GENERAL ARTICLES

TWO TYPES OF LEGAL UNCERTAINTY

Orlin Yalnazov*

Law-and-economics scholars analyse legal uncertainty as a choice between rules and standards. In doing this, they focus on individual laws that regulate and sanction conduct, or what Hart would call 'primary rules'. Hart also spoke of 'secondary rules', that is, rules that determine the validity and precedence of other rules. Here, I introduce secondary rules into the law-and-economics framework. Two types of uncertainty emerge. I call the one covered in the literature 'applicative uncertainty' and the 'new' one 'hierarchic uncertainty'. I show that the two always co-exist and, further, that there is a trade-off between them. I sketch out the economics of that trade-off and I discuss its implications for legal certainty in general.

Keywords: law-and-economics, rules versus standards, legal uncertainty

TABLE OF CONTENTS

I. LITERATURE REVIEW	13
II. APPLICATIVE AND HIERARCHIC UNCERTAINTY	17
1. <i>Applicative Uncertainty</i>	19
2. <i>Hierarchic Uncertainty</i>	21
3. <i>Penalty Clauses: An Example</i>	25
III. THE RELATIONSHIP BETWEEN APPLICATIVE AND HIERARCHIC INDETERMINACY	30
1. <i>Complexity</i>	30

* PhD Candidate, European Doctorate of Law and Economics (Erasmus University Rotterdam, Universität Hamburg, University of Bologna). My correspondence e-mail is orlin.yalnazov@edle-phd.eu. I am grateful to Michael Faure, Florian Faust, Elena Reznichenko, Janneke van Casteren, Fabrizio Esposito, and an anonymous reviewer for their comments on the innumerable drafts of this paper. My doctoral research, of which this paper forms part, was funded by the Deutscher Akademischer Austauschdienst.

2. <i>Metalaws and the Legal System</i>	33
3. <i>Penalty Clauses: A Second Example</i>	37
IV. CONCLUSION	45

When law-and-economics scholars discuss legal uncertainty, they problematise the use of vague language in the drafting of laws.¹ But when we say that the law is uncertain, we might also mean that two (or more) laws overlap, so that we do not know which is applicable to the facts of a particular case. I call the problem that the law-and-economics literature covers one of *applicative uncertainty* and the 'new' one *hierarchical uncertainty*. I will try to show that the two always coexist, and further that there is a trade-off between them: as the law becomes more applicatively certain, it also grows more hierarchically uncertain. I sketch out the economics of that trade-off and I discuss the implications for legal certainty in general.

I also propose a refinement to the standard mode of juridico-economic analysis. There is a distinction between laws and metalaws. An 'ordinary' law defines the conditions under which the state will exercise its coercive powers: murder triggers imprisonment, breach of contract generates an enforceable obligation to pay damages, those caught speeding are forced to pay fines, and so on. The orthodox approach in law-and-economics is to take one such 'ordinary' law and to analyse its allocative implications. Metalaws, on the other hand, have been overlooked in the literature. A metalaw is a law about laws: laws made by parliaments trump laws made by courts, laws in the

¹ See for example, Louis Kaplow, 'Rules versus Standards: an Economic Analysis' (1992) 42 *Duke Law Journal* 557; Ian Ayres and Eric Talley, 'Solomonic Bargaining: Dividing a Legal Entitlement to Facilitate Trade' (1995) 104 *Yale Law Journal* 1027; Jason Scott Johnston, 'Bargaining under Rules versus Standards' (1996) 11 *Journal of Law, Economics, and Organization* 256; Russel Korobkin, 'Behavioral Analysis and Legal Form: Rules v Standards Revisited' (2000) 79 *Oregon Law Review* 23; Isaac Ehrlich and Richard Posner, 'An Economic Analysis of Legal Rulemaking' (1974) 3 *Journal of Legal Studies* 257; Hans-Bernd Schäfer, 'Rules versus Standards in Rich and Poor Countries: Precise Legal Norms as Substitutes for Human Capital in Low-Income Countries' (2006) 14 *Supreme Court Economic Review* 113; Franziska Weber, 'European Integration Assessed in the Light of the 'Rules versus Standards' Debate' (2013) 35 *European Journal of Law and Economics* 187; Ezra Friedman and Abraham Wickelgren, 'A New Angle on Rules versus Standards' (2014) 16 *American Law and Economics Review* 499.

constitution trump laws in statutes, liability in the law of tort is concurrent to liability in the law of contract. I argue that applicative uncertainty is reduced through the adoption of more specific 'ordinary' laws. But as the law grows more complex, it also grows more hierarchically uncertain. Hierarchic uncertainty can only be managed by metalaws. From this analytical scheme, I derive a speculative theory of the optimal complexity of legal systems. I also discuss the hitherto overlooked use of metalaws in legal prediction.

I. LITERATURE REVIEW

Legal theorists and lawyer-economists mean different things when they use the term 'uncertainty'. What I call applicative certainty is extensively covered in both strands of the literature. Hierarchic uncertainty is new to law-and-economics, but not to legal theory. To avoid confusion and to contextualise this paper, it is best that I begin with a brief summary of the main authorities on the point.

Applicative uncertainty is the uncertainty that results from the need to apply general laws to specific facts. In law-and-economics, Kaplow, among others, distinguishes between standards and rules, and says they differ in their level of generality.² That conceptualisation is analogous to the positivist understanding of uncertainty. Kelsen speaks of a norm's 'frame of application'.³ All laws are written in human language, and human language is vague. Therefore, every law establishes a frame in which a number of outcomes can lie. Hart speaks of a core and a penumbra.⁴ Every legal norm has a core, that is, a set of factual matrices whose correspondence, or lack thereof, to the norm's text is self-evident. Norms also have penumbrae, that is, there are always factual situations to which the applicability of the norm is essentially contestable. Raz speaks of indeterminacy on a continuum: there are legal questions to which there are true or false answers, questions to which

² Kaplow (n 1) 560.

³ Hans Kelsen, *Introduction to the Problems of Legal Theory* (Bonnie Litchewski Paulson and Stanley L Paulson trs, Oxford University Press 2002) 80.

⁴ HLA Hart, 'Positivism and the Separation of Law and Morals' (1958) 71 *Harvard Law Review* 593, 607-15.

there is no true or false answer, and questions to which it is unclear whether there is a true or false answer.⁵

Applicative uncertainty in law-and-economics is analogous to uncertainty in legal positivism. In the writings of Kelsen, Hart, and Raz, all laws produce uncertainty because they are partly inexact. In law-and-economics, the inevitability of uncertainty is taken as a given and the central question is how much effort ought to be exerted in limiting its scope. In Kelsenian terms, law-and-economics scholars try to find the optimal width of a law's frame of application. The same point can be translated into Hart's terminology by saying that the juridico-economic concern is with the optimal size of a law's penumbra. And in the language of Raz, the purpose of the enterprise is to ascertain the optimal number of questions to which there is a true or false answer.

I now turn to hierarchic uncertainty. Barring a footnote in Kaplow's work,⁶ law-and-economics is silent on the point. In legal theory, on the other hand, hierarchic uncertainty is amply covered. Kelsen spoke of a 'chain' of legal norms which exist in a hierarchical relationship.⁷ He was also of the view that higher-order norms have frames of application, leaving open the possibility of there being a number of equally plausible lower-level norms.⁸ Hart distinguished between primary and secondary rules. These two types of rules correspond closely to laws and metalaws in the language of this paper.⁹

⁵ Joseph Raz, *The Authority of Law: Essays on Law and Morality* (Oxford University Press 1979) 72-4.

⁶ Kaplow (n 1) footnote 150 *infra*.

⁷ Kelsen (n 3) 71-5.

⁸ *Ibid* 78.

