

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

HARVARD

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

LAW MATTERS—LESS THAN WE THOUGHT

Daniel Klerman
Holger Spamann

Discussion Paper No. 1015

08/2019

Harvard Law School
Cambridge, MA 02138

This paper can be downloaded without charge from:

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

The Social Science Research Network Electronic Paper Collection:
<https://ssrn.com/abstract=3439526>

Law Matters – Less Than We Thought

Daniel Klerman[^] Holger Spamann[^]

Abstract: In a pre-registered 2×2×2 factorial between-subject randomized lab experiment with 61 federal judges, we test if the law influences judicial decisions, if it does so more under a rule than under a standard, and how its influence compares to that of legally irrelevant sympathies. The judges were given realistic materials and a relatively long period of time (50 minutes) to decide a run-of-the-mill auto accident case. We find weak evidence for the law effect, stronger evidence that rules constrain more than standards, and no evidence of a sympathy effect. Unexpectedly, we find that judges were more likely to choose the law that fully compensates an injured plaintiff.

August 13, 2019

[^] Edward G Lewis Professor of Law and History, University of Southern California Gould School of Law. dklerman@law.usc.edu.

[^] Lawrence R. Grove Professor of law, Harvard Law School. hspamann@law.harvard.edu.

We thank Scott Altman, Oren Bar-Gill, Iavor Bojinov, Nancy Gertner, Dan Kahan, Chris Robertson, Dan Simon, Alex Stremitzer, Symeon Symeonides, and participants in the Bolch Judicial Institute (Duke), Peking University Law School faculty workshop, Hong Kong University law and economics seminar, and at CELS 2018 (Michigan) and ALEA 2019 (NYU) for their advice and suggestions, Brooke Stanley and Colby Wilkinson for outstanding research assistance, John Goldberg and Denise Neary for facilitating the conduct of the experiment at the HLS/FJC Seminar, and especially the participating judges for their time. This research was approved by Harvard University's Committee on the Use of Human Subjects (#17-0428), and pre-registered on aspredicted.org (#3686). All data and code of this paper are available at <https://doi.org/10.7910/DVN/KXXA5C>.

1 Introduction

The extent to which legal rules bind judges, and, conversely, the extent to which judges can and do make decisions based on their preferences, ideology, and/or biases, has been vigorously debated in academia since at least the legal realists (e.g., Kantorowicz 1906; Frank 1930; Llewellyn 1940) and regularly features in popular discussions of the judiciary, particularly around judicial appointments. While judicial decisions exhibit regularity and claim to be based on authoritative legal sources and principles, the extent to which these authorities matter and induce that regularity is very much an open question. To date, however, much of this debate has remained anecdotal, based primarily on unsystematic observation of published judicial decisions. To the extent the literature has been systematically empirical, it has been divided into two main strands.¹ One strand analyzes published decisions to detect correlations with judges' political preferences (e.g., Segal and Spaeth 2002). Another strand performs vignette experiments with judges to investigate if judges are subject to standard psychological and/or ideological biases (e.g., Wistrich et al. 2014; Kahan et al. 2015). That is, the empirical literature on judicial decision-making to date has been focused on the effects of *non-law*. The only experiment with judicial subjects that directly tested the effect of *law* found no effect (Spamann and Klöhn 2016). Nevertheless, the legal variation in Spamann and Klöhn (2016) (a precedent) was mild, and the experimental setting (international law) was unfamiliar to most of the participants (U.S. federal judges). What if the legal variation were more powerful and the subject matter more familiar? In addition, to the extent the law does have an effect, does it matter whether it is framed as a rule or a standard (see Kaplow 1992; Schauer 2009)? Finally, even if there is a law effect, is it stronger than other effects that should be legally irrelevant?

In this paper, we present the results of an experiment that suggests a tentative “yes” to all three question. We find weak evidence that the law matters, stronger evidence that rules constrain more than standards, and no evidence that the pre-registered non-legal factor we tested for (sympathy) matters. On the other hand, the effect of the law is far from determinative even under a clear rule, and judges appear to be influenced by a preference for full compensation, a non-legal consideration we did not anticipate in our pre-registration. At the outset, we emphasize that, like any experiment, ours can only make statements about effect sizes for the particular variations we study. For example, there might be other legal rules and fact patterns that induce a stronger law effect. That said, we believe that our legal variation is quite powerful, so our finding of a muted law effect is revealing and perhaps concerning.

We conducted a pre-registered 2×2×2 factorial between-subject randomized lab experiment in which 61 federal judges spent up to 50 minutes deciding a run-of-the-mill auto accident case. Judges were given realistic materials, including opposing legal briefs and the full text of all cases, statutes and other sources cited in the briefs. The fifty minutes the judges were given to decide the case was much longer, and the materials much fuller, than comparable vignette experiments, which makes the experiment more realistic. We randomized the location (U.S. state) of the accident, the forum (U.S. state) of the litigation, and sympathy (legally irrelevant characteristics of the plaintiff and defendant that make one more appealing than the other).

