance of liability	Yes Incentive to comply because sure shipwreck characterizes earlier behavior as noncompliant	Yes Incentive to comply because bright-line rule characterizes earlier behavior as noncompliant

C. Probability of Liability Could Be Greater than 0% for Safe Harbor, Less than 100% for Sure Shipwreck

The discussion in Part II.A presents two safe harbor and sure shipwreck examples based on legal regimes that apply to drunk driving and lobbying expenditures. In both examples, the hypothetical safe harbors reduced the probability of liability to 0%, and the hypothetical sure shipwrecks increased the probability of liability to 100%. However, there is nothing magic about the 0%, or 100%, liability probability.

It is the reduction in the probability of liability when a safe harbor appears that produces a discontinuity and drives convergence from both sides of a safe harbor. And it is the increase in the probability of liability when a sure shipwreck appears that produces a discontinuity and drives bunching from the noncompliant side of a sure-shipwreck boundary. In fact, the actual probability of liability for most safe harbors, and especially for sure shipwrecks, does differ from 0%, or 100%, respectively.

A person that fits within a safe harbor may nevertheless face liability because of the application of an anti-abuse rule. An anti-abuse rule

applies explicitly, for example, in the case of the Rule 10b-1 insider trading safe harbor.⁶⁰ Imperfect enforcement may also incorrectly categorize safe harbor behavior as illegal, which imposes costs on the person subject to the safe harbor even if liability does not result in the final analysis.

For sure shipwrecks, a variety of imperfections in an enforcement regime will generally produce a less-than-100% probability of liability. These imperfections include imperfect detection, enforcement discretion, and difficulty of proof. Many drivers who drive after four drinks make it home without getting stopped, let alone convicted.⁶¹

For both safe harbors and sure shipwrecks, the essential feature is that the probability of liability is less (for safe harbors) or more (for sure shipwrecks) than it would be without the safe harbor or sure shipwreck. If driving after one drink has a probability of liability of 1% without a safe harbor and 0.05% with a safe harbor, the one-drink option will be more attractive, on the margin, both to persons who otherwise would abstain and to persons who otherwise would have two drinks. If driving after four drinks has a probability of liability of 1% without a sure shipwreck and 2% with a sure shipwreck, then on the margin the sure shipwreck will encourage drivers who otherwise would drink four drinks to drink less.

D. Who Are the “Persons Subject to” a Safe Harbor or Sure Shipwreck?

When a safe harbor or sure shipwreck appears, only some of the persons subject to the regime will change their behavior in response. The strength of the core thesis of the Article — which predicts convergence at safe-harbor boundaries and bunching immediately on the compliance side of sure-shipwreck boundaries — depends in a particular case on the number and identity of persons who change their behavior. To determine whether to change behavior, a person would consider the value of the compliance option right at the safe-harbor or sure-shipwreck boundary compared to the compliance option the person had chosen before the appearance of the safe harbor or sure shipwreck.

⁶⁰ See 17 C.F.R. § 240.10b5-1(c)(1)(ii) (2016) (providing that “affirmative defense” applicable only if used “in good faith and not as part of a plan or scheme to evade the prohibitions of this section”).

⁶¹ See, e.g., Polinsky & Shavell, *The Theory of Public Enforcement of Law*, *supra* note 52, at 421-23 (noting general problem of low probability of detection and citing evidence about the low probability of enforcement of drunk driving law).

A utility curve would capture a person's different expected values (aside from the possibility of liability) for different compliance and noncompliance options.⁶² For example, a person for whom expected value increases for more noncompliant choices is a person who might "go up to the line" drawn by a safe harbor. Also, a person relatively close to a safe-harbor or sure-shipwreck boundary probably faces less reduction in utility as a result of adopting more compliant behavior so as to fit within a safe harbor or avoid fitting within a sure shipwreck.⁶³

Consider again the one-drink safe harbor that protects against a drunk driving offense. Drivers who prefer one drink but drink less because of concerns about drunk driving liability in the absence of the safe harbor make up the relevant population for purposes of the going-up-to-the-line incentive to drink one drink rather than zero before driving. There is no practical going-up-to-the-line incentive for an individual who abstains for medical or religious reasons, because such an individual does not place a higher value on having one drink compared to having zero drinks.

To further illustrate this point about the distribution of persons subject to a safe harbor or sure shipwreck, consider another safe harbor, which exempts an individual taxpayer from the disadvantageous passive activity loss tax rules if the taxpayer spends 500 hours a year on a business activity.⁶⁴ A taxpayer who chooses 495 hours as the optimal amount of time to spend on a business activity in the absence of the safe harbor will likely increase her hours to 500 to take advantage of the safe harbor. This is because she needs to make only minor changes to her schedule to fit within it, and she likely assigns similar values to working 495 hours per year on the business activity and working 500 hours per year on the business activity. Taxpayers who before the safe harbor spend less time — say 450 hours, or 400, or 350 — will be increasingly less likely to change their behavior to fit within the safe harbor.⁶⁵

⁶² See *supra* note 49 (explaining Article's assumption about different persons' different expected value functions).

⁶³ See David A. Weisbach, *Line Drawing, Doctrine, and Efficiency in the Tax Law*, 84 CORNELL L. REV. 1627, 1661 (1999) (noting that if similar behavior is treated differently, a large incentive for taxpayers to change their behavior to the favorable side of a rule will result).

⁶⁴ See Temp. Treas. Reg. § 1.469-5T(a) (2015) (providing 500-hour safe harbor and also "regular, continuous, and substantial" material participation standard).

⁶⁵ See Cauble, *supra* note 2, at 1420 (arguing that, compared to a standard, a safe harbor "exacerbates distortions for taxpayers whose non-tax preference lie close to but outside of the bounds of the safe harbor").

Consider finally a new example of a sure shipwreck. This hypothetical sure shipwreck requires the provision of health insurance to janitors who work 40 hours a week or more (and leaves open the question of whether such health insurance must be provided for janitors working less than 40 hours per week). This might well cause employers to reduce janitors' hours to less than 40 hours per week, because employers might value the services of two part-time janitors at not much less than the services of one full-time janitor. In contrast, a sure shipwreck that requires the provision of health insurance to medical residents who work 40 hours a week or more (and leaves open whether medical residents who work less are treated as students or employees)⁶⁶ will not encourage many hospitals to reduce residents' hours so as to fall outside the sure shipwreck. This is because most residents work far in excess of 40 hours per week, and because the long workweeks of medical residents are built into the economics and culture of health care. The hospitals highly value the residents' 80-hour weeks, in other words, and 40-hour weeks are not an acceptable substitute.

The example of janitors and residents in a hospital emphasizes that there are many reasons that may cause persons subject to a legal regime to choose their level of compliance or noncompliance behavior. In this case, personal, labor market, socioeconomic and other factors result in the different work schedules of janitors and medical residents at hospitals. The possibility that the hospital might be required to provide health insurance for different workweek lengths is only one of many factors considered by the hospital when setting work schedules for its employees.

Under the first-order analysis in this Part II, the possibility of a legal violation within the safe harbor and/or sure shipwreck is the only factor that changes when a safe harbor or sure shipwreck appears. So far, the analysis overlooks the possibility that other factors that affect compliance decisions may change as a result of the appearance of a safe harbor and/or sure shipwreck. Perhaps the appearance of the safe harbor and/or sure shipwreck affects factors that are not directly related to the possibility of a violation. And certainly the appearance of a safe harbor or sure shipwreck may affect the possibility of a violation in the general standard space. Part III considers this endogeneity.

