Analogical Legal Reasoning: Theory and Evidence

Joshua C. Teitelbaum
Georgetown University Law Center

Prelimimary draft: January 2010
Please do not cite or quote without permission

Abstract

The paper offers a model of analogical legal reasoning. Under the model, the outcome of the case at hand is a weighted average of the outcomes of prior cases, where the weights are a function of the fact similarity and precedential authority of the prior cases. The paper then investigates whether the analogical model provides a better fit than a simple rule-based model (represented by a fractional polynomial) to the reported decisions by federal judges in U.S. maritime salvage cases from 1880 to 2007. The principal finding is that the rule-based model fits the data better than the analogical model. The paper also presents a regression tree analysis of the maritime salvage cases as a supplement to the main empirical analysis. Finally, it discusses the implications and limitations of the empirical analysis.

1 Introduction

How do judges reason about the law? There are many theories. The canonical theory is that judges reason by analogy from case to case (Levi 1949; Weinreb 2005). This method of reasoning is known as analogical legal reasoning to jurisprudence scholars and case-based legal reasoning to scholars in the field of artificial intelligence and law.[1]

In its purest form, analogical legal reasoning (ALR) involves reasoning directly from prior cases to the case at hand—the judge evaluates the similarities and differences between prior cases and the case at hand and reaches a decision through application of the principle that like cases should be treated alike (Alexander and Sherwin 2008). Notably, ALR operates without invoking a legal rule that governs the decision in the case at hand (Sunstein 1993, 1996).[2]

A leading alternative theory is that judges reason deductively from legal rules (Alexander and Sherwin 2008; Schauer 2009).[3] In other words, they engage in rule-based legal reasoning (RLR). In its purest form, RLR operates without reference to prior cases—the judge simply applies the governing legal rule to the case at hand. At most, the judge uses prior cases to infer (perhaps abductively or inductively) the governing legal rule; however, she does not reason directly from case to case.

Both ALR and RLR constitute "legalist" theories of judicial behavior (Posner 2008). According to the legalist theory, "judges decide cases through systematic application of the external, objective sources of authority that classically comprise the law" (Cross 2003).[4] Although the legalist theory is the traditional theory of judicial behavior in legal circles, it has many critics. Perhaps the leading criticism of the legalist theory is that it suffers from theoretical indeterminacy (Cross 2003), presumably due to the paucity of formal models in the jurisprudence literature. ALR has been especially targeted by

[1] Some commentators argue that its use of analogy makes legal reasoning a distinctive form of reasoning (e.g., Fried 1981; Weinreb 2005). The mystical notion that legal reasoning is a distinctive form of reasoning was famously articulated by Sir Edward Coke, the Chief Justice of England, who denied the authority and competence of the King of England to render legal judgments on the grounds that legal questions "are not to be decided by natural reason but by the artificial reason and judgment of law" (*Prohibitions Del Roy*, 77 Eng. Rep. 1342 (1607)).

[2] On the different forms of analogical legal reasoning, see generally Macagno and Walton (2009).

[3] See also Westen (1982), Eisenberg (1988), Posner (1990, 1995, 2006, 2008), Schauer (1991), and Alexander (1996, 1998).

[4] Of course, there are many other theories of judicial behavior. Posner (2008) identifies no fewer than nine theories, including most notably the legalist theory, the attitudinal theory, which posits that judges decide cases according to their ideological preferences (e.g., Segal and Spaeth 1993, 2002), and the economic and strategic theories, which posit that judges decide cases strategically, taking into account the responses of other actors, to promote their ideology (e.g., Epstein and Knight 1998; Smith and Tiller 2002), enhance their reputation or career prospects (e.g., Miceli and Coşgel 1994; Levy 2005), or further some other specified objective.

critics, with one commentator complaining that "it is infrequently described with any rigor or care" (Alexander 1996).[5]

This paper has two objectives. The first objective is to offer a formal model of ALR. I model ALR as similarity-weighted averaging of prior outcomes. More specifically, the model posits that the outcome in the case at hand is a weighted average of the outcomes of prior cases, where the weights are a function of the fact similarity and precedential authority of prior cases. The full specification of the model appears in Section 2. The main theoretical result of the paper is an axiomatization of a key feature of the model, which appears in Section 3.

The ALR model is closely related to the "empirical similarity" model of Gilboa et al. (2006), as well as the wider literature on case-based decision theory.[6] Case-based decision theory is an axiomatic model of reasoning by analogy to past cases (Gilboa and Schmeidler 2001).[7] Empirical similarity theory is a closely-related axiomatic model in which assessments are made according to similarity-weighted averages of prior assessments (Gilboa et al. 2006).[8] I compare and contrast the ALR model and empirical similarity theory in the course of specifying the ALR model in Section 2.

Empirical similarity theory is closely related to various methods in computer science, statistics, and related fields, including, most notably: kernel methods (Pagan and Ullah 1999), which are commonly used in nonparametric estimation; nearest neighbor methods (Dasarathy 1991; Devroye et al. 1996), which are commonly used in machine learning and pattern recognition; and conditional autoregressive (CAR) and simultaneous autoregressive (SAR) models (Banerjee et al. 2004), which are commonly used in the analysis of areal and other spatial data.[9] In addition, ALR is studied in the artificial intelligence and law literature, which contains a number of computational models of case-based adjudication (Rissland 1990; Rissland et al. 2003, 2006; Ashley and Brüninghaus 2006), as well as various theoretical models that are directed towards providing algorithmic or logical underpinnings for the computational models (Bench-Capon et al. 2004; Sartor 2005; Walton 2005; Bench-Capon and Prakken 2006; Bench-Capon et al. 2009).

[5] Notable exceptions include Sunstein (1993, 1996), Brewer (1996), and Weinreb (2005).

[6] In case-based decision theory, the term "case" is used generically; it does not refer to a legal case.

[7] See also Gilboa and Schmeidler (1995, 1996, 1997, 2000, 2002, 2003) and Gilboa et al. (2002). Case-based decision theory was inspired by work on case-based reasoning in artificial intelligence (Riesbeck and Schank 1989) and harkens back to the notion that all human "reasonings concerning matter of fact are founded on a species of Analogy" (Hume 1748).

[8] See also Billot et al. (2005), Gayer et al. (2007), Billot et al. (2008), Lieberman (forthcoming), and Gilboa et al. (forthcoming).

[9] I expand upon the connection between empirical similarity theory and kernel regression in Section 4.1. For discussions of the relationship between empirical similarity theory, on the one hand, and nearest neighbor methods and conditional autoregressive models, on the other hand, see Lieberman (forthcoming) and Gilboa et al. (forthcoming).

The second objective of the paper is to empirically evaluate the ALR model by testing whether it has more explanatory power than a simple RLR model. For a simple model of RLR, I turn to fractional polynomial regression (Royston and Altman 1994).[10] A fractional polynomial is an extension of a conventional polynomial that allows for noninteger and negative powers. Fractional polynomial regression is a flexible parametric method for approximating unknown functions using few parameters. Under the view that legal rules are functions (which map facts to outcomes),[11] fractional polynomial regression provides a flexible yet parsimonious method for modeling legal rules.[12] The full specification of the RLR model appears in Section 4.

Using data on U.S. maritime salvage cases, I compare the ALR and RLR models according to their Bayesian information criteria (BIC).[13] Proposed by Schwarz (1978), BIC is a standard criterion for comparing and selecting among nonnested models. The maximum likelihood estimates for both models and the results of the BIC test appear in Section 4. The main conclusion is that the RLR model fits the data better than the ALR model.

As a supplement to the main empirical analysis, Section 4.4 presents a regression tree analysis of the maritime salvage cases. Regression tree analysis is a nonparametric method for analyzing the relationship between categorical or continuous independent variables and a continuous dependent variable (Bierman et al. 1984).[14] Although the regression tree analysis does not shed light directly on the question of which model better fits the data, it serves as a robustness check of the coefficient estimates for both models.

In Section 5, I discuss implications and limitations of the empirical analysis. I also discuss a conceptual issue that underlies the enterprise of the paper, namely the extent to which ALR and RLR are theoretically distinct methods of legal reasoning. Concluding remarks appear in Section 6. The Appendix provides an overview of U.S. maritime salvage law.

[10] See also Royston and Altman (1997) and Royston and Sauerbrei (2008).

[11] See, e.g., Kornhauser (1992a,b), Cameron et al. (2000), Cameron and Kornhauser (2005, 2009), and Kastellec (forthcoming).

[12] The use of mathematical and statistical methods to model legal rules is the enterprise of the fact-pattern analysis literature in political science (Kort 1957, 1963, 1968, 1973; Kort and Mars 1957; Mackaay and Robillard 1974; Segal 1984; Cameron and Kornhauser 2005; Kastellec forthcoming).

[13] In Section 4.3, I explain why U.S. maritime salvage cases provide a fertile testing ground for comparing the ALR and RLR models.

[14] A closely related method—classification tree analysis—is used when the dependent variable is categorical. Kastellec (forthcoming) conducts classification tree analysis of search and seizure cases decided by the U.S. Supreme Court and confession cases decided by the Courts of Appeals.

2 A Model of Analogical Legal Reasoning

Let $\mathcal{K}$ denote the set of judges or *courts* in the legal system. The courts in $\mathcal{K}$ are ordered in accordance with the hierarchy of courts in the legal system. Let $\mathcal{Q}$ denote the set of *questions* of law that may be presented to a court. For each question $q \in \mathcal{Q}$, there exists a set of *conclusions* of law $\mathcal{Y}_q$ that a court may reach with respect to question q and a vector of *issues* of fact $\varphi_q = (\varphi_{q1}, ..., \varphi_{qn})$ that the court must resolve in order to reach a conclusion with respect to question q.[15] For each issue φ_{qi}, there exists a set of *findings* of fact Φ_{qi} that the court may make with respect to issue φ_{qi}. Accordingly, each question $q \in \mathcal{Q}$ induces a *fact space* $\Phi_q = \Phi_{q1} \times \cdots \times \Phi_{qn}$. Each element $\phi = (\phi_1, ..., \phi_n) \in \Phi_q$ is a *fact pattern*. Given question q, the set of conclusions $\mathcal{Y}_q$, the vector of issues φ_q, and the fact space Φ_q are known and unique. A *case* involving question q is a triple $c = (\phi, \kappa, y)$, where $\phi \in \Phi_q$, $\kappa \in \mathcal{K}$, and $y \in \mathcal{Y}_q$. Define $x = (\phi, \kappa)$ as the *inputs* and y as the *outcome* of the case. The set of all possible cases involving question q is $\mathcal{C}_q = (\Phi_q \times \mathcal{K}) \times \mathcal{Y}_q$. I shall assume throughout the paper that the inputs and outcomes of cases are or may be represented as real variables: $\Phi_q = \mathbb{R}^n_+$, $\mathcal{K} = \mathbb{R}_+$, and $\mathcal{Y}_q = \mathbb{R}$.

At time $t \in \mathbb{N}_{++}$, a court is presented with question q and a body of evidence. Based on the evidence, the court makes findings of fact $\phi_t \in \Phi_q$ with respect to issues φ_q. The court has access to a q-relevant *case history* $C_t = (c_1, ..., c_{t-1})$, where each $c_j = (x_j, y_j) \in \mathcal{C}_q$ is a prior case involving question q. How the court reaches its conclusion y_t depends on whether the court engages in ALR or RLR. What fundamentally distinguishes ALR and RLR is that under ALR the outcome of the case at hand is a function of the inputs of the case at hand as well as the history of prior cases, $y_t = Y(x_t, C_t)$, whereas under RLR the outcome of the case at hand is a function of the inputs only, $y_t = Y(x_t)$.[16]

I model ALR as similarity-weighted averaging of prior outcomes. Formally,

$$y_t = Y\left(x_t, C_t\right) = \sum_{j<t} \left(\frac{s\left(x_t, x_j\right)}{\sum_{j<t} s\left(x_t, x_j\right)} \right) y_j, \tag{1}$$

where

$$s\left(x_t, x_j\right) = \exp\left(-\mu\left(x_t, x_j\right)\right), \tag{2}$$

$$\mu\left(x_t, x_j\right) = v\left(x_t, x_j\right) d\left(\phi_t, \phi_j\right), \tag{3}$$

[15] Note that n (the dimension of φ_q) is a function of q.