In the pre-registered interpretation of our experimental treatments, the location treatment varies the applicable (state) law via choice-of-law principles, the forum treatment varies these principles from rule

¹ For an excellent recent survey, see Rachlinski and Wistrich (2017). For a more extensive discussion of the literature's relation to this paper's basic experimental design, see Spamann and Klöhn (2016)

to standard, and the sympathy treatment is legally irrelevant. In this interpretation, we find that the law effect is significantly stronger under the rule than the standard. Varying the location of the accident has virtually no effect under the standard, but even the very clear rule is not followed 23% of the time. This pre-registered interpretation is valid if judges do not have choice-of-law preferences. If they have such preferences (e.g. if they prefer the *lex loci delicti* rule to the Restatement 2nd standard), only the interaction of location and forum identifies a law effect (which we find, albeit of small magnitude), and we cannot test whether judges are more likely to follow the law under a rule or standard.

We find no evidence for a sympathy effect. Judges were no more likely to cap damages when the defendant was a caring paramedic than when he was a racist and convicted criminal (see Table 1 below for a full list of the differences between the two defendant/plaintiff pairs).

Nevertheless, in a test that we did not pre-register, we find that judges were much more likely to choose the law that fully compensates injured plaintiffs. That is, judges did not use their discretion to favor the more sympathetic party, but rather to implement the legal rule (full compensation) that they preferred. One interpretation of the results relating to sympathy and full-compensation is that judges are more influenced by ideological concerns (policy/legal preferences) than psychological factors (sympathy for a party based on appealing behaviors and characteristics). The fact that judges were not influenced by legally irrelevant characteristics of the parties provides some evidence that judicial training, selection, or norms induces judges to act in ways consistent with the idea that justice should be “blind” and treat all persons equally.

The key strengths of our design are its relatively high degree of realism and our manipulation of forum and accident location to induce legal variation. The latter overcomes potential problems with other ways of inducing legal variation. For example, the use of openly fictitious law would undermine the realism rationale that motivated us to employ real judges as our experimental subjects. Another alternative approach, the use of international law (as in Spamann and Klöhn (2016)), may not lead to accurate measurements of law following because American judges may not be familiar with international law or may not view it as “real” law. Finally, judges, as legal experts, might see through attempts by experimenters to manipulate the law by making legal assertions that judges may suspect or know to be false. By randomizing the forum and accident location, we can vary the “correct” legal result while still giving judges accurate legal materials in a familiar domestic context. Even though judges in our experiment are asked to decide the case under the law of states that they are not likely to be familiar with – Wyoming, South Dakota, Kansas, and Nebraska – that is not an unusual task for federal judges, who are often required as part of their diversity jurisdiction to decide cases under the law of various U.S. states.

A relatively high degree of realism is important because judges are highly trained, experienced, and selected individuals who operate in a uniquely structured and solemn environment with strong professional norms about judicial neutrality and adherence to the law. In particular, judges are especially trained, selected, and expected to discern legal commands from complex legal materials, and to ignore non-legal factors. At the same time, judges could also use their command of legal arguments to evade what to a layperson might appear to be a clear legal rule (Kennedy 1998).² Indeed, the interesting question

² Cf. Adam Liptak, *An Exit Interview with Richard Posner, Judicial Provocateur*, available at <https://www.nytimes.com/2017/09/11/us/politics/judge-richard-posner-retirement.html> (<https://perma.cc/2NP4-B6JM>) (quoting Judge Posner: “When you have a Supreme Court case or something similar, they’re often extremely easy to get around.”).

is not whether judges formally follow the law in the sense of citing legal sources to support their decisions (they do). Rather, the question is whether the overabundance of plausible legal arguments from statutes, precedents, regulations, and general principles allows the judge to reach whatever decision he or she wants while formally paying homage to the law. To capture this crucial aspect of judicial decision-making requires embedding the experimental task in rich facts and legal materials and giving judges sufficient time to weigh the facts and legal arguments. Alternatively, it may be precisely the legal materials and the parties' prompts in the briefs that trigger the judges' professional norms of rule-following and impartiality. Similarly, ample time may be necessary for judges to "think slow," distinguish weak from strong legal arguments, and set aside their prejudices and sympathies.

Whether one thinks that aspects of our experiment – active federal judges, realistic materials, and ample time – are likely to increase or decrease law following, it enhances the external validity of our experiment that our subjects are real judges, our materials are realistic, and the subjects were given nearly an hour to process the materials. Future experiments may show that a valid experimental design can abstract from certain features of our experiment because judges behave identically in the more abstract environment, because judges are not different from lay people, or because time does not affect this kind of decisionmaking. Or psychology may advance to a point where we can decompose judicial behavior into various building blocks from general psychology. For the time being, however, we have to treat judicial decision-making largely as a black box, so it is helpful for experimental purposes to recreate a realistic decision-making context.