⁹ The correspondence is not, however, perfect. To Hart, secondary rules were subdivided into rules of recognition, rules of change, and rules of adjudication: see HLA Hart, *The Concept of Law* (2nd edn, Oxford University Press 1994) 94-5. My concept of metalaws captures rules of recognition and some rules of change, but not rules of adjudication. A metalaw, in my framework, identifies other applicable laws. The rule of recognition does precisely that. A rule of change can serve as a metalaw indirectly: it is sometimes necessary to know who and when can change the law to determine whether some law is in fact valid at the moment. To give an example, in *Willers v Joyce* [2016] UKSC 43 & 44, the dispute turned on whether the Privy Council in *Crawford Adjusters (Cayman) v Sagicor General Insurance (Cayman)* [2014] AC 366 had overruled the House of Lords in *Gregory v Portsmouth City Council* [2000] 1 AC 419. That, in

Secondary rules serve to identify primary rules. To Hart, secondary rules were subject to the ordinary constraints of human language, including its 'open texture'.¹⁰ Cases may arise where a choice has to be made between two primary rules by reference to a secondary rule which is too vague to guide that choice.¹¹ In those circumstances, there is hierarchic uncertainty: a clash of two primary rules with no indication of which ought to prevail. Raz is also alive to the problem. He posits that sometimes a legal decision must be made, even though the law supplies conflicting reasons.¹² There are situations in which the law gives no indication of the reason to be preferred, others in which the two are equally matched, and others still in which the reasons are incommensurate. Such gaps, Raz shows, are inescapable in any legal system.¹³

Here I follow Hart and Raz, that adhesion manifesting in two specific assumptions. Firstly, I assume that there are facts to which two (or more) laws apply and that those laws may point to two (or more) different outcomes. Secondly, and more controversially, I assume that such conflicts can, but need not be, resolved only through metalaws produced by a lawmaker. I treat the Sources Thesis as a given. That thesis, in short, postulates that the content of the law can be determined from its sources – no moral reasoning is necessary.¹⁴ Why might the adoption of that thesis be controversial? Dworkin, in challenging Hart, has made the argument that although uncertainty exists, it is always soluble by reference to 'principles' immanent in the juridico-political order.¹⁵ Principles, in Dworkin's terminology, are closely analogous to metalaws in mine. My theory, however, is not consistent

turn, turned on a rule of change: could the Privy Council overrule the House of Lords on a point of English law? The identification of the correct rule to apply turned on the rule of change. However, rules of adjudication – say the definition of a court – do not identify the law to apply in particular cases. Therefore, they are not metalaws.

¹⁰ Hart (n 9) 125.

¹¹ Ibid 134-6.

¹² Raz (n **Error! Bookmark not defined.**) 74-6.

¹³ Ibid 77.

¹⁴ See Raz (n 5) 47ff; Joseph Raz, 'Legal Principles and the Limits of Law' (1972) 81 Yale Law Review 823.

¹⁵ See Ronald Dworkin, 'The Model of Rules' (1967) 35 University of Chicago Law Review 14; Ronald Dworkin, 'No Right Answer?' (1978) 53 New York University Law Review 1; Ronald Dworkin, *Law's Empire* (Harvard University Press 1986) 87-114.

with his. If it is always possible to discover a metalaw that gives the 'correct' outcome of a case, then evidently there is in every case an ideal solution to hierarchic uncertainty. Even though there is hierarchic uncertainty, it is not a problem that needs to be solved through the creation of new metalaws. There is only the issue of finding them. I add nothing to that argument here, other than noting that full acceptance of Dworkin's thesis would, for better or for worse, render most of the law-and-economics literature on the subject meaningless.¹⁶ Applicative uncertainty is unproblematic if whenever it arises, a canon of interpretation exists which delivers an unambiguous resolution to the controversy at hand. If one agrees with Dworkin's version of uncertainty, then one is unlikely to find the kind of analysis developed here very useful.

Lastly, I should mention another non-positivist account of law to which I think my theory applies, but to which I have not sought to contribute here. Legal realism implies, loosely speaking, that judicial ideology matters.¹⁷ Theorists in that tradition are generally divided on the extent to which a judge's personal ideological preferences are determinative of legal outcomes. Based on a 'hard' legal realist conception, ideology preordains all judicial outcomes. To this, my account of hierarchic uncertainty has little to contribute, for the obvious reason that a 'hard' realist would have little time for the interrelationship between laws within the system. A milder realist account of judicial decision-making posits that legal materials, such as laws, can be fashioned into arguments for one outcome or another through legal work.¹⁸ The pliability of the materials is evidently determinative, though only

¹⁶ That of course does not refute Dworkin's theory. Nor would Dworkin, who is highly sceptical of law-and-economics, be too troubled by it: see Ronald Dworkin, 'Is Wealth a Value?' (1980) 9 *Journal of Legal Studies* 191.

¹⁷ Joseph William Singer, 'The Player and The Cards: Nihilism and Legal Theory' (1984) 94 *Yale Law Journal* 1; Jerome Frank, *Law and the Modern Mind* (Transaction Publishers 1930).

¹⁸ Duncan Kennedy, 'A Left/Phenomenological Alternative to the Hart/Kelsen Theory of Legal Interpretation' in Duncan Kennedy (ed), *Legal Reasoning: Collected Essays* (Davies Pub 2008). Kennedy rejects the notion of there being cores and penumbrae, but does not dispute that norms vary in their determinacy. At the root of his disagreement with Hart and Kelsen is their notion that indeterminacy is an inherent property of legal texts, which to him overlooks the role of the interpreter. The analysis developed here is agnostic on this point, so it might be applicable to arguments in the realist tradition.

partly, of the range of potential outcomes. Hierarchic uncertainty is thus a property of those materials that matters to judicial output. Although I think what follows might be applicable to theories grounded in that milder realist position, I do not examine the implications here.

II. APPLICATIVE AND HIERARCHIC UNCERTAINTY

Let me begin by defining 'a law'. By that term, I do not mean 'all law' or 'the law', nor does my usage cover the entire spectrum of statements which in ordinary parlance are called laws. Nor do I propose to build a novel theory of law in these pages. I use the term merely to denote the specific type of provision that forms the subject matter of my analysis. To my ends here, a law is a sanctionability condition. It is a description of the conditions in which the state will deploy its coercive power against some individual. This might on first impression appear to limit the analysis developed here to criminal and administrative law. That impression is misleading. For example, the law of contract guarantees that if parties choose some arrangement, the state will facilitate it through the threat of coercive enforcement, or ultimately through its execution. Many of its substantive provisions: remedies, penalty clauses, and others, refer to sanctions for non-compliance. They are laws in the sense in which I defined the word here. The same is true in torts and in property.¹⁹

Laws thus defined may and most commonly are positive, as is the case when the state declares that certain actions or inactions²⁰ will trigger some sanctions. Laws can also occasionally be negative, as they are when the state declares that some course of conduct will attract no liability. For expositional

¹⁹ This definition excludes a number of propositions which we would ordinarily dub legal. For example, the law of procedure does not indicate any sanctions – so it is not a law in the sense in which I use the term here. Rules which prescribe the conditions under which a marriage or a will is valid are also excluded, for the same reason. By this exclusion, I do not mean to deny that those are laws – it is merely the case that they fall outside the ambit of the analysis which I develop.

²⁰ English law does not generally recognise liability for 'pure' omissions: see *Stovin v Wise* [1996] UKHL 15 and *Smith v Littlewoods Organisations* [1987] 2 AC 241. The position is different in civilian countries: French law imposes criminal liability for failure to rescue: see Art 223-6 *Code pénal*.

convenience, let me now give three examples of statements that can constitute laws:

- (A) If a person negligently harms another, they shall have to pay compensation.
- (B) If a person causes death, they shall be liable to imprisonment.
- (C) If a person drives a car at an unreasonable speed, they shall be liable to a fine.²¹

Legal uncertainty, in the umbrella sense of that term, exists when it is impossible to predict whether a particular set of facts will trigger a sanction with a probability of one.²² For example, suppose that we are told that the law of some imaginary polity comprises the three laws above, and no others. You and I are both subject to the laws of that polity. The following events transpire: I am driving my car at 45 km/h while gesticulating at my tearful girlfriend. While this is ongoing, I run over your cow. The force of the impact is such that the cow dies on the spot. The matter winds up before a court. Before the case is decided, we cannot be sure what its outcome will be. It is possible that I will be found liable for negligence, or for killing, or fined for speeding, or a combination of the three, or that I will go home free. But it is

²¹ I ask the reader to abstract, for now, from principles such as *lex certa* or the American 'void for vagueness' doctrine (*Grayner v City of Rockford* (1972) 408 US 104), which would render the last provision unenforceable.

²² See Kaplow (n 1). On the possibility of the law being certain in this sense see Cass Sunstein, 'Problems with Rules' (1995) 83 California Law Review 953. For a different interpretation of legal certainty, see Mark Greenberg, 'How Facts Make Law' (2004) 10 Legal Theory 157, 162. I should note that ambiguously-worded laws are not the only potential source of uncertainty about the *outcome* of legal proceedings: an outcome might be uncertain because some facts are unknown, because it is unclear whether some evidence will be sufficient to prove the existence of a material fact, because private parties cannot gauge the quality of counsel they retain, because the identity and the ideological preferences of the judge are unknown, and so on. These may very well be related to vagueness: for example, if a law is perfectly clear, the ideology of the judge will not matter since he would enjoy no discretion. If a law is very vague, investments in legal representation may in the aggregate be higher because more depends on representation – see on this point Wickelgren and Friedman (n 1) 9. In this paper, I restrict myself to discussing only uncertainty which has to do with the way in which laws are drafted.

impossible to ascribe to any one of these eventualities a probability of one. Therefore, my legal position is uncertain.