[16] Stated another way, under RLR the outcome depends on a bounded number of parameters, whereas under ALR the number of parameters increases with the prior case history (cf. Gayer et al. 2007).

$$v\left(x_t, x_j\right) = \begin{cases} \cos\left(\arctan\left(\frac{\beta}{d\left(\phi_t,\phi_j\right)}\right)\right) & \text{if } \phi_t \neq \phi_j \ \& \ \kappa_t < \kappa_j \\ 1 & \text{otherwise} \end{cases}, \ \beta \geq 0, \tag{4}$$

and

$$d\left(\phi_t, \phi_j\right) = \sqrt{\sum_{i=1}^{n} \omega_i \left(\phi_{ti} - \phi_{ji}\right)^2}. \tag{5}$$

The model posits that the outcome y_t in the case at hand is a weighted average of the outcomes $y_1, ..., y_{t-1}$ of prior cases. The weight placed on the outcome y_j of a prior case depends on the degree to which the inputs x_j of the prior case are similar to the inputs x_t of the case at hand. The degree of input similarity is given by s. The greater is the input similarity of a prior case, the greater is the weight given to the outcome of the prior case in the determination of the outcome of the case at hand. Hence, I interpret s as measuring the *precedential influence* of a prior case on the case at hand. I assume that input similarity—and, therefore, precedential influence—is an exponentially decaying function of the distance μ from the inputs of the prior case to the inputs of the case at hand.[17] In turn, I assume that input distance is a proportional function (with nonconstant proportionality factor v) of the weighted Euclidean distance d between the facts ϕ_j of the prior case and the facts ϕ_t of the case at hand.[18] The proportionality factor v is less than one if the prior case was decided by a superior court ($\kappa_t < \kappa_j$) and equals one otherwise. All else equal, therefore, prior cases decided by a superior court receive greater weight—and, therefore, have more influence—than prior cases decided by a parallel or inferior court in the determination of the outcome of the case at hand. The size of this influence advantage, however, is smaller the greater is d (the distance between the prior case and the case at hand in fact space); how much smaller is determined by the shape parameter β.

The notion of similarity-weighted averaging as a model of analogical reasoning was introduced by Gilboa et al. (2006), who provided an axiomatization of similarity-weighted averaging with a generic similarity function. Although they did not specify a particular similarity function or even a particular functional form, Gilboa et al. (2006) were interested in similarity functions that depend on a weighted Euclidean distance. The notion of a similarity function that decays exponentially as a function of distance was

[17] The assumption that influence decays exponentially with distance seems natural and appears in other contexts (Bolhuis et al. 1986; Nosofsky 1986; Shephard 1987; Glaeser et al. 2003; Billot et al. 2008). Several studies provide evidence of exponential decay with time of the precedential influence of legal cases in U.S. federal courts (Post and Eisen 2000; Fowler and Jeon 2008; Black and Spriggs II 2009).

[18] Note that the weights $\omega_1, ..., \omega_n$ in the weighted Euclidean distance d reflect the relative importance of the n issues of fact that the court must resolve in order to reach a conclusion with respect to the legal question at issue.

introduced by Billot et al. (2008), who provided an axiomatization of an exponential similarity function based on a generic metric (i.e., a symmetric distance function). Billot et al. (2008) also axiomatized an exponential similarity function based on a weighted Euclidean distance, which is a special case of a metric.

In my model of ALR, the similarity function s is based on the input distance μ, which is a quasimetric (i.e., an asymmetric distance function) (see Section 3).[19] This generalization of the standard similarity function is instrumental to my purposes,[20] because the precedential influence of a prior case in a legal system with hierarchical courts depends not only on the *fact similarity* of the prior case, which depends on the distance from the prior case in fact space, but also on the *precedential authority* of the prior case, which depends on the position in the judicial hierarchy of the court that decided the prior case. Although fact similarity is symmetric,[21] precedential authority is not symmetric. All else equal, the precedential authority of a case decided by a superior court is greater than the precedential authority of a case decided by a parallel or inferior court. Therefore, if the prior case was decided by a superior court, its influence on the case at hand ought to be greater than the influence of the case at hand on the prior case (under the counterfactual that the case at hand was decided before the prior case). Specifying a similarity function that is based on quasimetric allows the model to capture this distinctive feature of precedential influence in law.

Figure 1 displays the relationship in the model between precedential influence (s), fact similarity (d), and precedential authority (v).[22] As the figure illustrates, the precedential influence of a prior case is at its maximum ($s = 1$) when the facts of the prior case are identical to the facts of the case at hand ($\phi_t \neq \phi_j \Leftrightarrow d = 0$). As fact similarity decreases (i.e., as d increases), precedential influence decays exponentially at rate v—i.e., $s = \exp(-vd)$. The rate of decay differs depending on the precedential authority of the prior case. If the prior case was decided by a parallel or inferior court ($\kappa_t \geq \kappa_j$), the rate of decay equals one ($v = 1$). If, however, the prior case was decided by a superior court, the rate of decay is $v = \cos(\arctan(\beta/d)) < 1$.[23] All else equal, therefore, the precedential

[19] Strictly speaking, μ is a quasimetric provided the parameter β is sufficiently small (see Section 3).

[20] Like Gilboa et al. (2006) and Billot et al. (2008), Gayer et al. (2007), Lieberman (forthcoming), and Gilboa et al. (forthcoming) contemplate similarity functions based on a metric, usually a weighted Euclidean distance.

[21] That is, the distance in fact space from the prior case to the case at hand is equal to the distance from the case at hand to the prior case.

[22] Note that fact similarity and precedential authority are negatively related to d and v, respectively.

[23] The specification of v is motivated as follows. For any two points $x_1, x_2 \in \mathbb{R}^n_+$, $x_1 \neq x_2$, a standard generic measure of the asymmetric distance from x_1 to x_2 is $f(\theta_{12})d(x_1,x_2)$, where: $d(x_1,x_2)$ is the Euclidean distance between x_1 and x_2; θ_{12} is the polar direction from x_1 to x_2; and $f(\theta)$ is a monotonically increasing function on $(0, \pi/2)$, typically chosen or normalized such that $f(0) = 0$ and $f(\pi/2) = 1$ (Drezner and Wesolowsky 1989). Turning to our setting, take any $x_j = (\phi_j, \kappa_j)$ and $x_t = (\phi_t, \kappa_t)$ in the

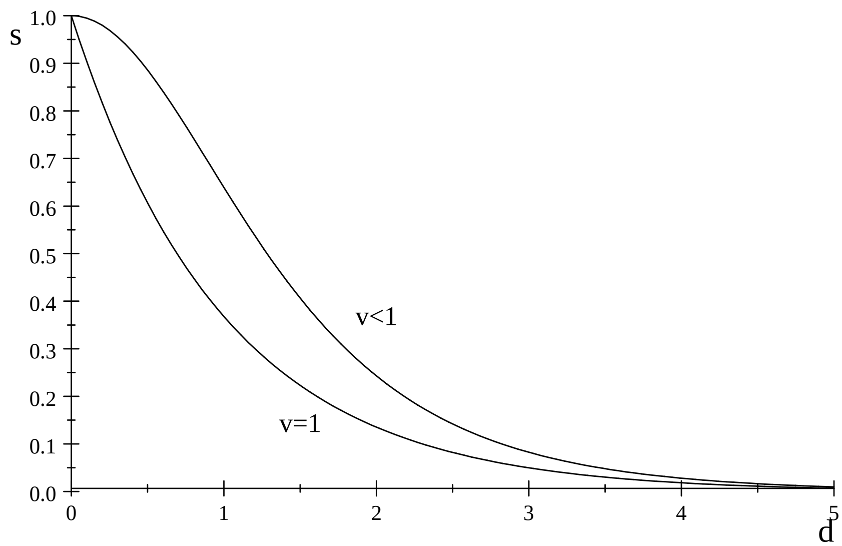

Figure 1: $s = \exp(-vd)$

influence of a prior case that was decided by a superior court is greater than precedential influence of a prior case that was decided by a parallel or inferior court. Moreover, the size of the influence advantage due to enhanced precedential authority (which, as noted above, is negatively related to v) increases with the degree of fact similarity (i.e., it increases as d decreases) at a rate determined by (and positively related to) the shape parameter β (see Figure 2).

3 Theoretical Results

In this section I prove two theoretical results pertaining to the model.

3.1 Similarity as a Quasimetric

The first result is that the function $\mu : (\Phi_q \times \mathcal{K}) \times (\Phi_q \times \mathcal{K}) \rightarrow \mathbb{R}_+$, on which the similarity function $s : (\Phi_q \times \mathcal{K}) \times (\Phi_q \times \mathcal{K}) \rightarrow \mathbb{R}_{++}$ is based, is a *quasimetric* provided that the shape parameter β is sufficiently small. Recall the definition of a quasimetric:

input space $\Phi_q \times \mathcal{K} = \mathbb{R}_+^{n+1}$, where $\phi_t \neq \phi_j$ and $\kappa_t < \kappa_j$ (implying $x_j \neq x_t$). If we normalize the distance in $\mathcal{K} = \mathbb{R}_+$ (judicial hierarchy space) between superior and inferior courts to one (so $\kappa_j - \kappa_t = 1$), then $\theta_{jt} = \arctan(\beta/d(\phi_t, \phi_j))$ is the β-scaled polar direction from x_j to x_t, and $v(\theta_{jt}) = \cos(\theta_{jt})$ is just the standard quasimetric described above, where $f(\theta) = \cos(\theta)$ and d is the weighted Euclidean distance in $\Phi_q = \mathbb{R}_+^n$ (fact space).

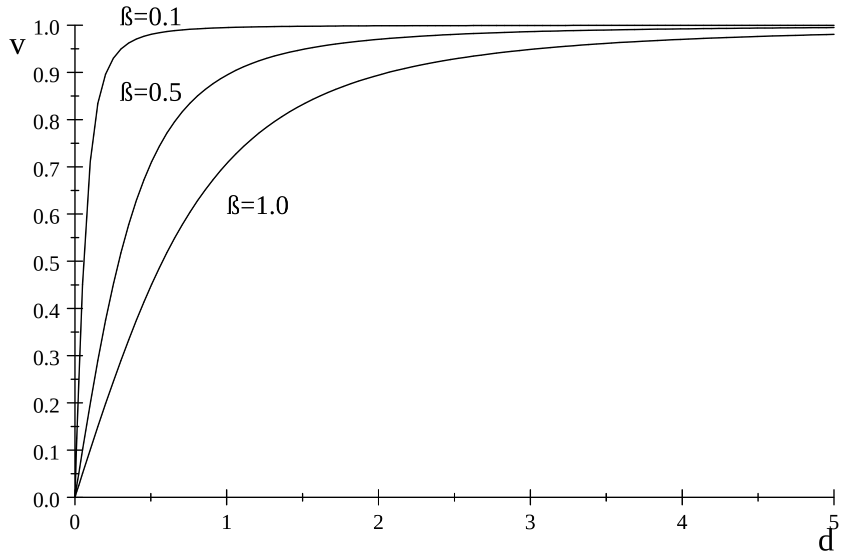

Figure 2: $v = \cos\left(\arctan\left(\frac{\beta}{d}\right)\right)$

Definition 1 (Quasimetric) *A function* $\xi : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}^n$ *is a* quasimetric *on* $\mathbb{R}^n$ *if for all* $x, y \in \mathbb{R}^n$*:*

(i) $\xi(x, y) = 0$ *if* $x = y$ *(identity of indiscernibles);*

(ii) $\xi(x, y) > 0$ *if* $x \neq y$ *(positivity);*

(iii) $\xi(x, z) \leq \xi(x, y) + \xi(y, z)$ *for any* $z \in \mathbb{R}^n$ *(triangle inequality).*

Note that a metric is a quasimetric that also satisfies symmetry: $\xi(x, y) = \xi(y, x)$. A quasimetric is not necessarily symmetric, i.e., in general $\xi(x, y) \neq \xi(y, x)$.

The following is a formal statement of the first result.

Theorem 1 *For all* $\phi, \psi \in \mathbb{R}^n$ *and* $\kappa, \iota \in \mathbb{R}$*, with* $w = (\phi, \kappa)$ *and* $z = (\psi, \iota)$*, let*

$$\mu\,(w, z) = v\,(w, z)\, d\,(\phi, \psi)\,,$$

where

$$v\,(w, z) = \left\{ \begin{array}{cl} \cos\left(\arctan\left(\frac{\beta}{d(\phi,\psi)}\right)\right) & \textit{if } \phi \neq \psi \textit{ \& } \kappa < \iota \\ 1 & \textit{otherwise} \end{array} \right. , \ \beta \geq 0,$$

and d *is a metric on* $\mathbb{R}^n$*. There exists* $\rho > 0$ *such that for all* $\beta < \rho$*,* μ *is a quasimetric on* $\mathbb{R}^{n+1}$*.*

Proof. First, take any $\beta \geq 0$. If $w = z$, then $\phi = \psi$ and $\kappa = \iota$. This implies $d(\phi, \psi) = 0$ and $v(w, z) = 1$, which in turn implies $\mu(w, z) = 0$. Thus, μ satisfies the identity of indiscernibles for every $\beta \geq 0$.