⁶⁶ Cf. *Mayo Found. for Med. Educ. & Research v. United States*, 562 U.S. 44, 58-60 (2011) (citing *Treas. Reg. § 31.3121(b)(10)-2(d)(3)(i)*) (deferring to a Treasury Regulation that treats an individual who works 40 hours or more per week as an employee, not a student; regulation applies a standard on other facts).

III. FACTORS ENDOGENOUS TO SAFE HARBOR AND/OR SURE SHIPWRECK

A. *Market Coordination*

What if the adoption of a safe harbor or sure shipwreck affects behavior because of market coordination? Here it is useful to compare the two main examples presented in Part II.A. In the first example, the legal regime regulated drunk driving. In this example, it is reasonably plausible to think that the different drivers respond independently to changes in the law.

But in the other example, relating to the rules applicable to entertainment of legislators by lobbyists, it is not plausible to assume that the lobbyists do not know each other. The project of networking is fundamental to the business of lobbying. And, in particular, we should assume that Lobbyist C, who paid \$90 per legislator dinner before the appearance of a \$100 fancy-dinner sure shipwreck, knows Lobbyist D, who paid \$110 per dinner before the sure shipwreck appeared. Moreover, we should assume that there are many Lobbyists C and also many Lobbyists D.

The first-order argument presented in Part II explained that only Lobbyists D had an incentive to shift behavior and pay \$100, not \$110, per dinner after the appearance of the sure shipwreck. The predicted result was a clustering of Lobbyists D at the \$100 dinner mark. But in the case of lobbyists and legislators, the first-order clustering of Lobbyists D at the \$100 dinner mark is not the end of the story.

Instead, in the case of lobbyists and legislators, the changed behavior of Lobbyists D could have a secondary impact and prompt a change in the behavior of other lobbyists. For example, if Lobbyists C and legislators observe the clustering of Lobbyists D at the \$100 dinner mark, Lobbyists C will be more likely to offer \$100 dinners, and legislators will be more likely to expect \$100 dinners. Restaurant owners might further the project of market coordination by offering \$100 dinner packages.⁶⁷

If a \$100 dinner market norm develops, then the expected value functions of Lobbyists C could change.⁶⁸ However Lobbyist C

⁶⁷ A separate issue is presented by the government's willingness to enforce the general standard at the \$100 mark. Perhaps the government will focus on the behavior prohibited by the sure shipwreck, and its silence as to the widespread practice of \$100 dinners will be taken to mean that the regulator does not object and will not enforce. But other enforcement strategies are also possible.

⁶⁸ Recall that Figure 1 does not show the expected value of different compliance and noncompliance choices (aside from the possibility of liability). So this change is not something easily shown with the Figure 1 graph.

balanced the pros and cons of a \$100 dinner before the sure shipwreck, Lobbyist C now likely adds another advantage of a \$100 dinner to the mix. This advantage is that a \$100 dinner will meet the market expectation of the legislator. As a result, Lobbyist C's behavior might change. Lobbyist C might increase spending on each legislator dinner from \$90 to \$100. Thus the second-order effect of market coordination might in some cases produce a convergence of behavior at the line drawn by the sure shipwreck,⁶⁹ even though the first-order analysis described in Part II only predicted bunching as a result of movement out of the sure shipwreck's zone of noncompliance.⁷⁰

B. Probability of Violation in General Standard Space May Increase, Decrease or Stay the Same

The discussion in Part II assumes that the appearance of a safe harbor or certain shipwreck only affects the probability of a violation within the boundaries of a safe harbor or sure shipwreck. In contrast, this Part III acknowledges that when a safe harbor or sure shipwreck appears, it could produce a change in the probability of a violation outside its boundaries, in other words in the general standard space.⁷¹ This could result in no discontinuity at the boundary of the safe harbor or sure shipwreck. It could also result in a greater discontinuity, one that more closely resembles the discontinuity of a bright-line rule, at the boundary of the safe harbor or sure shipwreck space.

The direction of a likely shift in the probability of a violation outside the boundaries of a safe harbor and/or sure shipwreck cannot be

⁶⁹ Cf. Raskolnikov, *From Deterrence to Compliance: Legal Uncertainty Reexamined*, *supra* note 7, at 8, 32 (arguing that the market for legal advice may foster increased noncompliance in the presence of legal uncertainty through the gathering of advice around increasingly aggressive positions).

⁷⁰ The possibility that a sure shipwreck might encourage convergence has implications for the regulation of bad behavior. See generally Alex Raskolnikov, *Irredeemably Inefficient Acts: A Threat to Markets, Firms, and the Fisc*, 102 *GEO. L.J.* 1133, 1143-50 (2014) (theorizing acts that are always inefficient). If sure shipwrecks only affect behavior on the noncompliant side of the line, as suggested by the first-order analysis in Part II, they may be ideal instruments to use when the main policy goal is to encourage less of a certain behavior, such as illegal fishing, drunk driving, or tax avoidance. But if a sure shipwreck has some possibility of morphing into a market rule and therefore including a going-up-to-the-line incentive, its use for the regulation of bad behavior becomes more complicated, and more contingent on the specific market at issue and on enforcement in the space of uncertainty adjacent to the sure shipwreck. (Safe harbors may also be useful in encouraging noncompliant types to increase compliance to the safe harbor level, but safe harbors also have the going-up-to-the-line incentive, which works against the goal of discouraging a behavior.)

⁷¹ See *supra* text accompanying note 57.

predicted as a general matter. Rather it is contingent on situation-specific factors. Factors that influence the discontinuity at the boundary of a safe harbor or sure shipwreck include: the framing given by the policy maker; the uniqueness of the facts within the safe harbor or sure shipwreck; the use of lists; the comparison between the safe harbor or sure shipwreck and previous beliefs about the content of the law; and whether the person interpreting the safe harbor or sure shipwreck considers the subtleties of the regime. Finally, it is important that many legal spaces include both safe harbors and sure shipwrecks. Where both safe harbors and sure shipwrecks exist in the same legal space, it is not logically possible for safe harbors and sure shipwrecks to behave as bright-line rules do.

C. Policy Maker Framing

To illustrate the importance of a policy maker's framing, consider a new example of a safe harbor. Assume that an administrative agency interprets a statutory requirement of "reasonable rates" for the delivery of packages to prisons⁷² with a safe harbor allowing delivery companies to charge the same rates as the postal service, such as \$20 for three-day delivery of a medium-size box. Does the appearance of the \$20 safe harbor change the possibility of liability when a company charges \$25 for three-day delivery of a medium-sized box? If so, does it increase or decrease the possibility of liability for a \$25 delivery? The answer depends in part on how the agency frames the safe harbor.

Say the guidance is accompanied by the agency's statement that the safe harbor is a preliminary guideline. Assume further that the agency invites delivery services to suggest additional rules protecting higher rates based on evidence of higher cost structures. With this message, a delivery company that charged \$25 could reasonably conclude that the appearance of the safe harbor decreased its chance of liability. Similarly, if a sure shipwreck is understood to open the way for liability for similar cases, it could increase the perception of the likelihood of liability outside the boundaries of the sure shipwreck.⁷³

⁷² The business of prison package delivery faces layers of regulation. See Alan Feuer, *The Cellblocks' Amazon.com*, N.Y. TIMES, Nov. 1, 2013, at M1 (reporting on specialized prison delivery firm sendapackage.com). Rules regulating charges for prison phone calls refer to industry-provided index benchmarks. Rates for Interstate Inmate Calling Services, 78 Fed Reg. 67956 (Nov. 13, 2013) (to be codified at 47 C.F.R. pt. 64) (explaining that permitted rates are based on industry data relating to costs).