1. *Applicative Uncertainty*

I opened with the proposition that there are two types of legal uncertainty and that the law-and-economics literature focuses on one but not the other. Let me start by describing the one which I believe to be amply covered. I call it applicative uncertainty. Suppose that I try to demystify my legal position by applying Laws A, B, and C to the facts of the case. I first check to see if I am liable to pay compensation under Law A. The word 'negligent' can mean one thing to me and quite another to someone else.²³ It is unclear whether driving while gesticulating at one's tearful girlfriend is negligent. Some may find it too dangerous, others may tolerate it, others still may say it is perfectly acceptable or even commendable. Consequently, I cannot tell with certitude whether I will be deemed to have been negligent.

I then move on to measure my behaviour against Law B, which prohibits killing. It is not in dispute that my actions have caused a death. But there may very plausibly be disagreement as to whether the killing of cows is criminal. Does 'death' in the sanctionability condition mean the death of a human or the death of any living being? Again, sensible persons may plausibly differ on the correct interpretation. I cannot know if I am liable.

Finally, was my speed unreasonable? This again depends on an unknown: unreasonable as judged by whom, against what, to which end? A speed of 45 km/h is acceptable to many. It gives pedestrians enough time to move out of the way when a vehicle approaches. Others may find it excessive – it evidently does not give distracted drivers enough time to swerve bovines. Each interpretation being as valid as the other, it is impossible to predict whether I will be fined.

What is the common trend? In all three cases, the law is vaguer than the facts. We may say that the law and the facts exist at different levels of generality,

²³ This is true of all words – see Ludwig Wittgenstein, *Philosophical Investigations* (GEM Anscombe tr, Oxford University Press 1953).

that of the law being much higher than that of the facts.²⁴ Applicative uncertainty arises because we cannot tell if the specific facts can be subsumed under the law. We lack information. The lawmaker has failed to provide us with enough premises to traverse the logical distance between his commands and reality.

Now, this is the very type of uncertainty which is ordinarily discussed under the rubric of rules-versus-standards. The point there is that the lawmaker can reduce the logical distance between laws and facts by producing more specific laws.²⁵ If, for example, the law says that 'if a person causes a death, they shall be liable to imprisonment', and then adds that 'death means the death of another human being', the additional premise makes the law less applicatively uncertain.²⁶ But the production of such additional premises requires the lawmaker to decide what it is that he wishes to prohibit: he must choose between a world where animals are killed and one in which they are not. To discriminate between the two alternatives, he needs information.²⁷

²⁴ The point is at its most explicit in Leonard Boonin, 'The Logic of Legal Decisions' (1965) 75 *Ethics* 179.

²⁵ For a model of law founded on this precept, see Vern R Walker, 'Discovering the Logic of Legal Reasoning' (2007) 35 *Hofstra Law Review* 1687.

²⁶ There is a contention here, to the effect that a premise can be as vague as the law which it defines, in which case applicative certainty remains the same. This is correct, provided that the premise is also circular. For example, Section 230(1) of the Employment Rights Act defines an employee as an individual who works under a contract of employment, and then Section 230(2) defines a contract of employment as a contract of service. Since service and employment mean the same thing, the effect is that Section 230(2) does not increase the applicative certainty of Section 230(1). Note, however, that this kind of duplication is rare, since the two vague laws must mean the exact same things. For example, if we are given a speed limit such as 'you must drive reasonably' and we are then told that 'reasonably means not too fast', then the two laws are obviously both vague. But the second still increases the applicative certainty of the first. Unreasonable driving can mean driving too slowly as well – the second law removes that interpretation from the possible set of interpretations, so that there are fewer possible interpretations in total, and the law becomes clearer.

²⁷ I use the term information very loosely, to denote both 'isses' and 'oughts'. For example, in the animal-killing example, the lawmaker would probably need to discover how many persons within the polity eat animals, how many animals are slaughtered, where the slaughter takes places, what methods are used to slaughter

Information being costly, it would be uneconomic to produce a set of premises sufficient for the resolution of all legal controversies. To do so, the lawmaker would have to acquire information about all possible states of the world and then decide which ones are desirable and which ones are not.²⁸ The cost would be prohibitive. Consequently, the optimally precise law is to some extent uncertain.

2. Hierarchic Uncertainty

I now turn to the novel approach that I advertised in the introduction. Uncertainty can also be hierarchic. It may be impossible, as between two (or more) laws, to tell which one has precedence. Suppose that I have again ran over your cow while arguing with my girlfriend, but we now have the following laws:

- (A) If a person negligently harms another, where negligence includes not paying attention to the road while driving a car, they shall have to pay compensation.
- (B) If a person causes the unlawful death of another, where another means any living creature, they shall be liable to imprisonment.

the animals, and so on. Those are simply descriptions of reality – which is what is ordinarily meant by the word information. But to decide whether to introduce this law, the lawmaker would also have to decide how to weigh the welfare of meat-eaters against the welfare of animals. He must decide whether 'humans take precedence over animals', or 'animals and humans are equal', or 'animals are more important than humans'. These statements are evaluative rather than descriptive – they do not have truth-values. While I do not deny this conceptual difference, I analyse them in a similar way. The reason is that to the lawmaker, both 'ises' and 'oughts' are costly. It would be easy for a dictator to decide whose welfare he prefers, but in most modern polities the lawmaker comprises numerous individuals whose policy preferences are different. It is costly for them to coordinate, which renders evaluative statements more like descriptive ones.

²⁸ The point, or a version thereof, is made in the authorities cited at n 1. There is also a very obvious analogy with the literature on incomplete contracts: for a flavour, see Benjamin Klein, Robert Crawford and Armen Alchian, 'Vertical Integration, Appropriable Rents, and the Competitive Contracting Process' (1978) 21 *Journal of Law and Economics* 297, 303.

- (C) If a person drives a car at an unreasonable speed, where unreasonable means in excess of 50 km/h in a city, they shall be liable to a fine.

The first point to note is that the law as it relates to the case under consideration is applicatively certain. It is beyond doubt that the killing of a cow falls within the remit of the amended Law B. I am definitely liable in tort. And having driven at a speed below that indicated in Law C, I am definitely not liable to a fine.

Does this make my legal position perfectly certain? I think not. The most pressing issue to me in practice is whether the death of your cow will cause me to be jailed. This in turn depends on whether the killing of the cow was 'unlawful'. The problem is that my behaviour is lawful when measured against Law C but unlawful when measured against Law A. My position, then, depends on the order of precedence between the speed limit and the negligence tort. If negligence trumps the speed limit, I go to prison; if the speed limit trumps negligence, I walk home free. But there is nothing in the text of either Law A, B, or C to suggest a ranking. The three laws are presumably equal. Therefore, one hierarchical ordering is as good as another. Accordingly, there is no way of telling whether I will go to prison. The law, even if applicatively determinate, remains uncertain.

Formally, we may define hierarchic uncertainty as the uncertainty that emerges when there are two (or more) laws which cover the same set of facts, but which point at mutually inconsistent outcomes. It may arise, as in the example just discussed, when laws refer to one another. It may also exist between a positive and a negative law. Take the following dyad:

- (A) If a person expresses his view freely, they shall not be persecuted in any way.
- (B) If a person blasphemes any God, they shall be liable to a fine.

It is easy to conceive of factual scenarios in which the two clash. Suppose that I am on trial for writing a newspaper article in which I claim to be a better weaver than Pallas Athena. My behaviour clearly falls within Law A. I am expressing a personal view. And my article is also quite clearly captured by Law B. It blasphemes the Greco-Roman pantheon. Whether I am liable does not depend on the construction of either law. It depends on the higher-order question of whether Law A is paramount to Law B. And since there is no

higher-order rule here telling the judge how to choose between Law A and Law B, the outcome of my trial is uncertain.

There is a lawyerly intuition here to the effect that the uncertainty is perfectly soluble. Since guarantees on free speech are usually contained in constitutions and prohibitions on blasphemy in statutes, and since constitutions usually trump statutes, a lawyer reading the foregoing passage might instinctively favour prioritising free speech over blasphemy, even if as in the hypothetical no such law is given.²⁹ Those intuitions are well-grounded, because laws which give the order precedence of other laws are a common feature of real-world legal systems. I said earlier that applicative certainty is solved by making laws more precise. Hierarchic uncertainty is solved through metalaws. In the hypothetical, we could not predict outcomes because we did not know the order of precedence between different laws. We would be able to make ironclad predictions if we had been given laws such as 'for the purposes of murder law, unlawful shall mean unlawful in the civil sense but not in the administrative sense', or 'free speech trumps blasphemy'. Note, however, that the statements to which I just adverted are not laws under the definition I adopted at the start. They do not contain any sanctionability conditions. They are laws about laws, or metalaws.