Next, suppose $\beta = 0$. Then $v(w, z) = 1$ (because $\arctan(0) = 0$ and $\cos(0) = 1$). This implies $\mu(w, z) = d(\phi, \psi)$. Because d is a metric, it follows that μ satisfies positivity and the triangle inequality for $\beta = 0$.

Finally, suppose $\beta > 0$. If $\phi = \psi$ or $\kappa \geq \iota$, then $v(w, z) = 1$ (by definition). This implies $\mu(w, z) = d(\phi, \psi)$. It follows that μ satisfies positivity and the triangle inequality for $\beta > 0$ when $\phi = \psi$ or $\kappa \geq \iota$.

If $\phi \neq \psi$ and $\kappa < \iota$, then $d(\phi, \psi) > 0$ (by positivity of d) and $v(w, z) \in (0, 1)$ (because $\arctan(\theta) \in (0, \frac{\pi}{2})$ for $0 < \theta < \infty$ and $\cos(\theta) \in (0, 1)$ for $0 < \theta < \frac{\pi}{2}$). It follows that $\mu(w, z) > 0$. (It also follows that $\mu(w, z) < d(\phi, \psi) = \mu(z, w)$.) Thus, μ satisfies positivity (but not symmetry) for $\beta > 0$ when $\phi \neq \psi$ and $\kappa < \iota$.

To complete the proof, we need to show that when $\phi \neq \psi$ and $\kappa < \iota$, there exists $\rho > 0$ such that μ satisfies the triangle inequality for every $\beta < \rho$. Take any $\vartheta \in \mathbb{R}^n$ and $\eta \in \mathbb{R}$, with $x = (\vartheta, \eta)$. Note that μ satisfies the triangle inequality iff $v(w, z)\, d(\phi, \psi) \leq v(w, x)\, d(\phi, \vartheta) + v(x, z)\, d(\vartheta, \psi)$.

Define $F(\beta) = v(w, x)\, d(\phi, \vartheta) + v(x, z)\, d(\vartheta, \psi) - v(w, z)\, d(\phi, \psi)$. Note that $F(0) = d(\phi, \vartheta) + d(\vartheta, \psi) - d(\phi, \psi)$ (because $v = 1$ for $\beta = 0$). Note further that $F(0) > 0$ (because d is a metric and $\phi \neq \psi$) and that F is continuous in β (because v is continuous in β). Because F is continuous in β, for every $\epsilon > 0$ there exists $\delta_\epsilon > 0$ such that $\beta < \delta_\epsilon$ implies $|F(\beta) - F(0)| < \epsilon$. Let $\epsilon = F(0) > 0$. It follows that there exists $\rho = \delta_{F(0)} > 0$ such that $\beta < \rho$ implies $F(\beta) > F(0) - \epsilon = 0$. It follows, in turn, that there exists $\rho > 0$ such that for all $\beta < \rho$, $v(w, x)\, d(\phi, \vartheta) + v(x, z)\, d(\vartheta, \psi) > v(w, z)\, d(\phi, \psi)$. ■

3.2 Axiomatization of an Exponential Similarity Function Based on a Quasimetric

The second result is an axiomatization of an exponential similarity function based on a quasimetric. The axiomatization closely follows Billot et al. (2008), which, as noted above, provided an axiomatization of an exponential similarity function based on a standard metric.

Let $\mathbb{C} = \cup_{t \geq 1} (\mathbb{R}^{n+2})^{t-1}$. Suppose there are given functions $Y : \mathbb{R}^{n+1} \times \mathbb{C} \to \mathbb{R}$ and $s : \mathbb{R}^{n+1} \times \mathbb{R}^{n+1} \to \mathbb{R}_{++}$ such that for all $x_t = (\phi_t, \kappa_t) \in \mathbb{R}^{n+1}$ and $C_t = (x_j, y_j)_{j<t} \in \mathbb{C}$,

$$y_t = Y(x_t, C_t) = \sum_{j<t} \left(\frac{s(x_t, x_j)}{\sum_{j<t} s(x_t, x_j)} \right) y_j \tag{6}$$

and s is normalized such that $s(x,x) = 1$ for all $x \in \mathbb{R}^{n+1}$.

I impose three axioms on Y.

Axiom 1 (Ray Monotonicity) *For all $w, x, z \in \mathbb{R}^{n+1}$, $x \neq 0$, $Y\left(w, \left(\left(w + \lambda x, 1\right), \left(w + z, -1\right)\right)\right)$ is strictly decreasing in $\lambda \geq 0$.*

Axiom 1 considers a world with two prior cases, $c_1 = (\phi_1, y_1) = (w + \lambda x, 1)$ and $c_2 = (\phi_2, y_2) = (w + z, -1)$. In such a world, equation (6) would generate an outcome between y_1 and y_2, i.e., $Y \in [-1, 1]$. Axiom 1 states that as λ increases—whereby $\phi_1 = w + \lambda x$ moves further away from w (along a ray through w), thereby becoming less similar to w—Y decreases, i.e., moves away from $y_1 = 1$ and toward $y_2 = -1$.

Axiom 2 (Ray Shift Invariance) *For case histories of the form $C = \left(w + \alpha_j x, y_j\right)_{j<t} \in \mathbb{C}$, where $w, x \in \mathbb{R}^{n+1}$, $y_j \in \mathbb{R}$ $(j < t)$, and $\alpha_j \geq 0$ $(j < t)$, $Y\left(w, \left(w + \left(\alpha_j + \delta\right) x, y_j\right)_{j<t}\right) = Y\left(w, \left(w + \alpha_j x, y_j\right)_{j<t}\right)$ for all $\delta > 0$.*

Axiom 2 considers a world in which the inputs of all prior cases lie on a ray through the inputs of the case at hand, w. Axiom 2 requires that a shift δ along this ray leaves the outcome Y unchanged.

Axiom 3 (Self-relevance) *For all $w, x, z \in \mathbb{R}^{n+1}$, $Y\left(z, \left(\left(w, 1\right), \left(x, 0\right)\right)\right) \leq Y\left(w, \left(\left(w, 1\right), \left(x, 0\right)\right)\right)$.*

Axiom 3 considers a world with two prior cases, $(w, 1)$ and $(x, 0)$. In such a world, equation (6) would generate an outcome $Y \in [0, 1]$. Axiom 3 requires that the outcome Y in a new case z be less than the outcome Y in a new case w. The idea is that any new case z must be judged less similar to w than w is to itself.

Theorem 2 states the second result.

Theorem 2 *The following are equivalent:*

(i) Y satisfies Axioms 1-3;

(ii) There exists a quasimetric μ on $\mathbb{R}^{n+1}$ such that $s(w,z) = \exp\left(-\mu\left(w,z\right)\right)$ for all $w, z \in \mathbb{R}^{n+1}$.

Proof. First, take any $w, x, z \in \mathbb{R}^{n+1}$ with $x \neq 0$. Ray Monotonicity (Axiom 3.1) holds iff

$$\frac{s\left(w, w + \lambda x\right) - s\left(w, w + z\right)}{s\left(w, w + \lambda x\right) + s\left(w, w + z\right)}$$

is strictly decreasing in $\lambda \geq 0$. Let $f(\lambda) = s(w, w + \lambda x)$. It follows that Ray Monotonicity holds iff

$$\frac{2s(w, w+z) \cdot f'(\lambda)}{[f(\lambda) + s(w, w+z)]^2} < 0,$$

which holds iff $f'(\lambda) < 0$.

Next, take $C = (w + \alpha_j x, y_j)_{j<t} \in \mathbb{C}$, where $w, x \in \mathbb{R}^{n+1}$, $y_j \in \mathbb{R}$ $(j < t)$ and $\alpha_j \geq 0$ $(j < t)$. Ray Shift Invariance (Axiom 3.2) holds iff for all $\beta > 0$,

$$\sum_{j<t} \left(\frac{s(w, w + (\alpha_j + \beta)x)}{\sum_{j<t} s(w, w + (\alpha_j + \beta)x)} - \frac{s(w, w + \alpha_j x)}{\sum_{j<t} s(w, w + \alpha_j x)} \right) y_j = 0.$$

This holds iff for all α_j and β,

$$\frac{s(w, w + (\alpha_j + \beta)x)}{s(w, w + \alpha_j x)} = \frac{\sum_{j<t} s(w, w + (\alpha_j + \beta)x)}{\sum_{j<t} s(w, w + \alpha_j x)} = g_\beta,$$

where $0 < g_\beta \leq 1$ (by Ray Monotonicity). Let $f(\lambda) = s(w, w + \lambda x)$, $\lambda \geq 0$. It follows that Ray Shift Invariance is equivalent to $f(\alpha_j + \beta) = f(\alpha_j)\, g_\beta$ for all α_j and β. Observe that $f(0) = 1$; therefore, $f(\beta) = g_\beta$ for all β. Accordingly, Ray Shift Invariance holds iff $f(\alpha_j + \beta) = f(\alpha_j) f(\beta)$ for all α_j and β. Because f is continuous and positive (by continuity and positivity of s), it follows that Ray Shift Invariance is equivalent to $f(\lambda) = \exp(\lambda h_x)$ for $\lambda \geq 0$, where $h_x \leq 0$ (by Ray Monotonicity) with equality iff $x = 0$ (by $s(w, w) = 1$). Defining $z = w + \lambda x$ and $\mu(w, z) = -\lambda h_x$, we conclude that Axioms 3.1-3.2 are equivalent to $s(w, z) = \exp(-\mu(w, z))$ for all $w, z \in \mathbb{R}^{n+1}$, where the function μ satisfies all the properties of a quasimetric, apart from the triangle inequality.

The last step is to show that Self-relevance (Axiom 3.3) holds iff μ satisfies the triangle inequality, $\mu(z, x) \leq \mu(z, w) + \mu(w, x)$ for all $w, x, z \in \mathbb{R}^{n+1}$. Take any $w, x, z \in \mathbb{R}^{n+1}$. Self-relevance holds iff $Y(z, ((w, 1), (x, 0))) \leq Y(w, ((w, 1), (x, 0)))$, which holds iff

$$\frac{s(z, w)}{s(z, w) + s(z, x)} \leq \frac{1}{1 + s(w, x)},$$

which, in turn, holds iff $s(z, x) \geq s(z, w)\, s(w, x)$. Because $s(w, z) = \exp(-\mu(w, z))$ for all $w, z \in \mathbb{R}^{n+1}$, this holds iff

$$\exp(-\mu(z, x)) \geq \exp(-\mu(z, w) - \mu(w, x)),$$

which, in turn, holds iff $\mu(z, x) \leq \mu(z, w) + \mu(w, x)$. ∎

4 Empirical Analysis

Given the result of Theorem 2, Axioms 1 through 3 may be interpreted as observable implications of similarity-weighted averaging with an exponential similarity function based on a quasimetric. However, the special case histories contemplated by the axioms are not ones that we would expect to observe in the real world. Therefore, evaluating the ALR model by testing the validity of the axioms is not a promising strategy.

I adopt an empirical approach to evaluating the ALR model. Namely, I test whether the ALR model has more explanatory power than a simple RLR model. First, I embed the ALR model specified in Section 2 in a statistical model. Second, I turn to fractional polynomial regression for a simple model of RLR. Next, I describe the data on U.S. maritime salvage cases, explain why the data provide a fertile testing ground for comparing the ALR and RLR models, and describe the criterion according to which I compare the models (namely, the BIC). I then present the maximum likelihood estimates for both models and the results of the BIC test. Lastly, I present a regression tree analysis of the maritime salvage cases as a supplement to the main empirical analysis.