⁷³ See Frederick Schauer, *Fear, Risk and the First Amendment: Unraveling the Chilling Effect*, 58 B.U. L. REV. 685, 693, 700-01 (1978) (suggesting that a law "directed at hard-core pornography" might "dete[r] . . . publishing . . . Lady

But the agency might alternatively state that after exhaustive research it has concluded that prison delivery companies “can make a fair profit” at the stated safe harbor rate. This framing suggests that the agency believes that its safe harbor covers just about all of the landscape. As a result, a delivery company that charged \$25 would reasonably conclude that the appearance of the safe harbor increased its chance of liability.

D. *Confined to Facts?*

The exceptional nature of a conclusion on certain facts also can affect how the conclusion influences the probability of liability on other facts. If a court decided that a railroad retained an abandoned right-of-way, rather than suffering the reversion of the right-of-way to the government, the court might explain that the decision arose from the unique or “*sui generis*” nature of railroad-government relations.⁷⁴ The case might produce a safe harbor for railroads holding similar property interests. But it should not affect the likelihood of liability for non-railroad persons who abandon right-of-ways, because the court has taken care to confine the case to its facts.⁷⁵

If the safe harbor or sure shipwreck takes the form of a list or “catalog,” this form also influences the probability of liability in the general standard space. A statute might prohibit leaving a “dog, cat or another pet” in an unattended car.⁷⁶ The statute, which explicitly anticipates application to other situations, will likely increase the probability of liability for leaving a pet rabbit in a car. In contrast, case

Chatterley’s Lover” and arguing that “[a]ny regulation will deter someone somewhere from engaging in conduct that the regulation does not purport to control”).

⁷⁴ Cf. *Marvin M. Brandt Revocable Trust v. United States*, 134 S. Ct. 1257, 1269 (2014) (Sotomayor, J., dissenting) (“[C]ourts have long treated railroad rights of way as *sui generis* property rights not governed by the ordinary common-law regime.”).

⁷⁵ It is possible that non-railroad persons will mistake such a unique holding as one that should apply more generally to allow the retention of property rights in abandoned easements. However, this would misunderstand the law. See Bert I. Huang, *Shallow Signals*, 126 HARV. L. REV. 2227, 2232, 2237-38 (2013) (exploring the possibility that “[w]hen the law quietly permits Actor 1 to act in a way that is usually forbidden, Actor 2 may be misled into taking the liberties with the law that he (mistakenly) perceives Actor 1 as taking” and suggesting that the presence of a licensed ticket seller on the sidewalk outside a theater may lead to the incorrect conclusion that ticket scalping is legal).

⁷⁶ See Parchomovsky & Stein, *supra* note 2, at 170 (giving the example of a sure shipwreck prohibiting leaving a “dog, cat, or another pet” in an unattended car and anticipating that courts will consider whether rabbits count as pets in future adjudications).

law, addressing the same issue, might conclude that the owner of a dog or a cat is liable under common law for animal endangerment as a result of leaving the animal in an unattended car. If the case decision is carefully tailored to the situation presented, it is likely easier to confine to the dog-and-cat facts.

The railroad and dog-and-cat examples both show the possibility that a court decision, rather than a government legislature or agency action, might provide a safe harbor or sure shipwreck. The examples highlight an important difference between case law, on one hand; and legislative or regulatory action, on the other hand. The accepted framework of case law development is the decision of specific cases.⁷⁷ In contrast, the project of legislation and regulation is broader. It is expected that a statute or regulation will apply to some range of situations.

All things equal, the specific case-by-case framework of judge-made law makes case law less likely to produce a discontinuous safe harbor or sure shipwreck, relative to legislative or regulatory law. A judge's decision in a specific case is often taken to support a similar decision in a closely analogous situation. In contrast, the protection of a statutory or regulatory safe harbor is easier to interpret (compared to a case law safe harbor) as protection for only those fact patterns described by the safe harbor, and not for closely analogous fact patterns. Similarly, a sure shipwreck provided in a statute or regulation is more readily understood (compared to a case) as assigning definite liability only for described facts, and not for analogous cases. The underlying premise for these comparative suggestions is that it is the job of a legislature or regulator to think not only about the safe harbor or sure shipwreck case presented, but also about other cases. If one assumes an obligation to think about a broader set of cases, one can more easily conclude that a legislator or regulator (compared to a judge) has deliberately left out cases in the general standard zone.

E. Comparison to Priors

The presence or absence of a gap between pre-existing beliefs about the content of the law and the result given in a safe harbor or certain shipwreck also influences whether the probability of liability changes

⁷⁷ See, e.g., OLIVER WENDELL HOLMES, *THE COMMON LAW* 162 (ABA ed., 2009) (1881) (“The tendency of a given act to cause harm under given circumstances must be determined by experience. And experience . . . is continually working out concrete rules . . .”).

in the adjacent standard space. If pet owners always assumed that it was illegal to leave a dog in an unattended car, then a case so holding likely will not change their understanding of the law applicable to leaving pet rabbits in cars. If they previously thought that leaving a dog in an unattended car was perfectly legal, the news that it is illegal likely will increase pet owners' concern that leaving a rabbit unattended in a car is also illegal.⁷⁸

The understanding of a safe harbor or sure shipwreck by persons subject to it, or by their advisors, also affects the perceived probability of liability in the general standard space. Perhaps persons subject to legal regimes expect bright-line rules, and fail to absorb the subtle structure of a safe harbor or sure shipwreck. Perhaps a lobbyist faced with a \$10 cup-of-coffee safe harbor will not consider the subtlety of the general standard space, and will not recognize that an \$11 cup of coffee might well comport with the underlying standard. Similarly, the lobbyist presented with a \$100 fancy-dinner sure shipwreck may simply see the sure shipwreck through the lens of his expectation that law takes the form of bright-line rules, and conclude that \$100 dinners are not compliant but that \$99 dinners are.

F. When a Space Includes Both a Safe Harbor and a Sure Shipwreck

All this is to say that a safe harbor or sure shipwreck may cause a shift in liability in the general standard space in either direction. A shift in the probability of liability in the general standard space may reduce the discontinuity at the boundary of the safe harbor or sure shipwreck. This results if a safe harbor causes the chance of liability to decrease in the adjacent general standard space, or if a sure shipwreck causes the chance of liability to increase in the adjacent general standard space.⁷⁹

⁷⁸ Emily Cauble also analyzes the interaction between the appearance of a safe harbor and parties' preexisting expectations about results. See Cauble, *supra* note 2, at 1415-17 (arguing that unsophisticated taxpayers may benefit from safe harbors because safe harbors "retain underlying standards that may be more likely than arbitrary rules to coincide with a taxpayer's uninformed expectations").

⁷⁹ Convergence could result even if the probability of a violation for behavior immediately adjacent to a safe harbor equals the probability of a violation for behavior within the safe harbor. For example, the expected probability might be 5% both within and immediately outside the safe harbor. Even in this case, persons right next to the safe harbor might choose to change their behavior in order to fit inside it. They might do so because of risk aversion or uncertainty aversion. See FRANK H. KNIGHT, *RISK, UNCERTAINTY AND PROFIT* 233-34 (1921) (using "risk" to mean a measurable or mathematical uncertainty like that faced in a game of chance and "uncertainty" to mean an unmeasurable uncertainty); see also KENNETH J. ARROW, *ASPECTS OF THE*

Conversely, a shift in liability in the general standard space may increase the discontinuity at the boundary of the safe harbor or sure shipwreck, making it look more like a bright-line rule. This results if a safe harbor causes the chance of liability to increase in the adjacent general standard space, or if a sure shipwreck causes the chance of liability to decrease in the adjacent general standard space.