Hierarchic uncertainty thus arises when there is no metalaw, or when the metalaw is formulated vaguely. One may ponder why a rational lawmaker might skimp on metalaws. Metalaws are costly to produce. This might on first impression appear startling – it costs little to say that the constitution trumps statute, or that more specific laws trump more general ones. But that simplicity is illusory, since before such simple determinations can be made, one also has to consider the content of constitutions and those of statutes, and the content of specific laws and their more general counterparts, and so on. A choice has to be made between laws, and to make that choice the lawmaker needs information, and that information is costly.³⁰ In the

²⁹ Similarly, in the first example, the legally trained are instinctively moved to apply all three norms together, even though any ordering is permissible.

³⁰ This only holds if the choice is to be exercised meaningfully. It is possible to produce a metalaw costlessly by saying that 'if two laws clash, the one whose text is shorter prevails', or something to that effect. But to choose in this fashion, the lawmaker

examples that I gave earlier, the lawmaker would have to decide whether in the regulation of homicide civil law standards are more appropriate than administrative law ones. He would also have to choose between religion and freedom of speech as social values and between unlawfulness and intent as benchmarks of criminal culpability.³¹ Choices such as these are costly to make, information-wise. Once this feature of the problem is accounted for, it becomes apparent that it is not sensible to attempt the complete elimination of hierarchical uncertainty. Like the optimal law, the optimal metalaw is not perfectly precise.

I discuss the optimal precision of metalaws in much more detail elsewhere.³² Here, I limit myself to highlighting hierarchic uncertainty as being distinct from its applicative counterpart. There are, I think, two critical differences. Firstly, applicative uncertainty only concerns the relationship between one law and one set of facts. Hierarchic uncertainty operates between a set of facts and two laws. Defamation by itself cannot be hierarchically uncertain. It only becomes so once combined with the protection of free speech.

Secondly, applicative uncertainty is resolved by reducing abstraction; hierarchic certainty is achieved by increasing it. If we have the law 'murder is the killing of another', its applicative uncertainty is reduced by adding to it the premise 'another means human being'. The second premise is specific to the first – the new law is more precise than the old. If we have the laws 'free speech is guaranteed' and 'defamation is prohibited', their hierarchic indeterminacy is resolved by adding that 'free speech trumps defamation'. The last premise is more general than the first two. The body of laws in question gains an additional layer of abstraction.

would have to be indifferent to the real-world implications of the law, which would in turn make it difficult to explain why he is producing laws in the first place.

³¹ The problem is greatly compounded when reference is made to some changeable body of law. The lawmaker, in order to determine the appropriate hierarchy, would have to consider not only civil and administrative standards as they are at the time of setting the hierarchy, but also as they might become in the future.

³² Orlin Yalnazov, 'Metarules versus Metastandards' (under review, manuscript available on request).

3. Penalty Clauses: An Example

The analysis has so far been built around highly unrealistic hypotheticals. For this reason, I will now pause for an example. I propose to discuss the rule on penalty clauses in English law. I will explain why it is uncertain, and how it can be rendered more certain. I will then proceed to consider it in conjunction with another rule. I do this in the hope of demonstrating how hierarchic uncertainty may arise and how it differs from applicative uncertainty.

One of the earliest statements of the penalty clause rule comes from *Lord Elphinstone v The Monkland and Iron Company*.³³ In an oft-cited dictum, Lord Halsbury described a penalty clause as a contractual provision designed to be enforced *in terrorem*.³⁴ For convenience, I will write this out as a sanctionability condition or, in Hart's language, a primary rule:

Penalty Clause Law: If a promisor and a promisee enter into a contract and that contract contains a clause which is designed to be enforced *in terrorem*, the promisor will suffer no sanction under that clause.

Let me now show how that law is applicatively uncertain. It has a core and a penumbra. I propose to do so by the reference to the facts of *ParkingEye v Beavis*.³⁵ Beavis had parked his car at a parking lot operated by ParkingEye. The contract incorporated a clause to the effect that the claimant was allowed to use the parking space for two hours. Breach of that term would trigger a charge of £85. Beavis overstayed by fifty-six minutes. He then refused to pay the charge. ParkingEye sued and at trial Beavis contended that the clause was a penalty. The law would be applicatively certain if we could match the facts of *ParkingEye* to an outcome with a probability of one.³⁶ I do not think we can. Consider first some of the possible applications of the Penalty Clause Law to the facts:

³³ *Lord Elphinstone v The Monkland and Iron Company* (1886) 11 App Cas 332.

³⁴ Ibid 348.

³⁵ *ParkingEye v Beavis* [2015] UKSC 67. The case was decided along with *Cavendish Square Holding v Makdessi* and is often cited under that name.

³⁶ Of course, I base this on the assumption that we do not know the outcome of *ParkingEye* itself.

- (1) The judge may say that the charge of £85 is vastly disproportionate to the loss likely to have been suffered by ParkingEye from Beavis overstaying by fifty-six minutes, so that the clause was intended to terrorise him into compliance. In that case, Beavis is not liable.
- (2) The judge may say that the charge of £85 was intended to fund the operation of the parking lot, and although it may have had the incidental effect of scaring Beavis into compliance, that was not its main purpose. Beavis is liable.
- (3) The judge may say that although £85 is a disproportionate amount, the word 'terror' indicates a degree of fright far beyond that experienced by the average citizen when faced with a large parking charge. Beavis is liable.
- (4) The judge may say that even though £85 is a very high amount, it is possible that Beavis (or some other user) would value the right to remain at the parking lot for an additional fifty-six minutes at £1,000, so that they would be more than happy to pay £85. Beavis is liable.

The four interpretations are equally valid.³⁷ Each is a perfectly sensible way of measuring the facts against Lord Halsbury's dictum. We can only speculate about the choice that would be made if the facts of *ParkingEye* were to come before some particular judge. If we assume that one of the four interpretations listed here will eventually be chosen, then Beavis is liable under (2), (3), and (4). He is not liable under (1). Beavis is liable with a probability well short of 1.³⁸

³⁷ It is probably not difficult to think of more than four interpretations. But since parking charges can only hold so much fascination, I limit myself to the ones in the main text. For the purposes of my argument, it is entirely sufficient that there exists at least one interpretation under which Beavis is liable and at least one under which he is not.

³⁸ If the four interpretations that I set out were exhaustive and each was as likely to be chosen at trial as the other, he would be liable with a probability of 0.75. This is of course a very crude way of calculating legal uncertainty: in reality, there will be many other factors involved (see n 22 above). And I do assume that the interpretations that I have set out here are the only possible ones, which is very likely not true in practice. This does not, I think, detract too much from the utility of the example – I merely

How can this be fixed? The law becomes more certain – its penumbra shrinks – if the confidence with which a prognosis can be made is increased. The elimination of possible interpretations reduces uncertainty. When only one interpretation remains, the law is certain. Suppose that the Penalty Clause Law is amended as follows:

Amended Penalty Clause Law: If a promisor and a promisee enter into a contract and that contract contains a clause which is designed to be enforced *in terrorem*, where *in terrorem* means disproportionate to harm actually suffered, and where terror means any degree of fear irrespective of its severity, and irrespective of whether terror was the sole intent of the party relying on the clause, notwithstanding the possibility that the promisor is in fact happy to pay rather than terrorised, then the promisor will suffer no sanction under that clause.

This dispenses with interpretations (2), (3), and (4), leaving only (1). We prune interpretations by saying what the word 'terror' does and does not mean, that is by defining it. The individual definitions, both positive and negative, are more specific than the word 'terror' itself.

Why might a rational lawmaker choose the original Penalty Clause Law over the Amended Penalty Clause Law? He may do so in order to economise on information costs. For example, the definition that I adopted above expressly sanctions the possibility that a promisor who valued the benefit from breach more highly than the penalty will nonetheless be able to escape sanction. It is unclear whether this is desirable. In order to decide whether it is, the lawmaker must inform himself of the relevant aspects of the problem, its prescience in society, and so on.³⁹ Acquiring such information is costly. The instances in which promisees value the use of parking lots at more than £85 are rare. The outlay in information is wasteful. It would have been better to keep the law vague.

assume that we are given no other information other than that in the example. I do this to show how the particular types of uncertainty may be reduced, and what the interplay between the two is – those connections would be more obscure in a more realistic example.

³⁹ For an accessible overview of that debate, see Ugo Mattei, 'The Comparative Law and Economics of Penalty Clauses in Contracts' (1996) 43 *American Journal of Comparative Law* 427.