The rest of this paper is structured as follows. Section 2 describes our experimental design. Section 3 reports the results. Section 4 concludes. Our invitations and other pre-session materials for participants are reproduced in Appendices 1-4, our main experimental materials in Appendices 5-7, our exit survey in Appendix 8, and our pre-registration statement in Appendix 9. A complete set of our experimental materials, data, randomization code, and analysis code is available at <https://doi.org/10.7910/DVN/KXXA5C>.

2 Design

2.1 Setting and Subjects

We conducted the experiment at the 2017 and 2018 annual three-day Law & Society Program organized jointly by the Federal Judicial Conference (FJC) and Harvard Law School (HLS). We had hoped to complete data collection in 2017, but repeated the experiment in 2018 because the number of 2017 participants (39) fell just short of the minimum we had preregistered (40). The FJC-Harvard Law & Society Program involves lectures by Harvard faculty and is open to all judges of the federal judiciary who apply to attend, subject to space constraints. Participants at the workshop and thus in the experiment included circuit, district, bankruptcy, and magistrate judges, and one judge from the Court of Federal Claims, whom we grouped with magistrate judges in our analysis to preserve anonymity. Spamann and Klöhn (2016) conducted their experiment in 2015 at the same program.

In 2017, we conducted the experiment in the classroom as part of a session on "Theories and Empirics of Judicial Decision-Making" offered by one of us (Spamann). Two weeks before the event, the judges received a letter from Spamann through the FJC informing them that 50 minutes of the session would be dedicated to an "experimental study of judicial decision-making," and inviting them to participate (Appendix 1). Four registered judges had participated in the actual experiment or a presentation of the

results of Spamann and Klöhn (2016) in the previous two years. To minimize the possibility of interference by the prior study, we sent the email reproduced as Appendix 2 to these four judges in advance of the session asking them not to talk to other Program participants ahead of the session. Shortly before the beginning of the session, we distributed the paper materials according to our randomization scheme and judges' name cards on their desks (See section 2.4 below). When the judges arrived for the class, Spamann read opening remarks (Appendix 3) that invited them to participate in the study by perusing the paper materials placed on their desks. Of the 43 registered participants, all but 4 accepted the invitation.

In 2018, the FJC asked us to conduct the experiment remotely in advance of the session in order to preserve classroom time. The FJC gave us the mailing addresses of anticipated participants four weeks before the program. We mailed each judge a cover letter (Appendix 4) inviting him or her to participate in a "study" in preparation for the session on "Empirical Studies of Judicial Decision-Making in the U.S. and Abroad." Along with the invitation, we sent a packet of materials, with treatments assigned according to the randomization scheme described below (See section 2.4). The cover letter and instructions informed judges that they could submit their answers either online at their convenience or on paper at the start of the program. A little more than half the judges to whom materials were sent participated in the study (See section 2.4).

In 2017, the first document the potential participants saw was the consent form. In 2018, participants saw the cover letter (described above) and then the consent form. The next document was the instructions, followed by the parties' briefs and a decision form. In addition, participants received a binder with all the legal sources cited in the briefs, including cases, statutes, and secondary materials. These materials are described in greater depth in section 2.2 below. In 2017, we distributed the exit survey (Appendix 8) to participants only after they returned the main study materials including the decision form, and we linked the decision form and the exit survey by noting a number on both forms when receiving one and handing out the other. In 2018, the exit survey was on the back of the decision form.

We pool the data from both years because participants from both years applied the correct law in virtually identical proportions and applied the damage cap in statistically indistinguishable proportions. In addition, we have no a priori reason to think that judges in different years were different from one another in relevant respects. On average, the 2018 participants were presumably either more motivated to participate or were less busy because they had to do the task in their spare time, but we have no reason to think this characteristic is related in any way to our treatments or to judicial decisions. In any event, we account for the separate randomization (see section 2.4) in our randomization tests (see section 3.2 and 3.3), and the inclusion of year dummies does not affect our regression results (see section 3.4).

2.2 The Case and Experimental Instructions

All participants were given facts relating to the same fictitious tort damages case and decided the same legal question – which state's substantive law should apply. If Kansas law applied, non-economic damages were capped at \$250,000. If Nebraska law applied, non-economic damages were not capped, and the defendant was liable for the full \$750,000 of non-economic damages, about half of which was not covered by his insurance.

Plaintiff Beatrice Parker and defendant Gary Rogers were friends. Parker was severely injured in a car accident in which Rogers was driving and Parker was a passenger in his car. Depending on the treatment, Rogers and Parker either met and lived in Nebraska and the accident occurred in Kansas, or vice versa.

Because the parties did not live where the accident took place, there was a choice-of-law question: whether Kansas or Nebraska law should apply (See section 2.3 for more details).