But note that many legal regimes limit the capacity of safe harbors and sure shipwrecks to behave like true bright-line rules where the chance of liability is 0% on one side of the line and 100% on the other side of the line. This is because legal regimes often have both safe harbors and sure shipwrecks. For example, case law that sometimes finds for the plaintiff and sometimes for the defendant tends to produce both safe harbors and sure shipwrecks. Also, a regulatory structure that gives good and bad examples, or produces adjudications that sometimes favor the government and sometimes not, contains both safe harbors and sure shipwrecks.⁸⁰

THEORY OF RISK-BEARING 28 (1965) (providing risk aversion model); Sarah B. Lawsky, *Modeling Uncertainty in Tax Law*, 65 STAN. L. REV. 241, 259-61 (2013) (citing Daniel Ellsberg, *Risk, Ambiguity, and the Savage Axioms*, 75 Q.J. OF ECON. 643, 650-53 (1961)). Satisficing could also prompt this result. See Herbert A. Simon, *Rational Decision Making in Business Organizations*, 69 AM. ECON. REV. 493, 498 (1979) (connecting the resource constraints that limit the ability to fully analyze every decision with the use of rules of thumb to guide decision making).

The uncertainty aversion and satisficing tendencies of third parties also can encourage the use of safe harbors even if behavior immediately outside a safe harbor also has an equally low chance of liability. Recommending safe harbor use may serve advisors' individual reputational interests. For example, advisors appear to recommend Delaware incorporation because of satisficing or "familiarity" as well as because of the substantive features of Delaware law. See Brian J. Broughman, Jesse Fried & Darian Ibrahim, *Delaware Law as Lingua Franca: Theory and Evidence*, 57 J.L. & ECON. 865, 865 (2014) (arguing that Delaware's dominance as a place of incorporation is "in part due to investors' familiarity with its corporate law"). The use of safe harbors may also serve risk management or other interests of contract counterparties. See, e.g., *Renal Physicians Ass'n v. U.S. Dep't of Health & Human Servs.*, 489 F.3d 1267, 1271-72 (D.C. Cir. 2007) (explaining that dialysis clinics, rather than directly regulated nephrologists, adjusted contracts to comply with safe-harbor compensation levels).

In contrast, equivalent probabilities of liability on either side of a sure shipwreck boundary do not cause persons subject to the rule to choose to move inside the boundaries of the sure shipwreck. This is because a risk-adjusted and/or estimated 99% liability probability is *more* attractive than a for-sure 99% probability of liability. See Amos Tversky & Daniel Kahneman, *The Framing of Decisions and the Psychology of Choice*, 211 SCI. 453, 454 (1981) (presenting evidence of loss aversion or "risk seeking" when persons are faced with the choice of (a) a sure loss versus (b) a chance of losing a larger amount, even where (a) and (b) have equal expected value).

⁸⁰ Many schemes that only explicitly list a safe harbor (or sure shipwreck) also

When both a safe harbor and a sure shipwreck occupy a legal regime, a space of uncertainty occupied by the general standard space necessarily exists between the safe harbor and sure shipwreck. As a result, neither the safe harbor nor the sure shipwreck should behave as a bright-line rule.

Of course, the legal space may become more densely populated by additional safe harbors and sure shipwrecks, meaning increasing numbers of decisions, statutes, regulations and so forth that either condone or prohibit behavior. Perhaps eventually these decisions may come to occupy the entire legal space, so that all behavior in the space would be either prohibited or permitted.⁸¹ Then the result would be that the incremental safe harbors and sure shipwrecks had accumulated to produce a bright-line rule.

IV. WHEN AND HOW TO USE SAFE HARBORS AND SURE SHIPWRECKS

In Part II, I argued that safe harbors encourage behavior convergence; while sure shipwrecks encourage the bunching of behavior immediately on the compliance side of a sure shipwreck boundary, as a result of the incentive to move out of the sure shipwreck zone. This first-order result depends on a discontinuity at the safe harbor or sure shipwreck boundary. As Part III explains, market coordination and/or shifts in the probability of liability within the general standard space can complicate the first-order result.

The analyses in Parts II and III of this Article provided a theory of how safe harbors and sure shipwrecks work. They also prompt additional questions about when and how safe harbors and sure shipwrecks should be used. In this Part IV, my goal is to raise (but not fully analyze) some of these additional questions. I briefly outline the issues of overinclusion and underinclusion presented by safe harbors and sure shipwrecks, safe harbors and sure shipwrecks' combination of ex ante and ex post qualities, and the vulnerability of safe harbors to interest group influence. Future work might explore some or all of these points.

implicitly include a sure shipwreck (or safe harbor). One example is the safe harbor in *Bush v. Gore*, which provided that Congress would accept a list of electors from a state if submitted by December 12, 2000. See *Bush v. Gore*, 531 U.S. 98, 110 (2000). Since the Constitution requires the newly elected President to take office on January 20, January 20, 2001 provided a sure shipwreck for the submission of electoral votes on the facts of *Bush v. Gore*. See U.S. CONST. amend. XX, § 1.

⁸¹ See Frederick Schauer, *supra* note 4, at 804 (arguing that standards develop into rules over time).

A. Overinclusion and Underinclusion

Like bright-line rules, safe harbors and sure shipwrecks present problems of overinclusion and underinclusion.⁸² These problems relate to the quality or congruence of the safe harbor or sure shipwreck. In other words, just like other regimes, safe harbors and sure shipwrecks can reflect better and worse judgments about what behavior ought and ought not be punished according to the underlying policy and wisdom of the relevant law.

Overinclusion means that, relative to the intended or ideal policy that underlies a background legal standard, a rule protects some behavior that ought to produce a noncompliance result or punishes some behavior that ought to produce a compliance result. Underinclusion means the failure of a rule to protect some behavior that is compliant or the failure of a rule to punish some behavior that is noncompliant. Prior work has suggested that safe harbors and sure shipwrecks produce less serious problems of overinclusion and underinclusion compared to bright-line rules.⁸³

Consider the tax law safe harbor that provides that an individual who spends 500 hours a year on a business activity avoids the usually adverse “passive activity loss” rules. Is this a high-quality, or congruent safe harbor, neither overinclusive nor underinclusive? This depends on whether, among other considerations,⁸⁴ (1) we agree that spending a quarter of one’s presumed 2000 hours of annual working time on a business venture is a sufficient commitment given the purposes of the passive activity loss rule⁸⁵ and (2) acceptably low deadweight loss costs result when the safe harbor encourages taxpayers to increase their time spent to 500 hours for targeted business activities.

Another tax law safe harbor provides that a person avoids an adverse “constructive sale” result if she retains exposure to the value of an

⁸² See Sunstein, *Problems with Rules*, *supra* note 1, at 992-93. The idea that an ex ante rule and ex post standard could have exactly the same legal content, or and thus exactly the same congruence, underlies some prior literature. See Kaplow, *supra* note 1, at 586-90. Assuming identical legal content in rules and standards focuses attention on comparing ex ante and ex post costs. However, the assumption that rules and standards will have the same content is too strong for purposes of this Article.