The foregoing concerns applicative uncertainty, taken in isolation. What I propose to do now is take Lord Halsbury's *in terrorem* dictum in conjunction with Lord Dunedin's pronouncement in *Dunlop Pneumatic Tyre v New Garage and Motor Company*.⁴⁰ There Lord Dunedin said that a clause is enforceable if it stipulates liquidated damages, that is, some realistic estimate of the loss likely to be suffered by the promisee. The dictum in full reads:

The essence of a penalty is a payment of money stipulated as *in terrorem* of the offending party; the essence of liquidated damages is a genuine covenanted pre-estimate of damages.⁴¹

As before, I will formulate the law as a complete sanctionability condition, and I will consider it in conjunction with that announced in *The Monkland and Iron Company*:

Penalty Clause Law: If a promisor and a promisee enter into a contract and that contract contains a clause which is designed to be enforced *in terrorem*, the promisor will suffer no sanction under that clause.

Liquidated Damages Law: If a promisor and a promisee enter into a contract and that contract contains a clause which is designed to be a genuine covenanted pre-estimate of damages, then the promisor will be sanctioned as per the provisions of that clause.

Here one law defines a positive sanctionability condition and the other a negative one. The Liquidated Damages Law points to enforceability, while the Penalty Clause Law points to invalidity. The two overlap. A law can be designed to be enforced *in terrorem* while being a genuine pre-estimate of damages. Suppose that we modify the facts of *ParkingEye* slightly, so that the claimant can prove that £85 is a commercially justifiable estimate of its loss, say because even though its loss from individual infractions is small, its aggregate expenditure on administering and enforcing the system is large. At

⁴⁰ *Dunlop Pneumatic Tyre v New Garage and Motor Company* [1915] AC 79.

⁴¹ *New Garage* (n 40) 86, citing *Clydebank Engineering & Shipbuilding v Don Jose Ramos Yzquierdo y Castaneda* [1905] AC 6. The phrase 'genuine covenanted pre-estimate of damages' denotes the idea that the parties foresaw the sum in question as a compensation of loss actually likely to be suffered, rather than as a penalty that one would have to pay to the other in the event of breach. There is a loose analogy between this and the distinction between compensatory and punitive damages: if the sum was contemplated as compensation, it is valid; if it is punitive, it is not.

the same time, for the drivers who do get caught, an £85 fine may be said to be *in terrorem*, say because average earnings in the area are £200 a week. The controversy then turns on whether the Penalty Clause Law takes precedence over the Liquidated Damages Law, or vice versa. There are two possible orderings:

- (1) The Penalty Clause Law applies only when it is consistent with the Liquidated Damages Law, in which case the parking charge clause is enforceable.
- (2) The Liquidated Damages Law applies only when it is consistent with the Penalty Clause Law, in which case the parking charge clause is unenforceable.

We do not know if ordering (1) or (2) is correct. Without a metalaw, or in Hart's terms a secondary rule, the uncertainty is irreducible. The outcome becomes a matter for the individual judge. The lawmaker may resolve the problem by formulating a metalaw, such as:

Metalaw: Where two laws clash, that which points to validity shall prevail.

This serves to eliminate ordering (2). The law becomes certain. Might a rational lawmaker have good reason not to produce such a metalaw, or to produce a vague one, such as 'where two laws clash, that which points to validity shall *usually* prevail'? In order to commit himself to a specific metalaw, the lawmaker must also make a choice between the sanctity of contracts and the protection of parties with asymmetric bargaining power. Choosing between the two entails an investment in information. Making a choice is not costless – the lawmaker would have to decide whether he wants the court to upset contracts more often. He would have to study the potential economic impact of one solution or the other, its fairness implications, the availability of mechanisms for circumventing the prohibition, their cost, and so on. The evaluation, and the acquisition of the germane information, have a positive cost. The cost of information will sometimes preclude the lawmaker from specifying a metalaw. Therefore, the optimal metalaw will sometimes be less than perfectly specific.

I said that applicative and hierarchic uncertainty differ in two respects, namely origin and solutions. Regarding the former, recall that when I was discussing the Penalty Clause Law in isolation, a problem arose because the

facts of *ParkingEye* were specific and the Penalty Clause Law vague. The two could not be rendered commensurate. The Penalty Clause Law and the Liquidated Damages Law, on the other hand, are problematic because they coexist. If there had been no Liquidated Damages Law, there would have been no need to rank it against the Penalty Clause Law, and vice versa.

As regards solutions, recall that we made the Penalty Clause Law applicatively certain by adding further premises that were specific to the pre-existing ones. We solved hierarchic uncertainty in the two-law example by introducing a metalaw. The metalaw was more general than both laws. If at first a body of law contains a Penalty Clause Law and a Liquidated Damages Law and later that same body of law combines those two laws with the Metalaw, we may say that the body of law in question has become more, rather than less, abstract. Despite the increase in abstraction, however, the law become less uncertain.

III. THE RELATIONSHIP BETWEEN APPLICATIVE AND HIERARCHIC INDETERMINACY

I will now try to postulate a relationship between applicative and hierarchic uncertainty. To this end, I argue that legal complexity increases, applicative uncertainty decreases and hierarchic uncertainty increases. I will first say what I mean by complexity. Thereafter, I will examine the connection between complexity as I define it and the two kinds of uncertainty. Lastly, I will proffer some very general arguments about the shape and structure of the optimal system.

1. Complexity

The word 'complexity' in general parlance denotes complicatedness. It might not be immediately obvious why I associate it with lower applicative uncertainty. The reason is that I use the term in a more technical sense. Complexity here denotes the existence of several elements within a system interacting with one other.⁴² Since we are here concerned with laws, we may say that a legal system is complex when it comprises multiple laws that relate

⁴² For a more detailed exposition see Herbert Simon, 'The Architecture of Complexity' (1962) 6 Proceedings of the American Philosophical Society 467.

to one another. This definition is meaningless in a static sense. All bodies of law comprise more than one element and those elements always relate to one another, so all law is complex. I however use the term dynamically: a legal system becomes more complex if the number of elements that compose it increases, and less complex when that number decreases.⁴³

I now return to the proposition that if complexity increases, applicative uncertainty decreases and hierarchic uncertainty increases. We may say that if the number of laws within a body of law increases, each individual law will become easier to apply, but the laws in their totality will become more difficult to systematise, that is, it will become more difficult to decide which law to apply to any particular set of facts. Recall the hypothetical in which I ran over your cow while arguing with my girlfriend. I referred to two versions of a negligence law, one determinate in respect of those facts and the other indeterminate.

Indeterminate Law: If a person negligently harms another, they shall have to pay compensation.

Determinate Law: If a person negligently harms another, *where negligence means not paying attention to the road while driving a car*, they shall have to pay compensation.

The italicised part specifies the law. The definition of negligence also increases that law's complexity: where before we had only a law, we now have a law and a definition. We may equally say that we have two laws, one which says that negligence triggers compensation and another that negligence includes not paying attention to the road.⁴⁴ It may be thought that the

⁴³ Consider a fictional legal system which comprises one law - 'civil wrongs shall be compensated' - and the *Bürgerliche Gesetzbuch* (BGB), which makes the same point in 2385 Articles, many of which have sub- and sub-sub-clauses. We may say that the former statement is complex, in that it comprises two elements – civil wrongs and compensation. And evidently, so is the BGB. But the BGB is more complex than the original statement, plainly because it comprises many more elements.

⁴⁴ In the examples I use in this paper, I take a somewhat cavalier attitude to the issue of whether something is 'one' or 'two' laws. This is because the point is purely semantic and does not impact on the veracity of my conclusions. Let me take an example I used earlier – suppose that blasphemy is an offence and also that freedom of speech is guaranteed. We may formulate this like two laws, like I did before, and speak of hierarchic indeterminacy. We may also formulate the two provisions as a

Determinate Law here is perfectly certain. Indeed, it is certain when it is measured against the facts that I used before. But let us suppose that another case comes to trial in which your neighbour's cow was mowed down by my neighbour, who was having a heart attack while driving. We now have hierarchic uncertainty between the law and the premise. The law dictates that negligent behaviour should be sanctioned. No-one would say that a person who is having a heart attack is negligent for failing to watch the road.⁴⁵ But the premise defines negligence as failure to watch the road. The facts fulfil that condition.

Perhaps it is helpful to rewrite the Determinate Law as two laws:⁴⁶

- (A) If a person negligently harms another, they shall have to pay compensation.
- (B) If a person fails to pay attention to the road, they shall have to pay compensation.

The two laws define sanctionability conditions, but their interrelationship is unclear. It might be that (A) and (B) have to be satisfied for there to be

single law, such as 'Freedom of speech is guaranteed and blasphemy is prohibited'. This is generally mimicked in the way lawyers speak of legal doctrines. For example, an English lawyer might describe 'the law of murder' as comprising the elements of the offence and defences such as diminished responsibility. Or he might describe 'the law of diminished responsibility' which includes elements such as substantial impairment, rational judgment, and self-control. In the former case, 'murder' is treated as one law. In the other, it is treated as being separate from defences.