4.1 Empirical Specification of the ALR Model

The first step of the empirical analysis is to embed the ALR model specified in Section 2 in a statistical model. Following Gilboa et al. (2006) and its progeny,[24] I assume that y_1 is an arbitrary random variable and that for $t = 2, ..., T$,

$$y_t = Y\left(x_t, C_t; \theta_{ALR}\right) = \sum_{j<t} \left(\frac{s\left(x_t, x_j\right)}{\sum_{j<t} s\left(x_t, x_j\right)} \right) y_j + \varepsilon_t, \tag{7}$$

where $\varepsilon_t \overset{iid}{\sim} N\left(0, \sigma^2\right)$, s is defined by (2)-(5), and $\theta_{ALR} = \left(\beta, \omega_1, ..., \omega_n, \sigma^2\right)$. I estimate the $n+2$ parameter vector θ_{ALR} via maximum likelihood. The log-likelihood function is

$$l\left(\theta_{ALR}\right) = -\frac{t}{2}\log\left(2\pi\right) - \frac{t}{2}\log\left(\sigma^2\right) - \frac{y'S'Sy}{2\sigma^2},$$

[24] See also Gayer et al. (2007), Lieberman (forthcoming), and Gilboa et al. (forthcoming).

where $y = (y_1, ..., y_T)'$ and

$$\underset{(t\times t)}{S} = \begin{pmatrix} 1 & 0 & 0 & 0 & 0 & \cdots & 0 \\ -1 & 1 & 0 & 0 & 0 & \cdots & 0 \\ -\frac{s(x_3,x_1)}{\sum_{j<3} s(x_3,x_j)} & -\frac{s(x_3,x_2)}{\sum_{j<3} s(x_3,x_j)} & 1 & 0 & 0 & \cdots & 0 \\ -\frac{s(x_4,x_1)}{\sum_{j<4} s(x_4,x_j)} & -\frac{s(x_4,x_2)}{\sum_{j<4} s(x_4,x_j)} & -\frac{s(x_4,x_3)}{\sum_{j<4} s(x_4,x_j)} & 1 & 0 & \cdots & 0 \\ \vdots & & & & \ddots & & \\ -\frac{s(x_t,x_1)}{\sum_{j<t} s(x_t,x_j)} & \cdots & & & \cdots & -\frac{s(x_t,x_{t-1})}{\sum_{j<t} s(x_t,x_j)} & 1 \end{pmatrix}.$$

For the derivation of the log-likelihood function, as well as an explication of the asymptotic theory of model (7), which establishes a theoretical basis for simple hypothesis tests involving the model parameters, see Lieberman (forthcoming).

Before turning to the RLR model, let me say a few words about the relationship between model (7) and kernel regression.[25] Kernel regression assumes a data generating process of the form

$$y_i = g(x_i) + e_i, \;\; i = 1, ..., N, \tag{8}$$

where $e_i \overset{iid}{\sim} \left(0, \sigma^2\right)$ and g is an unknown function. A standard estimator for g is the Nadaraya-Watson estimator

$$\hat{g}(x) = \sum_{i \le n} \left(\frac{K\left(\frac{x_i - x}{h}\right)}{\sum_{i \le n} K\left(\frac{x_i - x}{h}\right)} \right) y_i, \tag{9}$$

where K is a kernel function (i.e., a non-negative function satisfying, among other regularity conditions, $\int K(z)dz = 1$) and h is a bandwidth parameter. Note that the right-hand side of equation (9) has the same form as the first term on the right-hand side of equation (7). Indeed, a direct mapping exists between them (Gilboa et al. forthcoming). Moreover, they operate in the same way. Each generates a new/predicted value of y by taking a weighted average of the observed values of y where the weights are a function the distance between the new/hypothesized x and the observed values of x. Notwithstanding these connections, however, there is an important distinction between model (7) and kernel regression. Kernel regression is a statistical technique that uses weighted averaging to estimate model (8), which assumes that the data are generated by a function (i.e., a rule), whereas model (7) assumes that that the data are generated by weighted averaging.[26]

[25] On kernel regression, see Pagan and Ullah (1999).

[26] For more on the relationship between model (7) and kernel regression, see Gilboa et al. (2006, forthcoming) and Lieberman (forthcoming).

4.2 Modeling RLR Using Fractional Polynomial Regression

The second step is to specify a simple RLR model. The essence of RLR is that it involves application of a governing legal rule to the case at hand. The source of the governing legal rule is not important; for example, the rule may be stated in or inferred from a statute or regulation or it may be stated in or inferred from prior cases. What is important is that the court invokes the rule in determining the outcome in the case at hand.

A legal rule may be viewed as a function which map facts to outcomes.[27] This view suggests a simple model of RLR—the outcome in case at hand is a function of the facts of the case at hand, $y_t = Y(\phi_t)$. A pragmatic, parametric approach to estimating this unknown function is fractional polynomial regression (Royston and Altman 1994).[28] A fractional polynomial is an extension of a conventional polynomial that allows for noninteger and negative powers. In reliance on Taylor's theorem, conventional polynomials are often used to approximate unknown functions. However, polynomial regression generally involves a tradeoff between flexibility (i.e., fit) and parsimony. Royston and Altman (1994) introduced fractional polynomial regression as a flexible parametric method for approximating unknown functions using few parameters.

The standard multivariable fractional polynomial (MFP) regression model may be expressed as

$$y_t = Y(\phi_t; \theta_{RLR}) = b_0 + \sum_{i=1}^{h} b_i \phi_{it} + \sum_{i=h+1}^{n} \sum_{j=1}^{m} b_{ij} \phi_{it}^{(p_j)} + \varepsilon_t,\ t = 1, ..., T, \tag{10}$$

where $\varepsilon_t \overset{iid}{\sim} N\left(0, \sigma^2\right)$ and $\theta_{RLR} = (b_0, b_1, ..., b_h, b_{h+1,1}, ..., b_{h+1,m}, ..., b_{n1}, ..., b_{nm}, \sigma^2)$. The first h covariates, $\phi_1, ..., \phi_h$, are binary, categorical, or ordinal, and the remaining covariates, $\phi_{h+1}, ..., \phi_n$, are continuous. The powers $p_1, ..., p_m$ are chosen from a predefined set $\mathcal{P}$ according to the MFP algorithm developed by Sauerbrei and Royston (1999).[29] The round bracket notation signifies the Box-Tidwell transformation,

$$\phi_{it}^{(p_j)} = \begin{cases} \phi_{it}^{(p_j)} & \text{for } p_j \neq 0 \\ \ln \phi_{it} & \text{for } p_j = 0 \end{cases} .$$

The degree m is predefined by the researcher. The researcher also predefines two significance levels: α_1, which determines the critical value for variable selection; and α_2, which

[27] See, e.g., Kornhauser (1992a,b), Cameron et al. (2000), Cameron and Kornhauser (2005, 2009), and Kastellec (forthcoming).

[28] See also Royston and Altman (1997) and Royston and Sauerbrei (2008).

[29] See also Royston and Sauerbrei (2008).

determines the critical value for model selection. The parameter vector θ_{RLR}, which has a maximum of $(n-h+1)m+h+1$ parameters, is estimated via maximum likelihood.

4.3 Data, Empirical Strategy, and Results

My empirical strategy is to compare the ability of the ALR and RLR models (models (7) and (10), respectively) to explain the time series of U.S. maritime salvage cases. Under federal maritime law, a salvor of imperiled maritime property on navigable waters is entitled to a monetary award from the owner.[30] There are two forms of maritime salvage: "contract" salvage and "pure" salvage. Contract salvage is rendered pursuant to a prior agreement. Pure salvage is rendered voluntarily in the absence of a contract. The data include only pure salvage cases.

In the United States, the federal courts have exclusive admiralty jurisdiction in cases involving claims for salvage awards. There are three elements of a valid pure salvage claim: (i) a marine peril; (ii) service rendered voluntarily (and not required by a pre-existing duty or contract); and (iii) success in whole or in part. The peril need not be immediately impending—reasonable apprehension of danger is sufficient. In addition, although a party may render salvage services without the request of the owner, it may not force its services upon an owner who refuses assistance. Finally, under the "no cure-no pay" rule, there can be no salvage claim if the property is lost, notwithstanding the efforts of the putative salvors.

In the case of a valid pure salvage claim, the court determines the proper award according to six factors enumerated by the Supreme Court in *The Blackwall*, 77 U.S. (10 Wall.) 1 (1869):

(1) the labor expended by the salvors in rendering the salvage service;

(2) the promptitude, skill, and energy displayed in rendering the service and saving the property;

(3) the value of the property employed by the salvors in rendering the service, and the danger to which such property was exposed;

(4) the risk incurred by the salvors in securing the property from the impending peril;

(5) the value of the property saved; and

(6) the degree of danger from which the property was rescued.

[30] The following is a bare bones description of U.S. maritime salvage law. A more detailed overview appears in the Appendix.

There is no precise formula for computing a salvage award on the basis of the *Blackwall* factors. The court has considerable discretion in weighing the factors and making its determination on a case-by-case basis. The award, however, is limited by the value of the property saved.

The data comprise 130 pure salvage cases from 1880 to 2007.[31] For each case, the data record the salvage award (in thousands of 1980 dollars),[32] the court's findings of fact with respect to each of the six *Blackwall* factors (high = 1 or low = 0),[33] and the position of the court in the judicial hierarchy (circuit court = 1 or district court = 0).[34] Tables 1 and 2 provide descriptive statistics. Table 1 displays summary statistics for each variable. For instance, it shows that the salvage awards range from $320 to $1,865,000 with a mean award of $122,000; the values of the salvaged property range from $1,2000 to $23,400,000 with a mean value of $2,340,000; the danger to the salvaged property was high in 59 percent of the cases; and 32 percent of the cases were finally adjudicated by a circuit court. Table 2 displays the mean salvage award expressed as a percentage of the value of the salvaged property, as well as conditional means given different findings of fact. For instance, it shows that the (unconditional) mean award percentage is 13.2 percent; the mean award percentage for cases in which the danger to the salvaged property was low is 8.9 percent; and the mean award percentage for cases in which the labor expended by the salvors was high is 21.2 percent.

There are several reasons why maritime salvage cases provide a fertile testing ground for comparing the ALR and RLR models. First, the outcome (the salvage award) is a continuous variable (a dollar amount) and the inputs (the *Blackwall* factors) are well defined and stable over time.[35] Second, awards in maritime salvage cases arguably are

[31] The cases were identified by two search methods. The first was "KeyCiting" and "Shepardizing" *The Blackwall* in Westlaw and LexisNexis, respectively. The second was searching three databases: Westlaw's Federal Maritime Law - Cases (FMRT-CS); LexisNexis' Admiralty Cases - Federal and State (MEGA); and American Maritime Cases (AMC), which is available on Westaw and LexisNexis. The searches were designed to locate all reported federal cases decided after December 31, 1869 and on or before December 31, 2007 that apply the *Blackwall* factors to determine a pure salvage award, whether or not the cases cited *The Blackwall*.

[32] Adjustments for inflation were made using Tom's Inflation Calculator, available at http://www.halfhill.com/inflation.html.

[33] I read each case and hand coded the data. For the vast majority of cases, it was straightforward to determine the salvage award and the court's findings of fact with respect to all six *Blackwall* factors. For a handful of cases, however, information necessary to determine one or more variables was missing; these cases are excluded from the data. Absent a good reason why a court's method of reasoning would be correlated with a case having missing information, there is no reason to believe that excluding these cases biases the results of the empirical analysis.

[34] For each case, the court is the court of final adjudication, and the data record the salvage award and findings of fact as determined by the court of final adjudication.

[35] In the words of the U.S. Court of Appeals for the Ninth Circuit, the *Blackwall* factors "have weathered the storms of the past century" (*Saint Paul Marine Transp. Corp. v. Cerro Sales Corp.*, 505 F.2d 1115

Table 1: Summary Statistics

	Variable	Mean	Std Dev	Min	Max
y	Salvage award	122	271	0.32	1,865
x_1	Labor expended by salvors	0.35	0.48	0	1
x_2	Skill displayed by salvors	0.58	0.49	0	1
x_3	Danger to salvors' property	0.26	0.44	0	1
x_4	Risk to salvors	0.25	0.43	0	1
x_5	Value of salvaged property	2,340	4,529	1.20	23,400
x_6	Danger to salvaged property	0.59	0.49	0	1
court	Circuit court indicator	0.32	0.57	0	1

Notes: 130 cases from 1880 to 2007. y and x^5 in thousands of 1980 dollars.