⁸³ E.g., Swire, *supra* note 2, at 373 (suggesting a comparison between safe harbors and bright-line rules).

⁸⁴ Another prerequisite is that the restriction on deducting passive activity losses in the first place is wise policy.

⁸⁵ Cf. Temp. Treas. Reg. § 1.469-5T(a) (2015) (providing a 100-hour rule for persons participating at least as much as any other individual in the enterprise).

asset between 100% and 125% of the price of the asset at the time of a derivative transaction is entered into. This is a low-quality, noncongruent safe harbor if we acknowledge the fact of material differences in asset volatility. It is overinclusive, because it allows a person to use a derivative transaction to accomplish an economic sale of a high-volatility asset without facing an adverse constructive sale result.⁸⁶ It is also underinclusive, because a derivative transaction involving a low-volatility asset might be economically identical to a sale even if the taxpayer retained exposure to the value of the asset for prices between 100% and 125% of the asset's price at the time the derivative transaction was entered into.

As a matter of logic, in contrast to the suggestions in prior literature, the problem of overinclusion is equally acute for safe harbors and sure shipwrecks as for bright-line rules. That is, the problem is identical assuming that the line is drawn in the same place. For example, the same mistakes of overinclusion will be made whether a bright-line rule or a safe harbor provides that an individual who spends 500 hours a year on an activity will avoid the adverse result of the passive activity loss tax rules. Assume for example, that an individual who spends 550 hours really should suffer the consequences provided by the passive activity loss rules. In this case, either a bright-line rule or a safe harbor will erroneously exempt him from those rules.

But as a practical matter, the problem of overinclusion may be less acute for safe harbors and sure shipwrecks, compared to bright-line rules, if it is true that a policy maker tends to draft safe harbors and sure shipwrecks more narrowly compared to bright-line rules. When a policy maker gives a bright-line rule, the policy maker must embrace both the result that behavior on one side of the line is compliant and the result that behavior on the other side of the line is not compliant. In contrast, safe harbors and sure shipwrecks invite policy makers to make decisions incrementally,⁸⁷ and to decide only that a narrow set of facts produces compliance (for safe harbors) or noncompliance (for sure shipwrecks).

For example, the European Union prohibits “illegal, unreported and unregulated” (“IUU”) fishing.⁸⁸ Its regulatory scheme includes port

⁸⁶ See *supra* notes 19–20 and accompanying text (explaining the constructive sale background standard and safe harbor); Thomas Brennan, *Law and Finance: The Case of Constructive Sales*, 5 ANN. REV. FIN. ECON. 259, 259-60 (2013) (modeling the failure of the existing safe harbor Revenue Ruling to account for volatility).

⁸⁷ See *infra* Part IV.B.

⁸⁸ Council Regulation 1005/2008, 2008 O.J. (L 286) 1 (EC), establishing a Community system to prevent, deter and eliminate illegal, unreported and

inspections and other enforcement mechanisms. The scheme also calls for blacklisting specific vessels and/or all vessels flagged in certain countries to prevent these vessels from selling into the European fish market.⁸⁹ The country blacklist serves as a sure shipwreck, not a bright-line rule, since fishing vessels flagged in a country that is not on the blacklist could still face sanctions for IUU fishing under other portions of the regulatory scheme. Because of the possibility of other sanctions, the EU is less likely to blacklist a country. In other words, the EU is less likely to make a mistake of overinclusion. For example, it has issued “yellow cards” to countries with specific suggestions for improvement, rather than putting the countries on the sure-shipwreck blacklist.⁹⁰

The analysis of underinclusion in safe harbors and sure shipwrecks compared to bright-line rules is more subtle. Both safe harbors and sure shipwrecks are intentionally underinclusive. Indeed, this is the key point that separates safe harbors and sure shipwrecks from bright-line rules. Safe harbors and sure shipwrecks acknowledge that they are not solving the questions presented by fact patterns outside their boundaries; they leave that to future decision makers, who apply the background standard.

Take the lobbyist-legislator transfer regime, with the \$10 cup-of-coffee safe harbor and the \$100 fancy-dinner sure shipwreck. This pair of provisions leaves open the possibility that a \$20 sandwich at a conference would comply with the inference of intent to influence standard despite the \$10 cup-of-coffee safe harbor. It also leaves open the possibility that an \$80 dinner at the best restaurant in a legislator’s small town would violate the inference of intent to influence standard despite the \$100 sure shipwreck.

The intentional underinclusiveness of safe harbors and sure shipwrecks might suggest that underinclusion is only a feature, and not a problem, for these rule-standard hybrids. This is not so. For safe harbors, at least, underinclusion is a systematic problem. This is because, as argued in Part II, safe harbors affect the behavior of

unregulated fishing, amending Regulations (EEC) No. 2847/93, (EC) No. 1936/2001, and (EC) No. 601/2004, and repealing Regulations (EC) No. 1093/94 and (EC) No. 1447/1999, 2008 O.J. (L 286) I.

⁸⁹ See Commission Regulation 1010/2009, 2009 O.J. (L280) 5 (EC), laying down detailed rules for the implementation of Council Regulation 1005/2008 establishing a Community system to prevent, deter and eliminate illegal, unreported and unregulated fishing.

⁹⁰ *Questions and Answers on the EU’s fight against illegal, unreported and unregulated (IUU) fishing*, EUR. COMMISSION (Dec. 12, 2014), http://ec.europa.eu/information_society/newsroom/cf/mare/itemdetail.cfm?item_id=19549&subweb=343&lang=en.

persons outside the boundary of the safe harbor. The more secure result of lower liability provided by a safe harbor makes the behavior outside the safe harbor less attractive.⁹¹

For example, the \$20 conference sandwich is less attractive relative to a \$10 coffee after the safe harbor appears. So some persons who otherwise might accept the \$20 sandwich will refuse it. The incentive is gentler compared to the convergence incentive for a bright-line rule, but it exists. This presents the problem of underinclusion if the \$20 conference sandwich is perfectly fine under the background “inference of intent to influence” standard and policy.

The problem of underinclusion means that safe harbors can cause problems when used to describe “easy cases.”⁹² Consider the Department of Education regulations that give colleges a safe harbor for compliance with Title IX. The safe harbor provides that a college does not violate Title IX if the ratio of male intercollegiate athletic participants in athletics to total participants is less than or equal to the ratio of male students to total students.⁹³ Some colleges respond by cutting men’s teams. When members of a cut wrestling team object to the safe harbor regulation, they make an argument of underinclusion, meaning that they argue that numerical proportionality is not the only way to satisfy Title IX’s policy goal of equity in collegiate athletics.⁹⁴

⁹¹ A systematic instance of this phenomenon of underinclusion is found in the publications put out by the Internal Revenue Service. Such guidelines often fit the definition of safe harbors: they give a certain result on certain facts, such as permitting a deduction on certain facts; but leave open the possibility that other facts would also give a similar result. Nevertheless taxpayers may adhere to these safe harbor guidelines as if they are the law, rather than investigating whether the more complex underlying law gives a better result. See Joshua D. Blank & Leigh Osofsky, *Simplexity and the Tax Law*, 66 EMORY L.J. (forthcoming 2017).

⁹² Morrison, *supra* note 2, at 174 (arguing that safe harbors should address “easy cases” against the background of a standard).