⁴⁵ This to some extent mirrors the argument in n 22. The difference is that Sunstein argues that to hold the stroke victim liable would be absurd, and this would make the law uncertain. My point is that even if we accept this absurdity and proceed to convict, the law will still be uncertain – the goal of perfect applicative certainty not only produces absurdity, but it also fails to guarantee that the law will become entirely certain.

⁴⁶ The reader might suspect simulation: to make my point, I transform one law into two. Is there any artifice involved? This would be true if there is some fundamental difference between the meaning of 'if negligence then compensation, where negligence means failing to watch the road' and the meaning of 'if negligence then compensation; if failing to watch the road then compensation'. The only difference is that the first formulation conceals the hierarchic indeterminacy through syntax: the clause 'where' introduces the premise, but it does not make clear its relationship with the law – does 'where' mean 'if' or 'if and only if' or 'and' or 'or'.

liability, or that (A) or (B) will suffice, or (A) so far as it is consistent with (B), or (B) so far as it is consistent with (A). There being no metalaw, we cannot predict the outcome. Generalising, we may say that when we specify a law, we add more words to it. Since those words are vague, they overlap with the words we used before. The more we add, the greater the overlap, and the greater the resultant uncertainty.

2. *Metalaws and the Legal System*

I have so far strived to show that applicative certainty translates into hierarchical uncertainty. I have not said anything about the magnitude of the problem. It could be that a small increase in complexity brings marginal gains in applicative certainty while greatly amplifying hierarchical uncertainty. Or it might be that the gains in applicative certainty are large and the loss of hierarchic certainty negligible. My argument in this respect is that hierarchic uncertainty can be expected to rise very dramatically with every additional law, whereas investments in applicative certainty are likely to exhibit diminishing returns.⁴⁷ And this will eventually lead me to conclude that for the system to be able to sustain complexity, the lawmaker must produce very precise metalaws.

⁴⁷ To make the law more applicatively certain, the lawmaker must invest in information. For example, it is informationally cheap to say 'drive reasonably' – the lawmaker must only know that there are some dangers from unreasonable driving. It is much more expensive to say 'drive at a speed lower than 50 km/h'. This requires the lawmaker to enumerate the dangers of driving fast, to balance them against the inconvenience from slower transportation, and to strike a balance between the two. Returns diminish because the reduction in applicative uncertainty becomes smaller with every additional clarification: we gain a lot in terms of prognostics from switching from 'driving reasonably' to 'driving under 50 km/h', but less from switching from 'drive under 50 km/h' to 'drive under 50 km/h in cities unless near a school, when you should drive at 30 km/h', less still if we switch from 'drive under 50 km/h in cities unless near a school, when you should drive at 30 km/h' to 'drive under 50 km/h in cities unless near a school, when you should drive at 30 km/h or unless visibility is poor, when you should drive at 20 km/h'. In each step, applicative uncertainty decreases, more and more facts are brought within the scope of the law. But the decrease is in each step smaller than in the previous – hence the point that investments in information exhibit diminishing returns.

It is best to begin by considering a metalaw-free system. Recall that hierarchic uncertainty arises when we are given some laws, but not their order of precedence. The probability that a particular ordering will govern in a case is given by dividing one by the number of possible orderings. And the number of possible orderings, absent metalaws, depends on the number of laws that can potentially be applied to a given case. If we are given one law, there is hierarchic certainty. If we are given two laws, A and B, they can be ordered as either A-B or B-A. Any prediction about their likely order of precedence is true with a probability of 0.5. If we are given three laws, A, B, and C, then there are six potential orderings – A-B-C, A-C-B, B-A-C, B-C-A, C-A-B, C-B-A. Any prediction is true with a probability of 0.16. If we are given four, the number of orderings rises to twenty-four, giving a probability of 0.04. Five laws can be ordered in one-hundred and twenty ways, and so on and so forth. The number of potential orderings is the factorial of the number of laws.

Let us now consider the role of metalaws. With perfectly precise metalaws, there would be no hierarchic uncertainty – the lawmaker's desired ordering would always be known in advance. A set of perfectly precise metalaws would be very expensive to produce – the lawmaker would have to anticipate every case in which two or more laws might overlap, and he would have to choose between these alternative laws for each set of facts that he identifies. A vaguer metalaw will say that one ordering will govern with some probability p and others will govern with a probability $1 - p$.⁴⁸ What becomes critical then is the probability that the default ordering will apply – the higher the value of p , the higher the hierarchic determinacy of the conflicting laws.

⁴⁸ Real-life metalaws do not, of course, tie the application of one law or another to probabilities. For example, in precedential systems, we are told that (A) previous decisions of the courts are binding, unless there is a good reason to depart from them, and (B) if statute and precedent clash, statute prevails. (A) is a vague metalaw: it establishes a presumption that previous judgments are binding, but that presumption can be rebutted 'when it is right to do so' (*Practice Statement* [1966] 3 All ER 77). Since the criterion for departure is vague, we may say that the presumption will govern with a probability of p and the exception with a probability of $1 - p$. (B) on the other hand is precise: statute *always* prevails. Neither metalaw speaks directly of probabilities. But for both, it is possible to say that the application of a particular law is specified with some probability, and also that that probability varies between the two metalaws.

For every additional law, applicative uncertainty decreases. The question is by how much. The literature on applicative certainty suggests that the process exhibits diminishing returns.⁴⁹ If the most general law is 'if negligence then compensation', then specifying this through 'negligence includes not paying attention to the road' makes the law considerably more certain. But if we then go on to specify this further to say that 'not paying attention to the road includes driving with one's eyes closed', the gain is smaller: the second clarification prunes fewer interpretations than the first.

What, then, is the optimal number of laws? Recall that every additional law decreases applicative uncertainty but increases hierarchic uncertainty. Therefore, we may only increase the number of laws up to the point where the loss of hierarchic certainty exceeds the corresponding gain in applicative certainty. Beyond that point, further specification will result in a gain in applicative certainty which is lower than the corresponding loss of hierarchic certainty, causing uncertainty, in the aggregate, to increase rather than decrease. Accordingly, increasing complexity beyond that point will be counter-productive. In mathematical terms, if we were to plot applicative

⁴⁹ The point is at its most explicit in Ehrlich and Posner (n 1) and Korobkin (n 1). It might appear that there is a counter-argument to the effect that as refinements and exceptions are introduced to some general standards, it becomes harder for the judge to determine whether the case falls under the standard or under its exceptions. For example, we may at first have a law which prohibits driving at a speed in excess of 50 km/h, and thereafter we may introduce an exception to the effect that necessity is a defence to a charge of speeding. Once the necessity defence is introduced, there is the possibility that counsel in most cases will argue that the defendant needed to speed. The point is made at length in Sunstein (n 22). It does not refute my point here. First, in my framework, the introduction of the necessity defence simply makes the law more hierarchically uncertain, and the loss of hierarchic certainty outweighs the gain in applicative certainty. We do know something which we did not know before, namely that the limit is subject to a defence, but we also face the problem of the offence and the defence overlapping. If this is not accepted, then it is also possible to say that the necessity defence increases the cost of litigation rather than uncertainty: as cases are decided, it will become apparent what 'necessity' means, and once there is clarity, the defence will only be pleaded where it has a realistic chance of success. Litigation around that point will of course be costly, but the cost of litigation is not the same as the cost of information to the lawmaker.

and hierarchic certainty as functions of the number of laws, then the optimum point would be the intersection of the two curves.

Figure 1: The Relationship between Applicative and Hierarchic Uncertainty

Now, without metalaws, hierarchic uncertainty rises factorially. The decline in applicative uncertainty is not so sharp. We may therefore expect that the optimal number of laws would be very low. A very abstract law would be preferable to a very specific one. The reader may validly ponder how it is possible for real legal systems to be as complex as the *Bürgerliches Gesetzbuch* (BGB) or the common law of property. I refer back to the discussion of metalaws – the slope of the hierarchic certainty curve is influenced by the probability with which a metalaw designates one ordering as being valid. If the metalaw is very certain, that is, if it designates some ordering as being valid with a very high probability, the law can be very complex. Two possibilities then emerge – either modern legal systems are radically indeterminate, that is, the outcome of any one case is entirely a matter for the judge,⁵⁰ or they contain very precise metalaws.⁵¹

⁵⁰ That would be the legal realist interpretation: see Duncan Kennedy, 'Form and Substance in Private Law Adjudication' (1976) 89 Harvard Law Review 1685.