As discussed further below, judges were assigned randomly to one of two fora – Wyoming or South Dakota – in order to vary the applicable choice-of-law principle, the traditional *lex loci delicti* rule in Wyoming or the modern Restatement Second standard in South Dakota. The instructions (Appendix 5) asked participants to:

imagine you are sitting by designation in the District of [Wyoming / South Dakota]. One of the judges in the District of [Wyoming / South Dakota], Judge Frederick Simmons, had a stroke and is on disability leave. You have been assigned his cases. One of Judge Simmons's cases is *Parker v. Rogers*. Judge Simmons held a bench trial in that case and drafted Findings of Fact and Conclusions of Law. Damages, however, depended on whether Kansas's statutory cap on non-economic damages applied. Judge Simmons, therefore, ordered both parties to brief whether Kansas or Nebraska law should apply. Unlike Kansas, Nebraska does not cap non-economic damages. Judge Simmons asked both parties to submit their briefs simultaneously, without response or reply.

Participants were instructed to spend no more than 50 minutes on the task. The decision form asked them to indicate the "law that should apply," and provided several lines for "Reasoning." More details about the case and the legal question involved can be found in Judge Simmons's Findings of Fact and Conclusions of Law, which are reproduced in Appendix 6.

In addition to Judge Simmons's Findings of Fact and Conclusions of Law, the main materials furnished to participants were the two briefs that Judge Simmons requested and a binder of all the legal sources cited in the briefs. All materials were distributed in paper form because we thought that better reflected how most judges usually work, improving in this respect on Spamann and Klöhn (2016). We wrote the briefs ourselves. Each brief is about 3 pages long, with half of the first page consisting of the caption. The binder contains every legal source referenced in the briefs, including sources that the judges would be very unlikely to consult because their content is obvious or tangential. We produced a separate binder for each forum (Wyoming or South Dakota), adding filler pages to the South Dakota binder to make the two binders equal in thickness (about an inch). Each binder contained a table of contents and tabs for ease of access to the materials. The briefs also differed by forum because the legal argument was different, see section 2.3 below, but were otherwise identical for each treatment except that the arguments of plaintiff and defendant were swapped depending on where the accident took place and whose interest it was to argue for application of the law of the place of the accident. One half-sentence mentioning the circumstances of the accident when the defendant was not sympathetic also differed by treatment (see section 2.3 below). Sample briefs from one experimental condition can be found in Appendix 7.

2.3 Treatments

Our 2×2×2 factorial design cross-randomizes three factors: (1) the forum (the state where the case is litigated), (2) the accident location, and (3) the sympathetic party (plaintiff or defendant). We defer to section 3 a discussion of how these treatments relate to our research question regarding the effect of law.

By way of background, applicable substantive law in our experimental scenario (Kansas law capping damages or Nebraska law providing full compensation) depends on choice-of-law principles. Each state chooses its choice-of-law principles (see next paragraph), and a federal court is required to apply the

choice-of-law principles of the state where it is located.³ Depending on applicable choice of law, a court in one state (e.g. Wyoming) may be required to apply the substantive law of another state (e.g. Kansas). In general, which state's law applies depends on factors such as where the accident occurred and where the parties reside.

To begin with **factor (1) (forum)**, Parker sues Rogers in the state where Rogers moved after the accident, which is either Wyoming or South Dakota. Wyoming employs a choice-of-law *rule* to determine the applicable law in tort cases, the so-called *lex loci delicti* rule.⁴ Under this rule, judges are instructed to apply the law of the state where the accident occurred. By contrast, South Dakota employs a choice-of-law *standard*, the "most significant relationship" standard of the Restatement (Second) of Conflict of Laws.⁵ When both parties reside in the same state, South Dakota case law and the official commentary to the Restatement instruct judges to choose the law of the place where both parties reside, even if the accident occurred elsewhere. While any rule or standard has exceptions, and the party disadvantaged by the rule or standard will always argue for these exceptions, none is plausible in our case. In particular, application of the law of the state of the parties' "common domicile" to traffic accidents is the most widely accepted situation in which modern choice of law differs from older methods. In particular, when the legal issue involves "loss distribution" (e.g. damages) rather than "conduct regulation" (e.g. rules of the road), the consensus is that the applicable law is the law of the state where both parties reside, not the law of the state where the accident occurred (Symeonides 2006). This application of the common domicile rule is one of the leading examples in the Restatement's official commentary, probably its most celebrated deviation from the prior Restatement, and clearly acknowledged in South Dakota case law. That is, one might say that "common domicile" has morphed into a rule (See Kaplow 1992), even while the Restatement's open-ended language offers more wiggle room for a judge inclined not to follow it.