Table 2: Conditional Mean Award Percentages

Variable	Obs	Mean
$y/x5$	130	0.132
$y/x5$ if $x1 = 0$	85	0.089
$y/x5$ if $x2 = 0$	54	0.092
$y/x5$ if $x3 = 0$	96	0.104
$y/x5$ if $x4 = 0$	98	0.104
$y/x5$ if $x6 = 0$	53	0.087
$y/x5$ if $x1 = 1$	45	0.212
$y/x5$ if $x2 = 1$	76	0.160
$y/x5$ if $x3 = 1$	34	0.211
$y/x5$ if $x4 = 1$	32	0.216
$y/x5$ if $x6 = 1$	77	0.162

apolitical legal questions. Moreover, it is hard to imagine that a maritime salvage case is an opportunity for a judge to advance strategic goals such as career advancement. Thus, if there is any setting in which we should expect "legalist" models of judicial behavior to be operative (and other models such as attitudinal or strategic models to be inoperative), it is maritime salvage cases. Third, the law of maritime salvage is federal common law, and, as noted above, federal courts have exclusive jurisdiction in cases involving claims for salvage awards. Accordingly, state variation in law or courts is not an issue. Fourth, it seems reasonable to treat the federal courts as a single adjudicative body for purposes of maritime salvage cases: there is no split among the circuits (*Blackwall* is controlling precedent for all circuits); there are no specialty courts for maritime cases; and it generally is believed that federal courts are reasonably uniform in quality. Lastly, as noted above, although the criteria for determining a maritime salvage award are well defined and stable through time, there is no explicit formula or rule. This leaves open the possibility that courts are engaging in ALR or RLR.

I compare the ALR and RLR models according to their Bayesian information criteria, or BIC (Schwarz 1978). For a given model, the BIC is

$$BIC = l(\hat{\theta}) - \frac{1}{2} k \log T,$$

where $l(\hat{\theta})$ is the maximized value of the likelihood function for the model, k is the number of model parameters (i.e., the dimension of the parameter vector θ), and T is the number of observations. The basic idea behind the BIC is that it selects the model with the highest likelihood value (or best fit), subject to a penalty for lack of parsimony (or overfitting).

Tables 3 and 4 present the maximum likelihood estimates and BIC for the ALR model and the benchmark RLR model, respectively. (Note that in the benchmark RLR model, I set $\mathcal{P} = \{-2, -1, -0.5, 0, 0.5, 1, 2, ..., 5\}$, $m = 5$, and $(\alpha_1, \alpha_2) = (0.05, 0.05)$. Thus, although I allow for a fifth degree fractional polynomial (in the one continuos predictor, $\ln x_5$) with powers ranging from -2 to 5, the MFP algorithm selects a simple linear specification.[36]) In both models, the dependent variable is the natural logarithm of the salvage award and the independent variables are the court's findings of fact with respect to the six *Blackwall* factors (where the value of the salvaged property (factor

(9th Cir. 1974)).

[36] As a check of this selection, I estimated three alternative specifications of the RLR model. The first was a fifth degree convential polynomial (in $\ln x_5$). In the second alternative specification, I set $\mathcal{P} = \{-2, -1, -0.5, 0, 0.5, 1, 2, ..., 8\}$, thereby enlarging the set of powers. In the third alternative specification, I set $\alpha_2 = 1$, thereby forcing the MFP algorithm mto fit the best possible fifth degree fractional polynomial (in $\ln x_5$). None of the alternative specifications outperformed the benchmark model.

Table 3: ALR Model (dependent variable: ln y)

	Variable	Coeff	Std Err
x_1	Labor expended by salvors	16.49**	5.56
x_2	Skill displayed by salvors	1.60	2.54
x_3	Danger to salvors' property	6.16	3.95
x_4	Risk to salvors	0.71	2.25
$\ln x_5$	Value of salvaged property	4.97**	1.50
x_6	Danger to salvaged property	3.38*	1.75
β	Shape parameter for v	0.00	0.29
	Loglikelihood	-244.16	
	BIC	-263.63	

Notes: *Significant at 5% level. **Significant at 1% level. $\hat{\beta} = 1.3821 \times 10^{-7}$.

Table 4: Benchmark RLR Model (dependent variable: ln y)

	Variable	Coeff	Std Err
x_1	Labor expended by salvors	1.02**	0.20
x_2	Skill displayed by salvors	0.23	0.20
x_3	Danger to salvors' property	0.35	0.27
x_4	Risk to salvors	0.16	0.29
$\ln x_5$	Value of salvaged property	0.60**	0.05
x_6	Danger to salvaged property	0.59**	0.20
	Constant	9.41**	0.15
	Loglikelihood	-177.99	
	BIC	-197.46	

Note: **Significant at 1% level.

5) is log-transformed). The estimates for both models suggest that three factors are statistically significant to the determination of the salvage award—the labor expended by the salvors (factor 1), the value of the salvaged property (factor 5), and the danger to the salvaged property (factor 6)—with factors 5 and 6 commanding roughly equal weight and factor 1 commanding greater weight than factors 5 and 6.[37]

While interesting in their own right, the coefficient estimates are not my main concern. Rather, my main concern is the result of the BIC test. The BIC for the ALR model is -263.63, whereas the BIC for the benchmark RLR model is -197.46. This suggests that that the RLR model fits the data better than the ALR model. I discuss the implications of this result in Section 5.

4.4 Regression Tree Analysis

The final step of the empirical analysis is to perform a regression tree analysis of the maritime salvage cases. Regression tree analysis is a nonparametric method for analyzing the relationship between categorical or continuous independent variables and a continuous dependent variable (Bierman et al. 1984). I present the regression tree analysis as a supplement to the main empirical analysis. Although the regression tree analysis does not shed light directly on the question of whether the ALR or RLR model better fits the data, it serves as a robustness check of the coefficient estimates for both models.

Table 5 summarizes the regression tree model. As before, the dependent variable is the natural logarithm of the salvage award and the independent variables are the court's findings of fact with respect to the six *Blackwall* factors (where the value of the salvaged property (factor 5) is log-transformed). The tree growing method partitions the data to minimize within-node variance ("impurity"),[38] subject to four criteria/limitations: (i) only binary splits are allowed; (ii) the minimum decrease in impurity required to split a node is is 0.0001; (iii) the maximum number of levels beneath the root node is five; and (iv) the minimum number of cases in each node is ten.

[37] In addition, note that the estimate for the shape parameter β in the ALR model is sufficiently small such that μ is a quasimetric on the input space.

[38] Formally, the measure of impurity is least squares deviation (LSD), and is computed as

$$\frac{1}{N(\tau)}\sum_{t\in\tau}\left[y_t-\bar{y}\left(\tau\right)\right]^2,$$

where is $N(\tau)$ is the number of cases in node τ, y_t is the salvage award in case t, and

$$\bar{y}\left(\tau\right)=\frac{1}{N(\tau)}\sum_{t\in\tau}y_t$$

is the mean salvage award for cases in node τ.

Table 5: Summary of Regression Tree Model

Dependent variable:	$\ln y$
Independent variables:	x_1, x_2, x_3, x_4, $\ln x_5$, and x_6
Growing method:	CART
Impurity measure:	LSD
Split type:	binary
Minimum improvement:	0.0001
Maximum tree depth:	5
Minimum cases per node:	10

Figure 3 presents the regression tree analysis.[39] Like the ALR and RLR model estimates, the regression tree analysis suggests that three factors are key to the determination of the salvage award: the labor expended by the salvors (factor 1), the value of the salvaged property (factor 5), and the danger to the salvaged property (factor 6). Unlike the ALR and RLR model estimates, however, the regression tree analysis suggests that the value of the salvaged property (factor 5) is the most important factor.[40] Moreover, the regression tree analysis suggests how the factors interact. When the value of the salvaged property is low, this fact alone appears to dictate a small award. A somewhat larger award is made when the value of the salvaged property is moderate and the labor expended by the salvors is low. Higher awards are granted when either the value of the salvaged property is moderate and the labor expended by the salvors is high or the value of the salvaged property is high and the danger to the salvaged property is low. The highest awards occur when the value of the salvaged property and the danger to the salvaged property are high.

5 Implications and Limitations

The main conclusion of the empirical analysis is that the RLR model fits the data better than the ALR model. This conclusion is based on a comparison of the BIC of the two models. The key implication is that it is more likely that the data were generated by rule-based legal reasoning than by analogical legal reasoning. This implication, however, is subject to several limitations.

First, data on the inputs and outcomes of legal cases provides only indirect evidence regarding the method of legal reasoning. Nevertheless, it arguably is the best available

[39] The risk estimate is the mean within-node variance across the terminal nodes. Note that the tree did not require pruning to avoid overfitting.

[40] A model importance analysis suggests factor 5 is five times as important as factor 1 or factor 6.

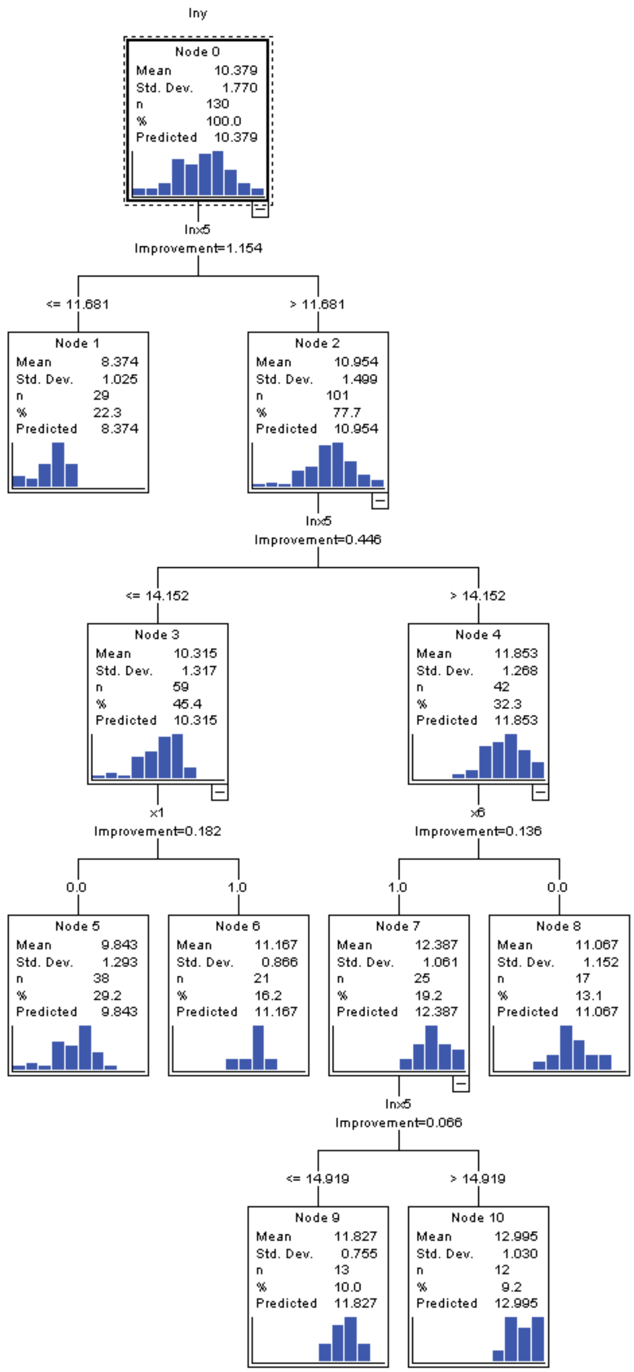

Figure 3: Regression Tree Analysis (risk estimate: 1.123)

evidence. In many cases, a court's written opinion offers no direct evidence regarding the method of legal reasoning. Even in cases in which the court's opinion offers some direct evidence, it rarely is definitive and, in any event, it arguably is of little probative value.[41]

Second, the two models are simplistic representations of ALR and RLR in their purest forms. Thus, not only are they highly stylized, they are rather extreme. It is quite possible that a combination or hybrid model, perhaps along the lines of a mixed SAR model (Anselin 1988), may fit the data better than either model. Examination of such a hybrid model would be intriguing direction for future research.

Third, both models take a representative agent approach. The ALR model assumes that all judges are equipped with the same similarity function, whereas the RLR model assumes that all judges apply the same legal rule. Allowing for heterogeneous judges surely would be more realistic, although tractability would require making strong assumptions about the structure of such heterogeneity.

Fourth, the ALR is less flexible than the RLR model. The ALR model is quite rigid in terms of structure. It specifies a particular method of assessment (similarity-weighted averaging) as well as a very specific functional form for the similarity function (equations (2)-(5)). The RLR model is more flexible. A fractional polynomial can approximate any sufficiently smooth function. It is conceivable that a more flexible specification (or even just a different specification) of the ALR could outperform the RLR model. Future research could explore the robustness of the BIC test result to alternative specifications of the ALR model.