⁵¹ I should make a clarification here. I do not claim that any law is hierarchically indeterminate when set against another. If we say that speeding triggers a fine and also that playing loud music at night is prohibited, the two are hierarchically

I believe the latter to be true, if anything as a matter of empirics. German justice is predictable because the BGB contains a rigid hierarchy. The same is true of the common law of property. Given the high number of laws in real legal systems and the fact that hierarchic uncertainty rises at a more-than-exponential rate, a very small change in the precision of the metalaw can have a tremendous impact on aggregate certainty. Metalaws, if I am right, set the boundaries of legal complexity.

3. *Penalty Clauses: A Second Example*

I will now try to illustrate this theoretical point. I will begin with a very brief history of the development of the liquidated damages rule in the last century. I will then show how the various points raised in the preceding section apply to it. Let us take as our starting point the dicta of Lord Dunedin and Lord Halsbury in *Dunlop* and *Monkton* respectively. The reader will recall that I formulated these as two laws, namely:

(To) Penalty Clause Law: If a promisor and a promisee enter into a contract and that contract contains a clause which is designed to be enforced *in terrorem*, the promisor will suffer no sanction under that clause.

Liquidated Damages Law: If a promisor and a promisee enter into a contract and that contract contains a clause which is designed to be a genuine covenanted pre-estimate of damages, then the promisor will be sanctioned as per the provisions of that clause.

In *Cooden Engineering v Stanford*,⁵² a car hire-purchase agreement provided that breach would entitle the claimants to repossess the car *and* to all outstanding instalments. The claimants argued that this was justified because the defendant Stanford had only paid eight instalments out of twenty-four that had been due. Lord Somervell rejected the argument. He was of the view that the determination of validity ought to be made at the time of

determinate because speeding cannot constitute playing loud music, and vice versa. Likewise, it is impossible to conceive of a factual situation in which all of the *Bürgerliches Gesetzbuch*'s 2385 Articles are potentially applicable, or of a case which requires the application of every single rule of the common law of property. My argument is only that every law is applicable to a case with some probability. Consequently, as the number of laws in a given set increases, so does the expected number of laws which are applicable to individual cases.

⁵² *Cooden Engineering v Stanford* [1953] 1 QB 86.

contracting. The defendant's actions after the contract had been signed could not be used to show that the clause was not penal.⁵³ Let us now rewrite the laws above to incorporate this proposition:

(T1) Penalty Clause Law: If a promisor and a promisee enter into a contract and that contract contains a clause which is designed to be enforced *in terrorem*, *where the operation of the clause is to be judged at the time at which the contract was made*, the promisor will suffer no sanction under that clause.

Liquidated Damages Law: If a promisor and a promisee enter into a contract and that contract contains a clause which is designed to be a genuine covenanted pre-estimate of damages, *where the operation of the clause is to be judged at the time at which the contract was made*, then the promisor will be sanctioned as per the provisions of that clause.

In *Campbell*,⁵⁴ the claimants argued that the disputed clause was not penal because the contract described it as 'agreed compensation'. They also argued that the defendant could not have felt any actual terror upon breaching. On the first point, Lord Racliffe said that 'the intention of the parties themselves is never conclusive and may be overruled or ignored if the court considers that even its clear expression does not represent 'the real nature of the transaction' or what 'in truth' it is to be taken to be'.⁵⁵ On the second, his Lordship pointed out that 'penalties may quite readily be undertaken by parties who are not in the least terrorised by the prospect of having to pay them and yet are, as I understand it, entitled to claim the protection of the court when they are called upon to make good their promises'.⁵⁶ Two additions were thus made to the law:

(T2) Penalty Clause Law: If a promisor and a promisee enter into a contract and that contract contains a clause which is designed to be enforced in *terrorem*, *where in terrorem does not mean actual terror or fear* and where the operation of the clause is to be judged at the time at which the contract was made, *irrespective of the description which the parties give to it*, the promisor will suffer no sanction under that clause.

Liquidated Damages Law: If a promisor and a promisee enter into a contract and that contract contains a clause which is designed to be a genuine

⁵³ *Cooden Engineering* (n 52) 94.

⁵⁴ *Campbell* [1962] AC 620.

⁵⁵ *Ibid* 622.

⁵⁶ *Ibid*.

covenanted pre-estimate of damages, where the operation of the clause is to be judged at the time at which the contract was made, *irrespective of the description which the parties give to it*, then the promisor will be sanctioned as per the provisions of that clause.

Then came *Imperial Tobacco*.⁵⁷ There, Lord Wright observed that 'a millionaire may enter into a contract in which he is to pay liquidated damages, or a poor man may enter into a similar contract with a millionaire, but in each case the question is exactly the same'.⁵⁸ From this dictum we may distil the proposition that the wealth of the parties does not enter into the *in terrorem* test.

(T3) Penalty Clause Law: If a promisor and a promisee enter into a contract and that contract contains a clause which is designed to be enforced in *terrorem*, where in *terrorem* does not mean actual terror or fear and where the operation of the clause is to be judged at the time at which the contract was made, irrespective of the description which the parties give to it *and irrespective of their wealth*, the promisor will suffer no sanction under that clause.

Liquidated Damages Law: If a promisor and a promisee enter into a contract and that contract contains a clause which is designed to be a genuine covenanted pre-estimate of damages, where the operation of the clause is to be judged at the time at which the contract was made, irrespective of the description which the parties give to it, then the promisor will be sanctioned as per the provisions of that clause.

In *Murray v Leisureplay*,⁵⁹ the defendants argued that a provision guaranteeing a year's salary to its director in the event of dismissal was a penalty clause – it made them afraid to dismiss him. That argument was rejected because the clause was 'commercially perfectly justifiable'.⁶⁰

(T4) Penalty Clause Law: If a promisor and a promisee enter into a contract and that contract contains a clause which is designed to be enforced in *terrorem*, where in *terrorem* does not mean actual terror or fear and where the operation of the clause is to be judged at the time at which the contract

⁵⁷ *Imperial Tobacco of Great Britain and Ireland v Parslay* [1936] 2 All ER 515.

⁵⁸ *Ibid* 523.

⁵⁹ *Murray v Leisureplay* [2005] IRLR 946.

⁶⁰ *Ibid* at para [14], citing Lord Woolf in the Privy Council case *Philips Hong Kong v AG of Hong Kong* (1993) 61 BLR 49, 58-9.

was made, irrespective of the description which the parties give to it and irrespective of their wealth, *whereas a clause which is commercially justifiable is not designed to be enforced in terrorem*, the promisor will suffer no sanction under that clause.

Liquidated Damages Law: If a promisor and a promisee enter into a contract and that contract contains a clause which is designed to be a genuine covenanted pre-estimate of damages, where the operation of the clause is to be judged at the time at which the contract was made, irrespective of the description which the parties give to it, *where genuine covenanted pre-estimate means a commercially justifiable provision*, then the promisor will be sanctioned as per the provisions of that clause.

Finally, let us consider the case of *ParkingEye*. There, the Supreme Court held that the law had become too entangled and formulated a new approach. Lord Mance, with whom the rest of their Lordships agreed on this point, said that whether a clause is enforceable depends on 'whether the sum or remedy stipulated as a consequence of a breach of contract is exorbitant or unconscionable when regard is had to the innocent party's interest in the performance of the contract'.⁶¹ We may formulate the resultant law thus:

(T5) Enforceability Law: If a promisor and a promisee enter into a contract and that contract contains a clause which stipulates an exorbitant or unconscionable remedy, where exorbitance and unconscionability are to be measured against the innocent party's interest in performance, then the promisor shall suffer no sanction under that clause.

I now come to the substantial points. I will discuss T₀-T₄ first, since there is a rather obvious pattern, and then I will speak of T₅, in which that pattern is reversed. First, between T₀ and T₄, the law becomes increasingly complex. At each step, a new element is added to the ones already present. A rather simple way of grasping this is to evaluate the position of a person minded to familiarise herself with the law of penalty clauses at T₀ and at T₄. Their task would obviously be much more laborious at T₄.

Second, I argued previously that as complexity increases, so does applicative certainty. We may, at T₀, ponder whether a person who was perfectly secure in his future when he made a contract but was reduced to fright at the time of breach would be able to escape the provision. At T₁, we know that this is not

⁶¹ *ParkingEye* (n 35), para 255.

the case – a further set of facts has been brought within the core of the provision. Likewise, we may at both T_0 and T_1 wonder whether the fact that the parties describe something as compensation matters – at T_2 , we know that wording and mentation are irrelevant. At T_0 , T_1 , and T_2 we do not know whether the parties' wealth makes a difference, but at T_4 we know to disregard it. In Hartian terms, with each step the core of the provision expands and its penumbra shrinks.⁶²

Third, I contended earlier that specification exhibits diminishing returns. We may say that this is true if the totality of facts brought within the core of the provision between T_0 and T_1 is larger than that brought within its core between T_1 and T_2 , and that one is larger than that between T_2 and T_3 , and so on. Does this hold here? I believe so. The temporal qualification in T_1 has the effect of reducing uncertainty a great deal. A clause may be *in terrorem* at some time between breach and suit and not in other times. A huge number of interpretations are possible at T_0 . At T_1 , all but one are rendered obsolete – the time of the determination is fixed. Between T_3 and T_4 , conversely, we know a lot about what a penalty clause is – there remains only a small residue of cases in which a commercially sound clause may still be thought to potentially contravene the rule. That issue is clarified by the T_4 law, but the gain in certainty is surely lower than that between T_0 and T_1 .