⁹³ Title IX of the Education Amendments of 1972, 44 Fed. Reg. 71,413 (1979) (to be codified at 45 C.F.R. pt. 86). The regulations provide a “Three-Part Test” including a safe harbor rule, which applies if intercollegiate athletic opportunities are “substantial[ly] proportionat[e]” to male and female enrollments; and two parts that are adjacent standards and allow the possibility of Title IX compliance if the institution has a sufficient practice of program expansion or if the institution has “fully and effectively accommodated” the interests of male and female students in sports participation. See Catherine F. Pieronek, *The 2010 “Dear Colleague” Letter on Title IX Compliance for College Athletic Programs: Pointing the Way to Proportionality . . . Again*, 38 J.C. & U.L. 277, 299-306 (2012) (describing “clarification” guidance issued in 1996, 2003, 2005 and 2010).

⁹⁴ In one case, the D.C. Circuit refused standing to a coalition of wrestlers and wrestling coaches who had sued the Department of Education on the theory that the safe harbor did not cause the college to take any action. See *Nat’l Wrestling Coaches*

Another example of an easy-case safe harbor guaranteed Medicare reimbursement for dialysis facilities if the referring physician's compensation did not exceed midpoints derived from survey evidence. Dialysis centers responded to the safe harbor by reducing nephrologists' compensation to safe harbor rates. When nephrologists challenged the safe harbor, they made an underinclusion argument, which contended that their previous compensation satisfied the underlying law.⁹⁵

Underinclusion for sure shipwrecks is more complicated. It is possible that a sure shipwreck presents no problem of underinclusion at all, so that, for example, a regulator could write a sure shipwreck prohibiting obviously noncompliant behavior without worrying about inadvertently condoning fact patterns not described by the sure shipwreck guidance. This is the result suggested by the first-order analysis provided in Part II of this Article. That is, if a sure shipwreck's only effect is to change the law applicable to the prohibited behavior, then the only result should be to encourage those within the sure shipwreck to move outside it, into the compliant zone.

But it is also possible that decision factors that apply to behavior outside the sure shipwreck are endogenous to the sure shipwreck. For example, as discussed in Part III, coordination might cause the benefit of behavior immediately outside the sure shipwreck to increase, for example because of the development of a market expectation of behavior that "goes up to the line" of a sure shipwreck. In addition, the appearance of a sure shipwreck may cause a decrease in the likelihood of liability in the general standard space adjacent to the sure

Ass'n v. Dep't of Educ., 366 F.3d 930, 938 (D.C. Cir. 2004). Other Title IX cases have extended standing to plaintiff organizations representing cut men's teams. *See, e.g.,* Equity in Athletics, Inc. v. Dep't of Educ., 639 F.3d 91, 101 n.7 (4th Cir. 2011) (granting standing because plaintiffs sued the university as well as the Department of Education and because of a lack of evidence that the college would have "eliminate[d] specific teams" in the absence of the Department of Education guidance); *see also* Bennett v. Donovan, 703 F.3d 582, 588 (D.C. Cir. 2013) (analyzing standing by considering whether, "with respect to the action at issue in a particular case" a third-party action is "independent of government policy" (emphasis in original)).

⁹⁵ The D.C. Circuit held that the nephrologists lacked standing for lack of a clear causal link between the safe harbor and dialysis center compensation decisions. *See* Renal Physicians Ass'n v. U.S. Dep't of Health & Human Servs., 489 F.3d 1267, 1277 (D.C. Cir. 2007) ("[W]e . . . have no way of knowing why the dialysis facility . . . reduced Dr. Anzalone's hourly wage . . ."). It held the causation and redress prongs missing. Standing does not present a similar causation question in the case of a bright-line rule or sure shipwreck. *See, e.g.,* Nat'l Wrestling Coaches Ass'n., 366 F.3d at 941-42 (noting cases in which parties alleging injury caused by a sure shipwreck-type rule had standing).

shipwreck. As Part III outlines, this depends on factors including the policy maker's framing, the uniqueness of the sure shipwreck fact pattern and the comparison between the sure shipwreck and prior beliefs about the content of the law. All of this means that sure shipwrecks' susceptibility to problems of underinclusion is contextual.

To illustrate, consider a sure shipwreck that prohibits the use of a number of listed tax shelters. Assume that there are also other, equally objectionable tax shelters, which do not make the list.⁹⁶ This sure shipwreck regime presents an underinclusion problem if taxpayers begin to use the unlisted shelters instead of the shelters on the list. Will taxpayers increase their use of the unlisted shelters? It depends. A market expectation that taxpayers will start to use the unlisted shelters would increase the risk of underinclusion. If the tax agency commits resources only to punishing the listed shelters, the risk of underinclusion increases. Considerations like the framing of the rule by the tax administrators and its comparison to the priors of taxpayers will affect the impact of the sure shipwreck. And factors can also interact with each other. For example, increased use of the unlisted tax shelters by taxpayers might increase⁹⁷ or decrease⁹⁸ the likelihood that the government will also attempt to punish users of the unlisted shelters.

B. *Ex Ante and Ex Post*

Safe harbors and sure shipwrecks have a temporal dimension. A safe harbor or sure shipwreck provides an *ex ante* answer for some factual situations covered by the legal standard. The remainder of the legal standard is left open for future development. The *ex post* space exists because later decision makers have the responsibility to determine the results of cases not covered by the safe harbor or sure shipwreck.⁹⁹

⁹⁶ For example, this could be because the government has not discovered the unlisted tax shelters yet.

⁹⁷ See Mark P. Gergen, *The Logic of Deterrence: Corporate Tax Shelters*, 55 *TAX L. REV.* 255, 263-65 (2002) (explaining incentives for users of corporate tax shelters under the assumption that additional iterations will increase the likelihood of detection).

⁹⁸ See, e.g., Sam Young & Lee Sheppard, *Korb Slams Textron Ruling*, *Wall Street Rule*, *TAX NOTES*, Oct. 15, 2007, at 204 (quoting then-IRS Commissioner Donald Korb as citing and criticizing the "Wall Street Rule" "creed" that says "if I can do enough deals before the commissioner finds out and acts, then I win").

⁹⁹ Rules-versus-standards choices also allocate decision making among different decision makers, for example between legislators and judges, or between current and future iterations of an administrative agency. See FREDERICK SCHAUER, *PLAYING BY THE RULES: A PHILOSOPHICAL EXAMINATION OF RULE-BASED DECISION-MAKING IN LAW AND IN LIFE* 158-62, 231-33 (1991).

The costs of an ex ante rule compared to the costs of an ex post standard properly influence a choice to use a safe harbor or a sure shipwreck. As prior literature has observed, if the content of an ex ante rule and an ex post standard are the same, then a rule is efficient when the rulemaking, compliance and enforcement costs of an ex ante rule are less than the costs of ex post decision making. A determination that the costs of an ex ante rule are lower for only some factual situations supports a decision to establish a safe harbor or sure shipwreck.¹⁰⁰

Relevant costs for the ex ante/ex post contrast include rulemaking, compliance and enforcement costs. These factors translate to several reasons why a decision maker might choose to replace only part of a legal standard with a safe harbor or sure shipwreck.¹⁰¹ For example, if persons subject to a legal regime are more numerous and/or homogenous, then an ex ante rule makes more sense, all else equal. If a court or agency faces an area featuring unsure and rapidly developing technology, then an ex ante rule makes less sense, all else equal.¹⁰²

The institutional capacity of an agency, court, or other decision maker also affects the decision whether to adopt a safe harbor or sure shipwreck. Sometimes, an agency may only have the capacity to immediately solve part of the problem before it.¹⁰³ This can result from resource constraints that include the challenge of too little

¹⁰⁰ See Kaplow, *supra* note 1, at 621 (contrasting ex ante costs of promulgating rules with ex post costs of applying standards). The distributive question of whether to weight costs differently depending on who bears them lies beyond the scope of this article. Cf. Louis Kaplow & Steven Shavell, *Why the Legal System Is Less Efficient than the Income Tax in Redistributing Income*, 23 J. LEGAL. STUD. 667, 667-68 (1994) (arguing that the income tax system is systematically better than regulatory systems at redistribution).