Lastly, the empirical analysis in Section 4 offers not a statistical hypothesis test but rather a model selection exercise. The objective is to choose the best of the two competing models, without regard to whether either model is false. As such, the result of the BIC test cannot be interpreted either as a rejection of the hypothesis that the data were generated by ALR nor as a failure to reject the hypothesis that the data were generated by RLR. Rather, it should be interpreted as statistical evidence favoring the RLR model over the ALR model.

A further limitation of the empirical analysis is that it speaks only to the ability of the ALR and RLR models to explain the awards in maritime salvage cases. It says nothing about the ability of either model to explain the outcomes of cases in other areas

[41] "As a rule, we conceive of the judge's writing of an opinion as a procedure in which he justifies his decision. The writing coincides neither necessarily nor realistically with the process by which he reaches his decision, the process of discovery" (Murray 1982). There are (at least) two reasons to think that a court might use the language of RLR to justify its decision even if it engages in ALR in reaching its decision. First, "the language of 'rules' is much more efficient and parsimonious than that of 'cases'" (Gilboa and Schmeidler 2000). Second, "[r]ules are excellent justification mechanisms" (Hunter 2001).

of law. Moreover, the fact that both models are legalist models of judicial behavior suggests that they may not be well suited to other areas of law, including, in particular, politically charged areas (to which we might expect attitudinal or strategic models to be better suited).

Finally, the import of the empirical analysis is subject to an underlying conceptual question regarding the theoretical distinction between "analogical" and "rule-based" methods. One might question the distinction at two levels. At a general level, one might ask, if judges, as a rule, decide new cases via similarity-weighted averaging of prior cases, then is this not a "rule-based" method? The answer is no. The question speciously trades on an ambiguity in the meaning of the word "rule" in ordinary language. When jurisprudence scholars refer to "rule-based" methods, they mean methods that entail invoking generalizations (Schauer 1991; Alexander and Sherwin 2008; Schauer 2009). They do not mean any method that judges generally or even invariably use to decide new cases.

At a deeper level, one might ask, does not the process of making similarity judgments entail invoking generalizations (and thereby render "rule-based" all putative "analogical" methods)? This question is the subject of active debate among jurisprudence scholars. Skeptics argue that judges cannot make similarity judgments without invoking generalizations (Dworkin 1997; Alexander and Sherwin 2008).[42] Alexander and Sherwin (2008) state the skeptical view thusly:

> "Our point is that [a judge] cannot *reason* that [two cases] should be decided alike because they similar. To reason that they should be decided alike, she must determine that they are *importantly* similar, and to reason that they are importantly similar, she must refer to some general proposition . . . . [I]n order to draw analogies . . . [a judge] must refer to some general proposition that supports the analogy. . . . [T]he rules or principles that govern similarity, rather than the outcome of the precedent case, determine the result of the new case."

Dworkin (1997) makes the point succinctly: "An analogy is a way of stating a conclusion, not a way of reaching one, and theory must do the real work."

Nonskeptics insist that judges can and do make similarity judgments without invoking generalizations (Kamm 1997; Hunter 2001; Weinreb 2005). In the words of Hunter (2001), "analogy is a one-to-one similarity comparison that requires no generalization to operative effectively." Responding directly to Dworkin, Kamm (1997) writes:

[42] See also Eisenberg (1988), Eisenberg (1988), Greenwalt (1992), Posner (1990, 1995, 2006, 2008), and Alexander (1996, 1998). Milder skeptics include Levi (1949), Sunstein (1993, 1996), and Brewer (1996).

> "I disagree with Dworkin when he says that analogy is only a way of stating conclusions. Analogy can be a way of reaching a conclusion. The relevance of an analogous case can be clear, even if one does not have a theory that links the analogous case and the original case, and even if one is initially uncertain about what one may permissibly do in the analogous case. While we may need a theory to *explain* why case A is really more like case B than case C, we may still, without deep theoretical justification, see that case A is more like B than C and use that conclusion to help us find a solution to case A. Indeed, sometimes one reaches a conclusion about a case by way of an analogous case and still does not provide an adequate theoretical justification of one's position in either case."

Weinreb (2005) offers an extensive defense of the nonskeptical view. He rejects the skeptical view on two grounds. First, citing research in cognitive science and psychology, Weinreb posits that "the capacity for analogical reasoning is hard-wired in us," including the "idea of *relevant* similarity." Moreover, he asserts that analogical reasoning "cannot be assimilated or reduced" to rule-based reasoning because the latter depends on the ability to discern relevant similarity. According to Weinreb: "Unless one is able to identify an object as a member of a class despite its differences from other members of the class, no deductive inference is possible."

Second, Weinreb contends that the skeptics' argument "proves too much." He says:

> "By the same reasoning that would require a rule that makes the similarity on which an analogy rests relevant, so also would there have to be a rule for each and every one of the innumerable other similarities and dissimilarities between the two things compared. Otherwise, how would one know that beside the similarity to which the rule referred, there was not some other feature of one or both that also was relevant to the outcome, which would be changed accordingly?"

Weinreb accuses the skeptics of having it "backwards" when they argue that "the rule precedes and is essential to the validity of any analogy on which the decision rests." Instead, he avers that "the 'rule of the case' . . . is a generalized statement of the decision, not the predicate on which the decision rests," and that "[r]ather than the analogy depending on the rule, the rule depends on the analogy."

I offer the ALR model as a formal representation of the nonskeptical account of analogical legal reasoning. In doing so, I do not stake out a position in the debate between the skeptics and nonskeptics. Rather, I simply allow the possibility that the nonskeptical account is correct, and develop a formal representation using the apparatus of case-based decision theory, which embraces the nonskeptical view.[43] Under case-based decision theory, "the notion of similarity is primitive" (Gilboa and Schmeidler 2001).[44] A case-based decision maker does not engage in *explicit* induction, whereby one formulates general rules. Rather, she engages in *implicit* induction (Gilboa and Schmeidler 2000, 2001), whereby "similar past cases are implicitly generalized to bear upon future cases" (Gilboa and Schmeidler 2000).[45] Similarity-weighting averaging is one example of implicit induction. Under similarity-weighted averaging, although the decision maker "does not explicitly resort to general rules and theories," she "can be viewed as someone who believes in a general rule of the form $Y = f(X^1, ..., X^m)$ but does not know the functional form of f and therefore attempts to estimate it by nonparametric techniques" (Gilboa et al. 2006).[46]

6 Concluding Remarks

The use of analogical reasoning in law is a central topic in the jurisprudence and artificial intelligence and law literatures. Contributing to these literatures, this paper presents and empirically evaluates a formal model of analogical legal reasoning. The model posits that the outcome of the case at hand is a weighted average of the outcomes of prior cases, where the weights are a function of the fact similarity and precedential authority of the prior cases. To evaluate the model, I test its ability to explain the outcomes in U.S. maritime salvage cases vis-à-vis a simple model of rule-based legal reasoning (for which I turn to fractional polynomial regression). Comparing their Bayesian information criteria, I find that the RLR model fits the data better than the ALR model. A regression tree analysis of the maritime salvage cases supplements the main empirical analysis.

The work presented in this paper points to several avenues of further research. For instance, I would like to explore alternative ways to model ALR, including, for example,

[43] It is noteworthy that Alexander and Sherwin (2008) concede to Weinreb that if "it is in fact psychologically possible . . . to intuit important similarity without referrring to a supporting generalization, this decision is geneuinely analogical." As a normative matter, however, they lament that this is "a very impoverished view of judicial decision making, which we are reluctant to attribute to judges adjudicating in good faith."

[44] See also Gilboa and Schmeidler (2000). Like Weinreb, Gilboa and Schmeidler (2001) suggest that "[o]ur ability to discuss counterfactuals . . . relies on our subjective similarity judgments."

[45] As noted by Hunter (1998), a number of accounts of ALR conflate or equate analogy and explicit induction (e.g., Levi 1949; Posner 1990, 1995, 2008). In other accounts of legal reasoning, the relation between analogy and explicit induction (as well as deduction and abduction) are more nuanced (e.g., Sunstein 1993; Brewer 1996; Sunstein 1996).

[46] Recall the discussion on the connection between empirical similarity theory and kernel regression in Section 4.1.

similarity-weighted versions of other statistics, such as the median or the mode. As mentioned in Section 5, I also would like to investigate a hybrid model along the lines of a mixed SAR model. In addition, I would like to probe the extent to which non-legalist theories of judicial behavior could be formalized using statistical models. For example, I believe one could profitably model an attitudinalist judge as a Bayesian nonparametric statistician. Finally, I would like to examine areas of law other than maritime salvage. Although this likely would require further data collection on my part, one area in which potentially suitable data already have been collected is U.S. criminal confession cases.[47]

Appendix: Overview of U.S. Maritime Salvage Law

Throughout the colonial period, royal English courts decided salvage cases (Mangone 1997). Following independence, the U.S. Constitution granted federal courts original jurisdiction in "all Cases of admiralty and maritime jurisdiction,"[48] which include salvage cases. By the end of the nineteenth century, most salvage law concepts were generally settled (Mangone 1997). Although the United States is a party to both the 1910 Brussels Salvage Convention[49] and the 1989 London Salvage Convention,[50] "U.S. courts usually decide salvage controversies under the principles of the general maritime law without reference to international conventions" (Force 2004; see also Gilmore and Black 1975).[51]

Jurisdiction and Types of Actions

By statute, subject matter jurisdiction to adjudicate salvage claims lies with the federal courts.[52] When the matter at controversy is whether salvage is due and, if due, the amount, a federal court applying admiralty law is the only one in which such questions

[47] See Benesh (2002) and Kastellec (forthcoming).

[48] U.S. Const. art. III, § 2.

[49] International Convention for the Unification of Certain Rules Relating to the Salvage of Vessels at Sea (1910). Congress codified the 1910 Convention, with minor changes, as the Salvage Act of 1912 (Gilmore and Black 1975; Mangone 1997).

[50] International Convention on Salvage (1989).

[51] See, e.g., *Sobonis v. Steam Tanker Nat'l Defender*, 298 F. Supp. 631 (S.D.N.Y. 1969) (allowing salvage awards without reference to the Salvage Treaty).

[52] "The district courts shall have original jurisdiction, exclusive of the courts of the States, of: (1) Any civil case of admiralty or maritime jurisdiction, saving to suitors in all cases all other remedies to which they are otherwise entitled. (2) Any prize brought into the United States and all proceedings for the condemnation of property taken as prize." 28 U.S.C. § 1333. The Supreme Court assumed that salvage claims were within federal admiralty jurisdiction as early as 1804. See *Mason v. The Blaireau*, 6 U.S. (2 Cranch) 240 (1804); *Treasure Salvors, Inc. v. Unidentified, Wrecked and Abandoned Sailing Vessel*, 640 F.2d 560 (5th Cir. 1981) ("Claims arising out of salvage operation—efforts to rescue or recover ships disabled or abandoned at sea or to retrieve their cargo—are, unquestionably, within the admiralty jurisdiction of the federal courts.").

can be tried.[53] By contrast, an action based on quantum meruit, or for a case where a contract has been formed, may be tried in a state court, which has the authority to "assess damages based upon contract but cannot make a salvage award" (Norris 2008). A salvage suit is generally brought in rem against a ship, its cargo, or both because salvors, under maritime law, have a lien upon the property salved; if the ship or cargo is unavailable—due to destruction or removal from the jurisdiction—a salvor may seek remedy from the owner directly or in personam.[54] Under current rules, a suit in rem and in personam may be joined.[55] The federal court of appeals, when hearing an appeal of a federal district court's salvage award, "will not disturb a salvage award unless it is based on erroneous principles or a misapprehension of the facts or is so grossly excessive or inadequate as to be deemed an abuse of discretion."[56]

Types of Salvage Services

Salvage services fall into two categories: "contract" salvage and "pure" salvage.

Contract salvage occurs when the owner of property enters into an agreement with a salvor to rescue imperiled assets (Norris 2008). A salvage contract may be entered into before any emergency or after the ship or cargo is already in peril (Mangone 1997). The most common contract of this sort is the Lloyd's of London Open Form (LOF), although there is no obligation on any party to utilize this document to form a valid contract (Mangone 1997).[57] Although a court will closely examine a contract to make sure that neither side has taken advantage of an emergency to subject the other party to "grossly unfair terms" (Mangone 1997),[58] a contract wherein an owner has struck a "hard bargain" or where "the service was attended with greater or less difficulty than was anticipated, will not justify setting [the contract] aside."[59] Should a contract be thrown out, however, due to misconduct on the part of a shipowner or captain, innocent

[53] *Houseman v. The North Carolina*, 40 U.S. (15 Pet.) 40 (1841).

[54] *The Sabine*, 101 U.S. 384 (1879).