As to **factor (2) (accident location / domicile)**, Parker and Rogers either met and lived in Kansas but had the accident in Nebraska, or, in the other treatment condition, place of accident and residence are reversed. The location of the accident and parties' domicile determines applicable law, although how it does so depends on the relevant choice-of-law principle, which, in turn, as explained above, is determined by the forum (factor (1)). For example, if suit was in Wyoming, then Kansas's damage cap should apply if and only if the accident was in Kansas, because Wyoming's choice-of-law principle (*lex loci delicti*) instructs judges to apply the substantive law of the place where the accident occurred. Conversely, if suit was in South Dakota, then Kansas's damage cap should apply if and only if parties resided in Kansas (and thus if the accident was in Nebraska), because South Dakota's Supreme Court has chosen the Restatement Second approach to choice of law, which instructs judges to apply the substantive law of the state of common domicile.

As to **factor (3) (sympathy)**, we make either the plaintiff (P) or the defendant (D) sympathetic, and the other unsympathetic. Defendant Rogers is either a paramedic who inherited a small house and later moved to the forum state to care for an ailing aunt (condition D), or a convicted drug dealer who was just released from prison and inherited a large estate, and later moved to the forum state to work as a bouncer

³ *Klaxon Co. v. Stentor Elec. Mfg. Co.*, 313 U.S. 487 (1941).

⁴ *Ball v. Ball*, 269 P.2d 302, 304 (Wyo. 1954); *Duke v. Housen*, 589 P.2d 334, 342 (Wyo. 1979); *Archuleta v. Valencia*, 871 P.2d 198, 200 (Wyo. 1994); *Boutelle v. Boutelle*, 2014 WY 147, ¶ 12, 337 P.3d 1148, 1153 (Wyo. 2014).

⁵ *Andrews v. Ridco, Inc.*, 863 N.W.2d 540, 554 (S.D. 2015); *Chambers v. Dakotah Charter*, 488 N.W.2d 63, 67 (S.D. 1992). The Restatements for this and other areas of law purport to summarize the law prevailing in most American states and are compiled by the American Law Institute.

in a cousin’s bar (condition P). Plaintiff Parker is either a real estate agent who owns an SUV (condition D) or a nurse who owns a car (condition P). Rogers rents a Chevrolet Cruze (condition D) or a Chevrolet Camaro (condition P). Plaintiff offers to drive defendant to the house he inherited in her vehicle because he usually borrows his roommate’s car but his roommate needed the car on the relevant days (condition D) or because defendant had not had time to buy a car since his release from prison (condition P). Most importantly, the unsympathetic party gets into an argument with the manager of the rental car agency and uses a racial slur. These tidbits of information are interspersed in the case facts in a way that we believe to be realistic and non-intrusive. In particular, the racial slur finds its way into the case facts because it triggers a relevant delay in the pair’s travel plans. Table 1 summarizes the differences between the two sympathy conditions. The sympathy manipulation was designed to be as strong as possible without making the experimental suspects suspicious so that we would have power to detect an effect, if it existed, even with our small number of observations.

Table 1. Sympathy manipulations

	Condition D (Sympathetic Defendant & Unsympathetic Plaintiff)	Condition P (Sympathetic Plaintiff & Unsympathetic Defendant)
Defendant is	Paramedic	Convicted drug dealer, recently released from prison
Defendant inherited	Small house	Large estate
Plaintiff offered to drive Defendant in her vehicle because	Defendant’s roommate needed the car defendant usually used	Defendant had not yet bought a car since his release from prison
Defendant rents	Chevrolet Cruze	Chevrolet Camaro
Defendant moves	To take care of ailing aunt	To be a bouncer
Plaintiff is	Real estate agent	Nurse
Plaintiff owns a(n)	SUV	Car
Who utters racial slur?	Plaintiff	Defendant

2.4 Treatment Assignment

We performed separate stratified permuted block randomizations for the 2017 and 2018 waves of the experiment. We stratified by type of judge (circuit, district, bankruptcy, or magistrate) to ensure approximate balance of judge types across treatment groups. Within strata, randomization was blocked to ensure approximate balance in the size of treatment groups, except, as discussed below, we underweighted Wyoming in 2017. Perfect balance was not achievable because judge strata sizes were not usually divisible by 8. To the extent there were “leftover” judges from different strata, we formed a separate block from those “leftovers” and permuted the 8 treatment combinations in a way that minimized differences in the number of participants – overall and from each stratum – assigned to each of the two levels of each factor.

Our final sample is more unbalanced, however, because logistical considerations forced us to assign potential participants to treatment conditions before obtaining participants' consent.⁶ Potential participants could, and did, decline to participate after our randomization. While only 4 out of 43 judges present did not participate when the experiment was conducted in the classroom in 2017, 21 out of 43 did not participate in 2018, when the experiment was offered as a pre-class activity. Moreover, 3 of the 4 non-participations in 2017 were concentrated in one treatment arm as described below. Nevertheless, we have no reason to think that non-participation was influenced by the treatment assignment and hence treat "missing" judges as random and hence ignorable in our analysis.⁷

Another problem stemming from the paper-based distribution of randomly assigned packages was that some packages must have been misdirected during distribution in 2017 because we ended up with two more participants in the South Dakota / Nebraska treatment arm than foreseen by our randomization (see table 2 below). We treat this as an accidental additional randomization and thus include these observations in our analysis, except in the randomization-based inference because it requires knowledge of the explicit assignment mechanism.