I cannot, however, show this conclusively. To do so I would have to draw up a complete list of possible interpretations and discuss the applicability of each – in other words, I would have to deliver a complete representation of all possible states of the world. Information costs prevent me from doing this in the same way in which they prevent the lawmaker from writing infinitely complete laws. I nonetheless think that the argument has some intuitive appeal. One counterargument is that if facts such as those of *Murray* had come to be decided before those of *Cooden*, then the position would have been reversed: a greater gain in applicative certainty would have been made at a later stage. It is possible to argue that had the facts of *Murray* come before the courts first, there would have been argument made by counsel about the time at which a clause ought to be assessed. This is, of course, speculative. Even if it is not accepted, there is also an argument from the law-and-economics literature to the effect that the more (applicatively) uncertain the

⁶² Hart (n 23) 607.

law, the likelier it is to be litigated.⁶³ The idea, summarised, is that if parties have the same expectations about the outcome of a lawsuit, they will always settle, since to settle is cheaper than it is to litigate. Consequently, litigation occurs because of the divergent expectations about outcomes. The more uncertain the law, the higher that divergence, and the greater the likelihood that the parties will litigate. Thus, a highly uncertain law – such as one that prohibits penalty clauses but does not fix a temporal reference point for that determination – will attract much litigation, whereas a law that does provide for such a temporal reference point but not for commercial justifications will attract less. This is of course a *ceteris paribus* argument, so that exceptions will exist, but in general it might be expected to hold.

Coming now to T₅, it would appear that at that point the law becomes less certain than it was at T₄ – there is an obvious reduction in complexity. The gains made between T₀ and T₄ are reversed. If, for example, we try to apply the T₅ law to a situation in which the parties have described their agreement as 'agreed compensation', there is nothing in the text of the T₅ law to guide us. Does this mean that T₅ is sub-optimal? I think not. Although the law at T₄ is very certain applicatively, it is also very uncertain hierarchically. How so? Even at T₀, we do not know what happens when a clause is simultaneously *in terrorem* and a genuine pre-estimate of loss. Let us now add to this the T₄ provisos that a clause which is *in terrorem* cannot have been commercially justifiable and that a genuine pre-estimate of loss is commercially justifiable.

Law 1: An *in terrorem* clause is unenforceable.

Law 2: A genuine estimate of loss is enforceable.

Law 3: A commercially justifiable clause is enforceable.

Like at T₀, we do not know what happens if a clause is *in terrorem* and a genuine pre-estimate. But to this, Law 3 adds further perplexities. For instance, what happens if a clause is *in terrorem*, not a genuine pre-estimate, but it is commercially justifiable? The introduction of a third law expands the

⁶³ See, for example, Guiseppe Dari-Mattiacci and Bruno Deffains, 'Uncertainty of Law and the Legal Process' (2007) 163 *Journal of Institutional and Theoretical Economics* 627.

set of possible orderings. Recall that with two laws, there were two possible orderings. With three laws, there are six:

- (1) If a clause is *in terrorem*, it is unenforceable, irrespective of whether it is a genuine estimate or commercially justifiable.
- (2) If a clause is a genuine preestimate, it is enforceable, irrespective of whether it is a *in terrorem* or commercially justifiable.
- (3) If a clause is commercially justifiable, it is enforceable, irrespective of whether it is a genuine estimate or *in terrorem*.
- (4) If a clause is a genuine preestimate and commercially justifiable, it is enforceable, irrespective of whether it is *in terrorem*.
- (5) If a clause is *in terrorem*, it is enforceable if it is commercially justifiable but unenforceable if it is a genuine pre-estimate.
- (6) If a clause is a genuine preestimate, it is enforceable if it is commercially justifiable but unenforceable if it is *in terrorem*.

Since nothing in our setup allows to determine that one of the six is to be preferred, each is as likely to be taken up by a judge as any other. At T_4 , therefore, any one prediction is likely to be correct with a probability of 16.66% – a third of the confidence that we may have had at T_0 , where any prediction would have been true in 50% of cases.

Once applicative and hierarchic uncertainty are considered together, the legal change between T_4 and T_5 actually makes the law more certain. Why so? Suppose that under the simple T_5 law, we know whether it applies to some facts with a probability of 20%, and under the complex T_4 one, that certainty is 90%. At T_4 , we only know what the applicable ordering is with a probability of 16%. At T_5 , that probability is 1. Therefore, the compound probability of a prognosis being true under the complex T_4 law is 0.9×0.16 , which gives 0.14. Under the simple T_5 law, it is 0.2×1 , or 0.2. The T_5 law is more certain in the aggregate, even on the wildly unfavourable assumption that it is three times as applicatively uncertain as the T_4 law.

Lastly, consider the role of metalaws. There are several that apply here: earlier High Court judgments bind future High Courts absolutely, the Court of

Appeal cannot depart from its own case law,⁶⁴ House of Lords or Supreme Court judgments bind all courts bar the highest but not Parliament, and the highest court can only depart from its own case law 'where it is right to do so'.⁶⁵ All of these metalaws are much more abstract than the rules contained in the cases themselves: the Supreme Court's case law includes not only *ParkingEye* but also judgments on contract damages, human rights, the rule on perpetuities, and the royal prerogative. The metalaw that tells us what happens when the Supreme Court's case law clashes with that of other courts is much more general than any individual rule contained in that case law.

How do these metalaws keep the system coherent? I said earlier that between T₀ and T₄, there emerged three laws which could be hierarchically arranged in six different ways. Without the set of metalaws I just described, the number of possible orderings would be considerably higher: in every case, the judge would be free to disregard previous judgments and introduce interpretative principles of his own. Moreover, the T₅ judgment would not reduce hierarchic uncertainty, since there would be no metalaw to indicate that if *ParkingEye* is at odds with any previous authorities, it is the former that prevails. Note, however, that the metalaws in place are not perfectly certain: the Supreme Court in *ParkingEye* was authorised to depart from the previous case law because it was 'right to do so'. But no-one could have predicted whether it would be right to depart from the previous case law before the case came before the Supreme Court. In other words, before *ParkingEye* was decided, any prediction of its outcome would have had to account for the possibility that it would be 'right' for that court to depart from precedent.⁶⁶ Thus, in the last analysis, it is metalaws that fix the boundaries of legal

⁶⁴ *Young v Bristol Aeroplane* [1944] KB 718.

⁶⁵ *Practice Statement* (n 31).

⁶⁶ The same is true of the metalaws that govern between T₀ and T₄, albeit in a more subtle way. The reason those are hierarchically uncertain is that there is a metalaw that permits the judiciary to *distinguish* between cases and also another that prohibits them from overruling previous judgments. Since distinguishing is a form of lawmaking – see, for example, Kelsen (n 3) 77 and Raz (n **Error! Bookmark not defined.**) 94–101 – it is inevitable that there will be a tension between the distinction and the law as originally promulgated. The problem is that although there is a metalaw which permits the British judiciary to distinguish cases, there is no indication in that metalaw of when distinguishing is permissible and when it is not.

prediction: what is known is knowable because there are metalaws, and the unknown is unknown because those metalaws are not perfectly certain.

IV. CONCLUSION

My purpose here was to reframe an old problem, rather than to solve it. It is nonetheless possible to synthesise the main points. First, there are two types of legal uncertainty. One comes from the vagueness of laws. The other has to do with the relationship between laws. Although their cause is the same – the scarcity of information – the solutions come in different forms. Applicative uncertainty is reduced by introducing more specific laws, hierarchic uncertainty through more abstract ones. There is a tension between the two, and a balance to be struck – so long as information remains costly, the law must remain partly uncertain. Second, once we account for hierarchic uncertainty, it becomes obvious why in law structure matters: metalaws keep the system coherent. Without them, we would be unable to have anything but the most general laws. The ancients found vague commandments to be perfectly serviceable bases for their legal regimes. Industrialisation, however, intensified the specialisation of labour and with it the specification of law. In our era, the thrust has been to ever-increasing complexity in the economy, society, and the law. Without metalaws to keep hierarchic uncertainty at bay, the welfare gains from economic specialisation – and surely there must have been at least a few – would have remained unrealised.