[55] *The G.L. 40*, 66 F.2d 764 (2d Cir. 1933).

[56] *Compania Galeana, S.A. v. Motor Vessel Caribbean Mara*, 565 F.2d 358 (5th Cir. 1978). See also *Oelwerke Teutonia v. Erlanger & Galinger*, 248 U.S. 521 (1919) ("Unless there has been some violation of principle or clear mistake, appeals to this Court concerning the amount of the allowance are not encouraged.").

[57] The LOF is four pages long, does not specify any sums, and proclaims its fundamental premise of "no cure-no pay" in large, bold letters. The LOF also contains provisions relating to enforcement of the contract through arbitration in London—a provision that U.S. courts have declined to enforce when the clause of the LOF is the only connection to an otherwise domestic operation between U.S. citizens. *Jones v. Sea Tow Services Freeport NY Inc.*, 30 F.3d 360 (2d Cir. 1994).

[58] This cuts both ways—the salvors could, of course, extort a favorable agreement, but the party in peril also could conceal the extent of danger or damage to its own advantage (Gilmore and Black 1975).

[59] *The Elfrida*, 172 U.S. 186 (1898).

crew members may still be entitled to an award.[60]

Pure salvage is rendered voluntarily in the absence of a contract. The reward for a person at sea who rushes to save another's property is "generously computed" as a matter of public policy (Gilmore and Black 1975). Because a salvor may never claim title to property by salvaging it, the need for incentive to save the property of another evolved into a salvage award (Mangone 1997).

Three elements are necessary for a valid pure salvage claim: (1) a marine peril; (2) service voluntarily rendered when not required as an existing duty or from a special contract; and (3) success in whole or in part, or that the service rendered contributed to such success.[61] The peril involved in a salvage operation does not necessarily have to be immediately impending—just subjecting the ship to the potential danger of damage or destruction.[62]

Because they already have a duty to aid their own ship in peril, crew members may not be considered "voluntary" salvors of their own ship[63]—nor may any person engaged in a profession which creates such a duty to salvage, such as firemen (Force 2004).[64] Regardless of heroic or costly measures to salvage property, any attempt that results in the complete loss of property will not be rewarded—nor will costs be reimbursed (Norris 2008).[65] A party may not force its services upon any owner or master of a vessel who

[60] *Jackson Marine Corp. v. Blue Fox*, 845 F.2d 1307 (5th Cir. 1988).

[61] *The Sabine*, 101 U.S. 384 (1879). Note the difference between simple "towage," which is for the simple convenience of another vessel in expediting her passage, and "salvage," which also includes an element of peril. *McConnochie v. Kerr*, 9 F. 50 (S.D.N.Y. 1881).

[62] *Fort Myers Shell & Dredging Co. v. Barge NBC 512*, 404 F.2d 137 (5th Cir. 1968). In the spectrum of peril, the degree of danger is immaterial—the degree "can affect the amount of the award, but not the establishment of a salvage service" (Norris 2008).

[63] They may, however, participate in salvage operations for reward if "their ship has been abandoned without hope of recovery, or the crew has been legally discharged from further services by the master." *Drevas v. U.S.*, 58 F. Supp. 1008 (D. Md. 1945). Common ownership of the salving vessel and the salved vessel does not necessarily preclude salvors from claiming award. 46 U.S.C § 80107(b). In contrast, passengers have "no duty to a vessel or its cargo" and cannot be compelled to assist in saving either (Mangone 1997). For this reason, passengers may "for extraordinary services, and the use of extraordinary means, not furnished by the equipment of the ship herself, by which she is saved from imminent danger . . . have salvage." *The Connemara*, 108 U.S. 352 (1883).

[64] *In re Iowa Fleeting Service, Inc.*, 211 F. Supp. 2d 794(M.D. La. 2002) ("Firefighters are precluded from obtaining a salvage award when the salvage work they perform is in the course of their existing duties as firefighters."). But see *Markakis v. S/S Volendam*, 486 F. Supp. 1103 (S.D.N.Y. 1980) (noting that even within certain professions, actions outside the line of duty may entitle a salvor to an award). Note, however, that nothing legally precludes the U.S. government from claiming salvage, even though it usually does not as a matter of policy (Gilmore and Black 1975).

[65] *Scott v. The Clara E. Bergen*, 21 F. Cas. 816 (D.S.C. 1882) (No. 12526a) ("All attempts, however costly, meritorious, or praiseworthy, go for nothing. In the event of failure, [the salvor] has to make his own repairs and pocket all losses, and he must give before he can get. He must save before he can ask to share what is saved. The owner, in fact and in law, can only be called upon to give to the salvor a portion of that very property which the salvor has saved for him; to restore only a portion of that which, but for the salvor, would have been lost to him. Thus it is the salvor who enables the owner to make

does not want assistance.[66]

The Supreme Court has made clear that only maritime property can be salvaged,[67] however the definition of maritime property can include any vehicle, cargo, or object with a nexus to traditional maritime activities.[68] The saving of life, in contrast to property, does not on its own confer an award of salvage (Gilmore and Black 1975).[69] To discourage disregard for human life in favor of saving property in times of emergency, however, U.S. law allows for salvors of human life to receive a "fair share" of the salvage award.[70] In contrast to the law of salvage, the law of "finds" controls property for which the title has been irrevocably lost and usually applies to ancient shipwrecks (Schoenbaum 2004). Historically, the court might award the title in place of a salvage award for property for which no owner came forward, but the Abandoned Shipwreck Act of 1987 transferred ownership of most such property to the government (Norris 2008).

Salvage Awards

In *The Blackwall*, 77 U.S. (10 Wall.) 1 (1869), the Supreme Court first listed the six factors for the determination of a salvage award: (1) the labor expended by the salvors in rendering the salvage service; (2) the promptitude, skill, and energy displayed in rendering the service and saving the property; (3) the value of the property employed by the salvors in rendering the service, and the danger to which such property was exposed; (4) the risk incurred by the salvors in securing the property from the impending peril; (5) the value of the property saved; and (6) the degree of danger from which the property was rescued. The value of any potential cargo of the salving ship is not a factor (Norris 2008).[71] By contrast, the value of freight and cargo of the salved ship are taken into

---

the payment.").

[66] *The Indian*, 159 F. 20 (5th Cir. 1908). In contrast, salvors may start operations on an abandoned vessel found at sea without prior authorization—with the hope of later reward (Mangone 1997).

[67] *Cope v. Vallette Dry-Dock Co.*, 119 U.S. 625 (1887) ("[N]o structure that is not a ship or vessel is a subject of salvage.").

[68] *Broere v. Two Thousand One Hundred Thirty-Three Dollars*, 72 F.Supp. 115 (1947) (finding that that money found on a human body floating on navigable waters was a proper subject of salvage); *Lambros Seaplane Base v. The Batory*, 215 F.2d 228 (2d Cir. 1954) (considering a seaplane which crashed in navigable waters a proper subject of salvage). But see *Provost v. Huber*, 594 F.2d 717 (8th Cir. 1979) (holding that house being dragged across a frozen lake had not embarked upon a "maritime adventure" and dismissing the salvage action for lack of a nexus with traditional maritime activities). Property considered salvable also traditionally includes any object which has been "thrown overboard" (jetsam), "found freely floating on the sea" (flotsam), "attached to buoys" (ligan), and "washed up to shore from the sea" (lagan) (Mangone 1997).

[69] Note that masters of ships at sea are obliged by statute to "render assistance to any individual found at sea in danger of being lost, so far as the master or individual in charge can do so without serious danger to the master's or individual's vessel or individuals on board." 46 U.S.C. § 2304(a)(1).

[70] 46 U.S.C. § 80107(a).

[71] See, e.g., *The Ereza*, 124 F. 659 (E.D. Pa. 1903).

account when computing the value of the property saved (Schoenbaum 2004). The burden of establishing the value of the salved property is on those seeking the award and should be assessed according to fair market value.[72]

There is no "precise formula" for computing salvage awards.[73] Although all of the factors should be considered in determining the award, each factor is not given equal weight (Force 2004). The trial court has considerable discretion in weighing the factors and making its determination on a case-by-case basis.[74] Courts do not view salvage awards "merely as pay, on the principal of a quantum meruit, or as a remuneration pro opere et labore, but as a reward given for perilous services, voluntarily rendered, and as an inducement to seamen and others to embark in such undertakings to save life and property."[75] Salvage awards are apportioned among co-salvors according to the “relative participation and risk” of each salvor (Norris 2008).[76] The statute of limitations for securing a salvage award is two years,[77] and the owners of both the ship and its cargo are liable to pay.[78]

As noted above, the salvage must be successful to merit an award. However, the actions of a salvor that worsen the position of the salved property may reduce the award, preclude it, or result in an award of damages (Force 2004). If the salvage is successful, but the operation causes some damage through ordinary negligence, the salvor is liable and the court will be reduce the award accordingly; if, however, the salvor causes damage through "gross negligence or willful misconduct," then the court may deny an award or even award affirmative damages.[79] Fraud or other dishonest conduct also may deprive a salvor of an award (Mangone 1997).

---

[72] *Nolan v. A. H. Basse Rederiaktieselskab*, 267 F.2d 584 (3d Cir. 1959).

[73] *Allseas Maritime, S.A. v. M/V Mimosa*, 812 F.2d 243 (1987).

[74] *The Emulous*, 8 F. Cas. 704 (C.C.D. Mass.) (No. 4480) ("And here, again, the court is asked to lay down some rules, by which to guide the parties in interest, underwriters as well as owners, in the ascertainment of the proper rate of salvage. That is asking the court to do, what it is utterly impracticable to do, to lay down rules, in cases admitting of an indefinite diversity of circumstances, and endless considerations of value, of perils, of services, and of merit. The subject is necessarily one, in which the reward must depend upon a just estimate of all the circumstances of each particular case."); *The Rescue v. the George B. Roberts*, 64 F. 139 (E.D. Pa. 1894) (“At best the award must be the result of an intelligent guess.”).

[75] *The Blackwall*, 77 U.S. (10 Wall.) 1 (1870). See also *B.V. Bureau Wijsmuller v. United States*, 702 F.2d 333 (2d Cir. 1983).

[76] See, e.g., *The Lydia*, 49 F. 666 (E.D.N.Y. 1892).

[77] 46 U.S.C. § 80107(c).

[78] In general, the owner of the cargo will be responsible for its own portion of the award and the owner of the ship for its portion; however, the court has discretion to apportion the award and the two owners may make other agreements between themselves before disbursement of the cargo (Schoenbaum 2004; Norris 2008).

[79] *Basic Boats, Inc. v. U.S.*, 352 F. Supp. 44 (E.D. Va. 1972); Schoenbaum (2004). Note that professional salvors are held to a higher standard of care than non-professionals (Mangone 1997).