¹⁰¹ See Kaplow, *supra* note 1, at 621-24 (listing and analyzing different ex ante and ex post costs).

¹⁰² Other solutions to developing technology also exist, including automatically dynamic rulemaking approaches. See, e.g., Lynn Blais & Wendy Wagner, *Emerging Science, Adaptive Regulation, and the Problem of Rulemaking Ruts*, 86 TEX. L. REV. 1701, 1731-37 (2008) (considering adjustment mechanisms including “contemporaneous revision planning,” “revision rulemaking” and periodic review); cf. *Indus. Union Dept. v. Am. Petroleum Inst.*, 448 U.S. 607, 607 (1980) (invalidating OSHA benzene exposure regulation interpreting a “safe and healthful” standard and reasoning that OSHA had insufficient evidence to show that an emission level of 1 ppm or less was required to accomplish the goal of a safe and healthful workplace).

¹⁰³ See Parchomovsky & Stein, *supra* note 2, at 172 (arguing that catalogs, or open-ended lists of situations that give sure legal results, have an advantage over rules and standards because they “can be expeditiously and cheaply adapted to accommodate changes while reducing information costs for actors”).

information, or because of information overload;¹⁰⁴ because of the “ossification” that can follow from agency risk aversion and the threat of judicial review¹⁰⁵ and/or because of political considerations outside the agency.¹⁰⁶ Courts may face similar information constraints, and also must operate within the limitation that requires judges to decide cases presented, not hypothetical cases.

The qualities of regulated parties and/or the institutional capacity of an agency or court can change over time, and safe harbors and sure shipwrecks accommodate this within the ex post standard space. In other words, safe harbors and sure shipwrecks can help decision makers use the dimension of time¹⁰⁷ to break rulemaking tasks into smaller, more manageable parts.¹⁰⁸ But using time in this way may also

¹⁰⁴ See, e.g., Douglas A. Kysar & James Salzman, *Foreword: Making Sense of Information for Environmental Protection*, 86 TEX. L. REV. 1347, 1347-48 (2008) (considering the problem of information for regulators).

¹⁰⁵ Compare Thomas McGarity, *Some Thoughts on “Deossifying” the Rulemaking Process*, 41 DUKE L.J. 1385, 1400-07 (1992) (identifying judicial and other oversight reasons for ossification), with Mark Seidenfeld, *Why Agencies Act: A Reassessment of the Ossification Critique of Judicial Review*, 70 OHIO ST. L.J. 251, 259-60 (2009) (describing a “purposive” method of setting an agency’s agenda under which the individuals in the agency retain considerable power).

¹⁰⁶ See, e.g., Elena Kagan, *Presidential Administration*, 114 HARV. L. REV. 2245, 2315-19 (2001) (exploring increased impact of presidential priorities on some rulemaking projects under the first George W. Bush and Clinton administrations); Jennifer Nou, *Agency Self-Insulation Under Presidential Review*, 126 HARV. L. REV. 1755, 1796-98 (2013) (noting agencies’ use of timing strategies to reduce the chance of close presidential review); see also Thomas O. McGarity, *Administrative Law as Blood Sport: Policy Erosion in a Highly Partisan Age*, 61 DUKE L.J. 1671, 1703-19 (2012) (describing strategies available for Congressional challenge of administrative policies).

¹⁰⁷ Others have observed that administrative rulemaking has a temporal dimension. See, e.g., Yoon-Ho Alex Lee, *An Options Approach to Agency Rulemaking*, 65 ADMIN. L. REV. 881, 888 (2013) (arguing based on option theory that agencies should carefully consider “policy reversibility” and should face a more lenient standard of judicial review for rules that permit “ex post exemptions”); Yair Listokin, *Learning Through Policy Variation*, 118 YALE L.J. 480, 499-500, 522-23 (2008) (arguing that reversible policies should pursue a high-risk, or high-variance approach, while a policymaker might choose to delay the implementation of an irreversible policy); see also Brian D. Galle, *In Praise of Ex Ante Regulation*, 68 VAND. L. REV. 1715, 1715 (2015) (arguing that ex ante regulation has important advantages and that heterogeneity can be addressed by establishing different zones of regulation for different cases).

¹⁰⁸ In other words, safe harbors and sure shipwrecks can combine features of “comprehensive rationality” and incremental “muddling through.” Colin S. Diver, *Policymaking Paradigms in Administrative Law*, 95 HARV. L. REV. 393, 394-95 (1981) (describing “comprehensive rationality” and “incrementalism”); Charles Lindblom, *The Science of Muddling Through*, 19 PUB. ADMIN. REV. 79, 82 (1959) (arguing that “the administrator focuses his attention on marginal or incremental values”). Various scholars recommend that agencies take a more comprehensive rationality view in

expose decision makers to the pitfall of increasing interest group influence over less visible, incremental decisions.¹⁰⁹

C. Safe Harbors and the Problem of Capture

Because safe harbors produce a good result for persons subject to them, there is an increased risk of interest group influence for safe harbor rulemaking. Interest-group theory predicts regulatory rules benefiting small, well-organized interest groups even if such rules present detriments to larger, more diffuse groups.¹¹⁰ It thus suggests

specific subject areas. See, e.g., Daniel J. Gifford, *Labor Law and Its Reform*, 80 IOWA L. REV. 201, 203 (1994) (arguing the NLRB should take a more comprehensive view); Allen Rostron, *Incrementalism, Comprehensive Rationality, and the Future of Gun Control*, 67 MD. L. REV. 511, 513-14 (2008) (discussing gun control laws); J.B. Ruhl & James Salzman, *Climate Change, Dead Zones, and Massive Problems in the Administrative State: A Guide for Whittling Away*, 98 CALIF. L. REV. 59, 119-20 (2010) (discussing administrative agency coordination and climate change); see also David A. Super, *Against Flexibility*, 96 CORNELL L. REV. 1375, 1380-81 (2011) (stating a preference for earlier decisions on the theory that this reduces policymaker bias and citing the case of disaster relief regulation). But the practical reality of muddling through is widely acknowledged. See, e.g., Lynn M. LoPucki, *The Systems Approach to Law*, 82 CORNELL L. REV. 479, 521 (1997) (acknowledging descriptive importance of incremental account of rulemaking). Other scholars acknowledge the policy potential of incrementalism. See, e.g., Rachel Brewster, *Stepping Stone or Stumbling Block: Incrementalism and National Climate Change Legislation*, 28 YALE L. & POL'Y REV. 245, 311-12 (2010) (arguing that incrementalism can create feedback effects that feature in a productive interaction between national legislation and international negotiation); Daniel R. Cahoy, *An Incrementalist Approach to Patent Reform Policy*, 9 N.Y.U. J. LEGIS. & PUB. POL'Y 587, 635-40 (2006) (advocating trial-and-error analysis in patent reform due to limited availability of empirical inputs); Robert L. Glicksman & Sidney A. Shapiro, *Improving Regulation Through Incremental Adjustment*, 52 U. KAN. L. REV. 1179, 1180-81 (2004) (suggesting greater reliance on back-end regulatory adjustment and recommending public access to process).