We initially underweighted the Wyoming forum in an attempt to maximize power. In 2017, we randomized only 2 judges from each stratum—8 in total—into the Wyoming forum treatment in a way that ensured that one judge from each stratum was assigned to each level of the other two main factors.⁸ Our idea was that Wyoming's choice-of-law rule was so clear that it would perfectly or nearly perfectly determine participants' decisions so that the residual variance under the rule would be zero or near zero. At the same time, balance of each judge type across the levels of the two remaining factors would have avoided confounding the two remaining main effects by judge type (See Mead, Gilmour, and Mead 2012, ch. 15). This would have allowed us to document a strong law effect under the rule while preserving most observations, and hence power, for the detection of law and sympathy effects under the standard where we thought variance would be much higher, and for comparison of law and sympathy effects under the standard to those under the rule.⁹ We would have had power to detect a law effect at the 5% level under the rule even if two of the 8 judges did not complete the experiment but the remaining six followed the law, or at the 10% level even if one did not decide in accordance with the law but all others did and completed the experiment. Alas, out of 43 judges present in 2017, all three who did not complete the experiment happened to be in the rule (Wyoming forum) condition. In addition, there were three judges who completed the experiment but did not provide his/her judge type and for whom we could not infer judge type from treatment, and one of these three was from the Wyoming (rule) treatment.¹⁰ To avoid

⁶ In 2018, this was inevitable as we had to mail treatment-specific packages to judges before knowing whether they would participate. In 2017, the tight schedule in the classroom and our inability to obtain binding participation promises did not allow us to determine participation before performing the randomization.

⁷ The materials in all treatment conditions were identical in overall appearance including length, and differences in treatment were sufficiently subtle that it seems unlikely that the choice not to consent was based on the treatment.

⁸ For example, the two bankruptcy judges randomized into Wyoming would have been assigned either to (Kansas, Sympathy) and (Nebraska, No Sympathy), or to (Nebraska, Sympathy) and (Kansas, No Sympathy), but not to, e.g., (Kansas, Sympathy) and (Kansas, No Sympathy).

⁹ If the residual variance under the rule had been zero, the variance for testing interaction effects of rule vs. standard and the other main effects would have been equal to the residual variance for the main effects under the standard, which we aimed to minimize by assigning as many observations as possible to that condition.

¹⁰ While the randomization had been stratified by judge type, for logistical reasons, we did not have an ability to note judge type on the materials we handed to each participant, relying instead on the exit survey.

such problems in 2018, we abandoned the underweighting we practiced in 2017 and aimed for equal group sizes in Wyoming (rule) and South Dakota (standard) treatments.

The first number in each cell of Table 2 below shows the number of each type of participant that we assigned to each of the 2x2x2=8 treatment combinations. The next number in each cell, in parentheses, shows the number who actually participated, and finally, in square brackets, the last number in each cell shows the number who chose Kansas law (the cap on damages). While we will discuss this last number (our results) in detail later, it may be helpful for the reader to remember that when the forum is Wyoming, judges are supposed to apply the traditional, *lex loci delicti* rule, so Kansas law should apply if and only if the accident was in Kansas. Conversely, if the forum is South Dakota, judges are supposed to apply the Restatement Second’s “most significant relationship” standard, which according to South Dakota Supreme Court precedents, means judges should apply Kansas law only if the accident was in Nebraska, because the experiment was set up so that the accident was in Nebraska if and only if parties lived in Kansas. When there were zero participants in a cell, we indicate the number choosing Kansas law as “[–]” because “[0]” might be misinterpreted as application of Nebraska law.

Table 2: Number of judges assigned (participated) and [chose Kansas law] by treatment and judge type

Forum	Treatment combinations							
	WY	WY	WY	WY	SD	SD	SD	SD
	KS	KS	NE	NE	KS	KS	NE	NE
Sympathy	P	D	P	D	P	D	P	D
2017								
Circuit	1 (0)[–]			1 (0)[–]	2 (2)[0]	2 (2)[1]	2 (2)[0]	2 (2)[1]
District		1 (1)[1]	1 (0)[–]		3 (3)[1]	5 (3)[0]	3 (3)[1]	4 (4)[1]
Bankruptcy		1 (1)[1]		1 (1)[1]	2 (2)[2]	1 (1)[0]	3 (2)[0]	1 (1)[1]
Magistrate (missing)	1 (1)[1]		1 (0)[–] (1)[0]		2 (2)[1]		1 (1)[1]	2 (2)[0] (2)[1]
2018								
Circuit	1 (0)[–]	1 (0)[–]	1 (0)[–]			1 (1)[0]	1 (0)[–]	1 (0)[–]
District	2 (1)[0]	1 (0)[–]	1 (1)[0]	1 (0)[–]	1 (1)[0]	1 (0)[–]	1 (0)[–]	2 (1)[0]
Bankruptcy	2 (0)[–]	2 (0)[–]	2 (2)[0]	3 (2)[0]	3 (2)[1]	2 (2)[1]	3 (3)[0]	2 (2)[1]
Magistrate (missing)	1 (1)[0]	1 (0)[–]	1 (0)[–]	1 (1)[0]	1 (1)[1]	1 (1)[0]	1 (0)[–]	1 (0)[–]
Total	8 (3)[1]	7 (2)[2]	7 (4)[0]	7 (4)[1]	14(13)[6]	13(12)[3]	15(11)[2]	15(12)[4]