References

Alexander, Larry. 1996. Bad beginnings. *University of Pennsylvania Law Review* 145: 57–87. 1, 2, 24

———. 1998. The banality of legal reasoning. *Notre Dame Law Review* 73:517–533. 1, 24

Alexander, Larry, and Emily Sherwin. 2008. *Demystifying legal reasoning*. Cambridge: Cambridge University Press. 1, 24, 26

Anselin, Luc. 1988. *Spatial econometrics: Methods and models*. Dordrecht: Kluwer Academic Publishers. 23

Ashley, Kevin D., and Stephanie Brüninghaus. 2006. Computer models for legal prediction. *Jurimetrics* 46:309–352. 2

Banerjee, Sudipto, Bradley. P. Carlin, and Alan E. Gelfand. 2004. *Hierarchical modeling and analysis for spatial data*. Boca Raton, FL: Chapman & Hall/CRC. 2

Bench-Capon, Trevor, James Freeman, Hanns Hohmann, and Henry Prakken. 2004. Computational models, argumentation theory and legal practice. In *Argumentation machines: New frontiers in argument and computation*, ed. Chris Reed and Timothy J. Norman, 85–120. Dordrecht: Kluwer Academic Publishers. 2

Bench-Capon, Trevor, and Henry Prakken. 2006. Argumentation. In *Information technology & lawyers: Advanced technology in the legal domain, from challenges to daily routine*, ed. Arlo R. Lodder and Anja Oskamp, 61–80. Dordrecht: Springer. 2

Bench-Capon, Trevor, Henry Prakken, and Giovanni Sartor. 2009. Argumentation in legal reasoning. In *Argumentation in artificial intelligence*, ed. Iyad Rahwan and Guillermo R. Simari, 363–382. Dordrecht: Springer. 2

Benesh, Sara C. 2002. *The U.S. court of appeals and the law of confessions: Perspectives on the hierachy of justice*. New York: LFP Scholarly Publishing. 27

Bierman, Leo, Jerome H. Friedman, Richard A. Olshern, and Charles L. Stone. 1984. *Classification and regression trees*. Belmont, CA: Wadsworth International Group. 3, 20

Billot, Antoine, Itzhak Gilboa, Dov Samet, and David Schmeidler. 2005. Probabilities as similarity-weighted frequencies. *Econometrica* 73:1125–1136. 2

Billot, Antoine, Itzhak Gilboa, and David Schmeidler. 2008. Axiomatization of an exponential similary function. *Mathematical Social Science* 55:107–115. 2, 5, 6, 9

Black, Ryan C., and James F. Spriggs II. 2009. The depreciation of precedent on the U.S. supreme court. Working Paper, Washington University in St. Louis, St. Louis, MO. 5

Bolhuis, Johan J., S. Bijlsma, and P. Ansmink. 1986. Exponential decay of spatial memory of rats in a radial maze. *Behavioral and Neural Biology* 46:115–122. 5

Brewer, Scott. 1996. Exemplary reasoning: Semantics, pragmatics, and the rational force of legal argument by analogy. *Harvard Law Review* 109:923–1028. 2, 24, 26

Cameron, Charles M., and Lewis A. Kornhauser. 2005. Modeling law: Theoretical implications of empirical methods. Working Paper, Princeton University, Princeton, NJ. 3, 14

———. 2009. Modeling collegial courts (3): Adjudication equilibria. NYU Law and Economics Research Paper No. 09-29, New York University School of Law, New York, NY. 3, 14

Cameron, Charles M., Jeffrey A. Segal, and Donald R. Songer. 2000. Strategic auditing in a political hierarchy: An informational model of the supreme court's certiorari decisions. *American Political Science Review* 94:101–116. 3, 14

Cross, Frank B. 2003. Decisionmaking in the U.S. circuit court of appeals. *California Law Review* 91:1459–1515. 1

Dasarathy, Belur V., ed. 1991. *Nearest neighbor (nn) norms: Nn pattern classification techniques*. Los Alamitos, CA: IEEE Computer Society Press. 2

Devroye, Luc, László Györfi, and Gábor Lugosi. 1996. *A probabilistic theory of pattern recognition*. New York: Springer. 2

Drezner, Zvi, and George O. Wesolowsky. 1989. The assymetric distance location problem. *Transportation Science* 23:201–207. 6

Dworkin, Ronald. 1997. In praise of theory. *Arizona State Law Journal* 29:353–376. 24

Eisenberg, Melvin Aron. 1988. *The nature of the common law*. Cambridge, MA: Harvard University Press. 1, 24

Epstein, Lee, and Jack Knight. 1998. *The choices justices make*. Washington, DC: CQ Press. 1

Force, Robert. 2004. *Admiralty and maritime law*. Washington, DC: Federal Judicial Center. 27, 29, 31

Fowler, James H., and Sangick Jeon. 2008. The authority of supreme court precedent. *Social Networks* 30:16–30. 5

Fried, Charles. 1981. The artificial reason of the law or: What lawyers know. *Texas Law Review* 60:35–58. 1

Gayer, Gabrielle, Itzhak Gilboa, and Offer Lieberman. 2007. Rule-based and case-based reasoning in housing prices. *B.E. Journal of Theoretical Economics* 7. 2, 4, 6, 12

Gilboa, Itzhak, Offer Lieberman, and David Schmeidler. 2006. Empirical similarity. *The Review of Economics and Statistics* 88:433–444. 2, 5, 6, 12, 13, 26

———. Forthcoming. A similarity-based approach to prediction. *Journal of Econometrics*. 2, 6, 12, 13

Gilboa, Itzhak, and David Schmeidler. 1995. Case-based decision theory. *Quarterly Journal of Economics* 110:605–639. 2

———. 1996. Case-based optimization. *Games and Economic Behavior* 15:1–26. 2

———. 1997. Act similarity in case-based decision theory. *Economic Theory* 9:47–61. 2

———. 2000. Case-based knowledge and induction. *IEEE Transactions on Systems Man, and Cybernetics—Part A: Systems and Humans* 30:85–95. 2, 23, 26

———. 2001. *A theory of case-based decisions*. Cambridge: Cambridge University Press. 2, 26

———. 2002. Cognitive foundations of probability. *Mathematics of Operations Research* 27:65–81. 2

———. 2003. Inductive inference: An axiomatic approach. *Econometrica* 71:1–26. 2

Gilboa, Itzhak, David Schmeidler, and Peter P. Wakker. 2002. Utility in case-based decision theory. *Journal of Economic Theory* 105:483–502. 2

Gilmore, Grant, and Charles L. Black. 1975. *The law of admiralty*. Mineola, NY: Foundation Press. 27, 28, 29, 30

Glaeser, Edward L., Bruce I. Sacerdote, and Jose A. Scheinkman. 2003. The social multiplier. *Journal of the European Economic Association* 1:345–353. 5

Greenwalt, Kent. 1992. *Law and objectivity*. New York: Oxford University Press. 24

Hume, David. 1748. *An enquiry concerning human understanding*. Oxford: Oxford University Press. 2

Hunter, Dan. 1998. No wilderness of single instances: Inductive inference in law. *Journal of Legal Education* 48:365–401. 26

———. 2001. Reason is too large: Analogy and precedent in law. *Emory Law Journal* 50:1197–1264. 23, 24

Kamm, F. M. 1997. Theory and analogy in law. *Arizona State Law Journal* 29:405–425. 24

Kastellec, Jonathan. Forthcoming. The satistical analysis of judicial decisions and legal rules with classification trees. *Journal of Empirical Legal Studies*. 3, 14, 27

Kornhauser, Lewis A. 1992a. Modeling collegial courts I: Path-dependence. *International Review of Law and Economics* 12:169–185. 3, 14

———. 1992b. Modeling collegial courts II: Legal doctrine. *Journal of Law, Economics, and Organization* 8:441–470. 3, 14

Kort, Fred. 1957. Predicting supreme court decisions mathematically: A quantitative analysis of the "right to counsel" cases. *American Political Science Review* 51:1–12. 3

———. 1963. Simultaneous equations and boolean algebra in the analysis of judicial decisions. *Law and Contemporary Problems* 28:143–163. 3

———. 1968. A nonlinear model for the analysis of judicial decisions. *American Political Science Review* 62:546–555. 3

———. 1973. Regression analysis and discriminant analysis: An application of R.A. Fisher's theorem in data in political science. *American Political Science Review* 67: 555–559. 3

Kort, Fred, and David Mars. 1957. *The administration of justice in connecticut*. Storrs, CT: Institute of Public Service, University of Connecticut. 3

Levi, Edward H. 1949. *An introduction to legal reasoning*. Chicago: University of Chicago Press. 1, 24, 26

Levy, Gilat. 2005. Careerist judges and the appeals process. *RAND Journal of Economics* 36:275–297. 1

Lieberman, Offer. Forthcoming. Asymptotic theory for empirical similarity models. *Econometric Theory*. 2, 6, 12, 13

Macagno, Francis, and Douglas Walton. 2009. Argument from analogy in law, the classical tradition, and recent theories. *Philosophy and Rhetoric* 42:154–182. 1

Mackaay, Ejan, and Piere Robillard. 1974. Predicting judicial decisions: The nearest neighbor rule and visual representation of case patterns. *Datenverarbeitung im Recht* 3:302–331. 3

Mangone, Gerald J. 1997. *United states admiralty law*. The Hague: Kluwer Law International. 27, 28, 29, 30, 31

Miceli, Thomas J., and Metin M. Coşgel. 1994. Reputation and judicial decision-making. *Journal of Economic Behavior and Organization* 23:31–51. 1

Murray, James R. 1982. The role of analogy in legal reasoning. *UCLA Law Review* 29: 833–871. 23

Norris, Martin J. 2008. The law of salvage. In *Benedict on admiralty*, ed. John D. Kimball and Thomas H. Belknap, Jr., vol. 3A. New York: Matthew Bender. 28, 29, 30, 31

Nosofsky, Robert M. 1986. Attention, similarity, and the identification-categorization relationship. *Journal of Experimental Psychology: General* 115:39–57. 5

Pagan, Adrian, and Aman Ullah. 1999. *Nonparametric econometrics*. Cambridge: Cambridge University Press. 2, 13

Posner, Richard A. 1990. *The problems of jurisprudence*. Cambridge, MA: Harvard University Press. 1, 24, 26

———. 1995. *Overcoming law*. Cambridge, MA: Harvard University Press. 1, 24, 26

———. 2006. Reasoning by analogy. Review of *Legal reason: The use of analogy in legal argument*, by Lloyd L. Weinreb. *Cornell Law Review* 91:761–774. 1, 24

———. 2008. *How judges think*. Cambridge, MA: Harvard University Press. 1, 24, 26

Post, David G., and Michael B. Eisen. 2000. How long is the coastline of the law? Thoughts on the fractal nature of legal systems. *Journal of Legal Studies* 29:545–584. 5

Riesbeck, Christopher K., and Roger C. Schank. 1989. *Inside case-based reasoning*. Hillsdale, NJ: Lawrence Erlbaum Associates. 2

Rissland, Edwina L. 1990. Artificial intelligence and law: Stepping stones to a model of legal reasoning. *Yale Law Journal* 99:1957–1981. 2

Rissland, Edwina L., Kevin D. Ashley, and L. Karl Branting. 2006. Case-based reasoning and law. *Knowledge Engineering Review* 20:293–298. 2

Rissland, Edwina L., Kevin D. Ashley, and R. P. Loui. 2003. AI and law: A fruitful synergy. *Artificial Intelligence* 150:1–15. 2

Royston, Patrick, and Douglas G. Altman. 1994. Regression using fractional polynomials of continuous covariates: Parsimonious parametric modelling. *Applied Statistics* 43: 429–467. 3, 14

———. 1997. Approximating statistical functions by using fractional polynomial regression. *Statistician* 46:411–422. 3, 14

Royston, Patrick, and Willi Sauerbrei. 2008. *Multivariable model-building: A pragmatic approach to regression analysis based on fractional polynomials for modelling continuous variables*. Chichester, UK: John Wiley. 3, 14

Sartor, Giovanni. 2005. Legal reasoning: A cognitive approach to the law. In *A treatise of legal philosophy and general jurisprudence*, ed. Enrico Pattaro, vol. 5. Dordrecht: Springer. 2

Sauerbrei, Willi, and Patrick Royston. 1999. Building multivariable prognostic and diagnostic models: Transformation of the predictors using fractional polynomials. *Journal of the Royal Statistical Society, Series A* 162:71–94. 14

Schauer, Frederick. 1991. *Playing by the rules: A philosophical examination of rule-based decision making in law and in life*. Oxford: Clarendon Press. 1, 24

———. 2009. *Thinking like a lawyer*. Cambridge, MA: Harvard University Press. 1, 24

Schoenbaum, Thomas J. 2004. *Admiralty and maritime law*. St. Paul, MN: Thomson/West. 30, 31

Schwarz, Gideon. 1978. Estimating the dimension of a model. *Annals of Statistics* 6: 461–464. 3, 18

Segal, Jeffrey A. 1984. Predicting supreme court cases probabilistically: The search and seizure cases, 1962-1981. *American Political Science Review* 78:891–900. 3

Segal, Jeffrey A., and Howard J. Spaeth. 1993. *The supreme court and the attitudinal model*. Cambridge: Cambridge University Press. 1

———. 2002. *The supreme court and the attitudinal model revisited*. Cambridge: Cambridge University Press. 1

Shephard, Roger N. 1987. Toward a universal law of generalization for psychological science. *Science* 237:1317–1323. 5

Smith, Joseph L., and Emerson H. Tiller. 2002. The strategy of judging: Evidence from administrative law. *Journal of Legal Studies* 31:61–82. 1

Sunstein, Cass R. 1993. On analogical reasoning. *Harvard Law Review* 106:741–791. 1, 2, 24, 26

———. 1996. *Legal reasoning and political conflict*. New York: Oxford University Press. 1, 2, 24, 26

Walton, Douglas. 2005. *Argumentation methods for artificial intelligence in law*. Berlin: Springer. 2

Weinreb, Lloyd L. 2005. *Legal reason: The use of analogy in legal argument*. Cambridge: Cambridge University Press. 1, 2, 24, 25

Westen, Peter. 1982. On "confusing ideas": Reply. *Yale Law Journal* 91:1153–1165. 1