¹⁰⁹ See Saul Levmore, *Interest Groups and the Problem with Incrementalism*, 158 U. PA. L. REV. 815, 855 (2010) (outlining the possibility of an interest group's plan to accomplish a regulatory goal incrementally over time and thus fragment possible opposition); see also Diver, *supra* note 108, at 399, 402-03, 406 (noting that making rules "piecemeal" over time is a more legitimate approach in the presence of "decentralized decisionmaking" where interested parties have access to decisionmakers).

¹¹⁰ See, e.g., James Q. Wilson, *The Politics of Regulation*, in SOCIAL RESPONSIBILITY AND THE BUSINESS PREDICAMENT 135 (James W. McKie ed., 1974) ("[R]egulatory constraints often arise out of a political situation in which a small, relatively homogeneous beneficiary group can make substantial gains by imposing unobtrusive costs on large numbers of others."); see also Cass R. Sunstein, *Interest Groups in American Public Law*, 38 STAN. L. REV. 29, 48 (1985) (suggesting a continuum where at one pole "interest-group" controls and where at the other pole faction pressures are unimportant).

the paradigm of an industry or other regulated-party group quietly and successfully pushing for a safe harbor while potential opponents remain paralyzed by collective action problems.¹¹¹

Of course, safe harbors sometimes face opposition from organized opposing groups, whether of industry competitors, public interest protectors, or otherwise. In addition, agencies can develop safe harbor rules despite the fact that no interest group requests them. Nevertheless, there remains a vulnerability: because safe harbors protect an identified group of regulated parties, it is likely that such protected parties will seek safe harbors; while, in general, others will have less reason to oppose the safe harbor compared to the protected parties' incentive to pursue it.

How serious is the problem of interest-group influence and safe harbors? The answer boils down to a familiar tension between public choice theory and accounts of agency independence and expertise. If public choice theory is correct that bureaucrats are captured by interest-group influence,¹¹² then we should expect safe harbors to bring out the most extreme forms of rent-seeking cronyism. This concern would be particularly strong if the safe harbors develop from informal and incremental guidance rather than full-fledged and fully public notice-and-comment rulemaking.¹¹³

Yet public choice theory is not necessarily a true account of how government agencies work.¹¹⁴ Perhaps bureaucrats face multiple motivations, including job security, institutional power, and good policy.¹¹⁵ Or perhaps administrators strive for good policy decisions subject to resource, institutional, exogenous political and other constraints.¹¹⁶ To the extent the expertise story is right, or that

¹¹¹ See MANCUR OLSON, *THE LOGIC OF COLLECTIVE ACTION: PUBLIC GOODS AND THE THEORY OF GROUPS* 144 (1965) (noting "political advantages of the small groups of large units").

¹¹² See DANIEL A. FARBER & PHILIP P. FRICKEY, *LAW AND PUBLIC CHOICE: A CRITICAL INTRODUCTION* 7-9 (1991) (explaining the idea that economic rent-seeking motivates bureaucrats' actions).

¹¹³ See *supra* note 109 and accompanying text; see also SIDNEY A. SHAPIRO & ROBERT L. GLICKMAN, *RISK REGULATION AT RISK: RESTORING A PRAGMATIC APPROACH* 158-68, 175 (2003) (stating support for exceptions and other "back-end adjustments" "only if the opportunities for participation afforded interested persons are essentially equivalent to those that govern adoption of regulatory standards in the first place").

¹¹⁴ Cf. STEPHEN G. BREYER, *REGULATION AND ITS REFORM* 9 (2009) (arguing that relative political power cannot explain regulation).

¹¹⁵ Cf. RICHARD F. FENNO JR., *CONGRESSMEN IN COMMITTEES*, at xiii (1973) (identifying goals of influence, re-election and policy for members of congressional committees).

¹¹⁶ See Thomas McGarity, *Regulatory Analysis and Regulatory Reform*, 65 *TEX. L.*

another solution to the problem of capture or interest group influence is available,¹¹⁷ agencies have interesting opportunities to use leverage over regulated parties who desire safe harbors.¹¹⁸ For example, an agency might force regulated parties to internalize some rulemaking costs,¹¹⁹ for example by providing better and more usable information.¹²⁰

CONCLUSION

Safe harbors and, as I call them, sure shipwrecks are rule-standard hybrids that appear throughout statutory, regulatory and case law. Safe harbors guarantee compliance, and also leave open the possibility of compliance for fact situations not described by the safe harbor. Sure shipwrecks describe behavior that violates the law as a matter of rule, and also leave open the possibility of a noncompliance result for fact situations outside the sure shipwreck.

This Article supplies a theory of safe harbors and sure shipwrecks that breaks new ground in the rules and standards literature. The first-order analysis shows that safe harbors and sure shipwrecks produce asymmetric behavioral incentives for regulated parties. Safe harbors encourage parties on both sides of the line drawn by the safe harbor to change their behavior so that it converges on the line drawn by the

REV. 1243, 1254 (1987) (defining “techno-bureaucratic rationality”).

¹¹⁷ See, e.g., Ronen Avraham, *Private Regulation*, 34 HARV. J.L. & PUB. POL’Y 543, 591-93 (2011) (recommending that private regulators develop medical standards of care and face liability “as determined from the ex ante perspective, for writing guidelines that are inefficiently risky”); Omri Ben-Shahar & Kyle D. Logue, *Outsourcing Regulation: How Insurance Reduces Moral Hazard*, 111 MICH. L. REV. 197, 197 (2012) (arguing that the insurance plus strict- or no-liability legal rules can substitute for government safety regulation).

¹¹⁸ For example, Bradley Karkkainen has analyzed California’s Proposition 65 emissions standard, Cal. Health & Safety Code §§ 25249.5–13, which invited regulators to avoid an unattractive default regime by proposing and defending safe harbor emissions limits. He reports that it took only “a few months” for California regulators to establish safe harbor levels for hundreds of substances under this regime. Bradley Karkkainen, *Bottlenecks and Baselines: Tackling Information Deficits in Environmental Regulation*, 86 TEX. L. REV. 1409, 1430 (2008) (“[B]y shifting the default position to one of uncertainty and risk on the part of polluters, Proposition 65 profoundly changes the dynamics of information flow.”).

¹¹⁹ See A. Mitchell Polinsky & Steven Shavell, *The Economic Theory of Public Enforcement of Law*, 38 J. ECON. LIT. 45, 57-58 (2000) (explaining rationale for requiring internalization of public enforcement costs).

¹²⁰ See, e.g., Wendy E. Wagner, *Administrative Law, Filter Failure, and Information Capture*, 59 DUKE L.J. 1321, 1419-20 (2010) (suggesting “information filter” measures such as length-limited policy briefs).

safe harbor. This is because of the advantage of a lower chance of liability at the safe harbor boundary. Sure shipwrecks encourage bunching immediately on the compliant side of the line drawn by the sure shipwreck as a result of the incentive to move out of the sure shipwreck space.

Safe harbors and sure shipwrecks are natural tools for incremental policy making over time. But they have design flaws as well. The first-order results of convergence for safe harbors and bunching for sure shipwrecks can shift in the presence of market coordination or changes in the possibility of liability outside a safe harbor or sure shipwreck. In addition, both safe harbors and sure shipwrecks can present problems of overinclusion like those presented by bright-line rules; and safe harbors confined to “easy cases” systematically present problems of underinclusion. The issue of interest group influence also influences safe-harbor and sure-shipwreck policy.