Judge type is “missing” for three participants in 2017 because four participants did not answer the corresponding question in the exit survey, and we did not have an ability to mark the judges’ type directly on their decision form. For one of these four, we could infer the type from a comparison of treatment assignments and the decision forms that we received (and that did contain a coded reference to the treatment assignment); this judge’s type is not listed as “missing” in Table 2. Of the three remaining participants whose type is “missing” in Table 2, the one in forum-location-sympathy combination WY-NE-P must be district or magistrate, given our assignment list, and we use both possibilities in our randomization-based tests (section 3.2 and 3.3) with identical results. The two participants with irretrievable judge type are omitted from the randomization-based tests.

2.5 Realism and Awareness Checks

To ensure that the task was realistic and feasible in the 50 minutes allotted, we performed several checks before and after the experiment.

Before the experiment, we pre-tested the case with several advanced law students and two former judges. All reported that the case seemed realistic and was entirely feasible – most reported spending far less than 50 minutes.

After the experiment, we asked our participants several questions in the exit questionnaire reproduced in Appendix 8. First, we asked “Did the case seem realistic to you? If not, why not?” All but five participants answered this question. All answered “yes,” except one answered “No,” two remarked that real facts tended to be messier, and two noted that time was too short (although in discussion after the 2018 session, the consensus in the room seemed to be that the time was unrealistically short only if a full opinion were expected). Second, the exit questionnaire also asked specifically if the participant had “ever had to choose between the law of the place of the accident or of the parties’ common domicile, as in *Parker v. Rogers*,” which 17 participants answered in the affirmative.

To check if participants guessed what we were trying to test, we also asked in the exit questionnaire: “What do you think this study was about?” Six participants answered that they thought we were testing for a sympathy effect, and three of them even mentioned Rogers by name (all of them were in the D condition where Rogers was unsympathetic). Four more gave a perhaps related, more abstract answer such as “Heart vs. application of the law.” Some judges answered we were testing the effect of the law, and very few even mentioned the effect of precedent. Most, however, either explicitly stated that they did not know what we were testing, or mentioned unrelated issues such as the order of documents presented as their guess as to our research question. To our surprise, many seemed to think that the case was a close one. Consistent with this, in response to another exit question, the judges expressed low certainty that other judges decided the case as they did: the modal 5% bin was 50%, the median 60%, and the mean 65%.

Due to the close relationship of our design to Spamann and Klöhn (2016), we also asked “I administered a similar experiment at this workshop two years ago, and talked about it last year. Did you take part in this workshop last year or the year before, or did you otherwise hear or read about my prior experiment?” 54 participants answered this question. One participant answered “yes,” another “yes – forgotten what we decided,” and a third “I did go to this workshop before, but did not participate in a study or recall any discussion;” everyone else answered “No.”

As we pre-registered, we exclude in robustness checks either or both of the participant who answered “yes” to the last question and the participant who answered “no” to the question whether the case seemed realistic. As our results are unaffected, we do not report these checks individually below.

3 Results

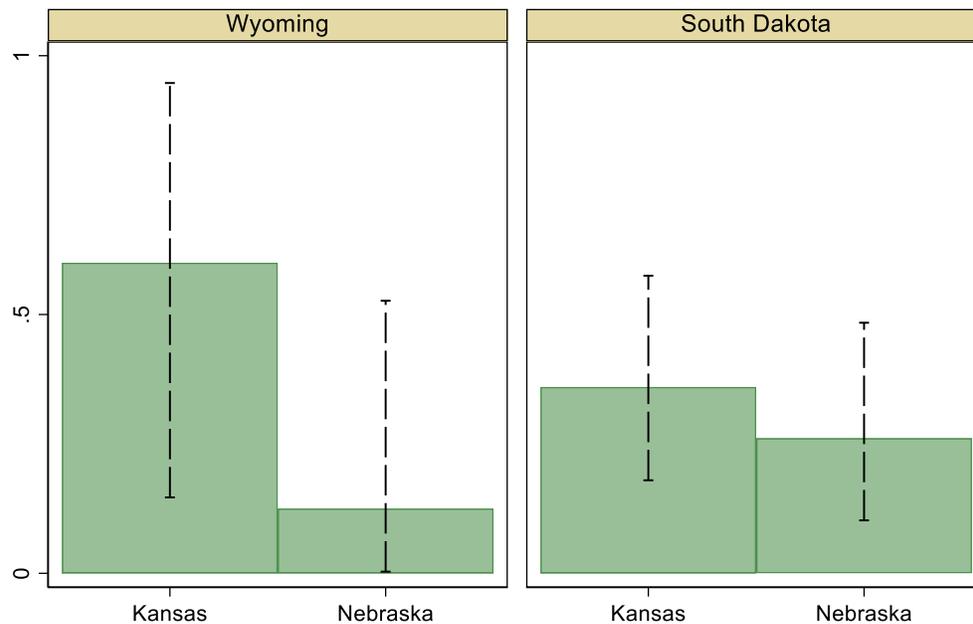
We first discuss our results informally using simple graphs (section 3.1). Then we analyze the data more rigorously using three kinds of statistical tests: pre-registered exact tests (section 3.2); exact tests under an alternative interpretation (section 3.3); and regressions (section 3.4). The results are largely the same, but we cannot determine whether law effects are larger under a rule or standard under the alternative interpretation.

As a preliminary matter, we note that, in principle, exact tests are preferable in a sample as small as ours because they do not rely on large-sample approximations or functional form assumptions. That said, in a factorial design, randomization-based tests for individual factors still require specifying a null for interactions with the other factors unless the sample is balanced and factorial effects are additive (See Dasgupta et al. 2015), neither of which is the case in our design. In particular, with binary outcomes, factor effects can hardly be additive.

3.1 Informal and graphical analysis

Our basic results can be gleaned by looking at Figure 1, which summarizes most of the last row of Table 2. The law appears to matter only under a rule (Wyoming), and even then it far from completely determines outcomes. In reading the graphs, the reader should bear in mind that the confidence intervals for individual groups are misleadingly wide for purposes of comparison of groups because they do not impose the restrictions of the null hypothesis. We discuss rigorous tests in the next sections.

Fig. 1: Rate choosing Kansas law (cap), by Forum and Accident state



Spikes indicate exact 95% c.i.

The left panel of Figure 1 shows the results when the forum was Wyoming, which applies the traditional, *lex loci delicti* rule, pooling together the two sympathy treatments. Kansas law is clearly chosen much more often when the accident was in Kansas (60%) than when it was in Nebraska (12.5%). This is as it should be under the *lex loci delicti* rule, and hence evidence of a law effect. In the next subsection, we show that this difference is almost statistically significant at conventional levels, notwithstanding the large overlap in the confidence intervals. At the same time, the law effect is not nearly as strong as we expected it to be. Given that the rule is clear and no exceptions plausibly apply to the experimental fact pattern, the left bar should extend to one, and the right bar should be at zero. That is, when the accident was in

Kansas, Kansas law should always apply. Conversely, when the accident was in Nebraska, Nebraska law should always apply, so Kansas law should never apply.

The right panel of Figure 1 shows the results when the forum was South Dakota, which applies the Restatement Second's "most significant relationship" *standard*. We again pool together the two sympathy treatments. The evidence of a law effect here is much weaker because the bars are of almost equal height and the difference between them, in fact, goes in the wrong direction. The right bar should be higher under South Dakota law, because the standard has been interpreted by the South Dakota Supreme Court to require application of the law of the place of common domicile. So the left bar should be at zero and the right bar should be at one. That is, when the accident was in Kansas, the judges should have applied Nebraska law because the fact pattern always had the parties reside in Nebraska when the accident was in Kansas. Conversely, the right bar should be at one, because when the accident was in Nebraska, the judges should apply Kansas law because the fact pattern had the parties reside in Kansas when the accident was in Nebraska. In fact, the judges applied Kansas law more often when the accident was in Kansas than when the accident was in Nebraska. It is clear that the judges did not apply the Restatement Second in accordance with South Dakota Supreme Court precedent.

The fact that, when the forum was South Dakota, judges applied Kansas law more often when the accident was in Kansas might suggest that judges ignored the instruction to apply the Restatement Second and instead applied the traditional *lex loci delicti* (place of accident) rule. This implies an alternative way of testing the law effect: do judges make different decisions under the Restatement Second than under the *lex loci delicti* rule? That is, instead of asking whether judges make different decisions based on where the accident was (the pre-registered question we explore in Section 3.2), we can ask whether judges make different decisions based on the forum's choice-of-law approach. We explore that question rigorously in Section 3.3.

One could, of course, argue that the Restatement Second's "most significant relationship" test gives judges lots of discretion and that even two South Dakota Supreme Court cases holding that the place of common domicile was the place with the most significant relationship still gave the judges some wiggle room. Perhaps, for example, judges could legitimately take into account the fact that the defendant changed his domicile after the accident, so, at the time of adjudication, the parties no longer had a common domicile. Or perhaps renting the car in the accident state tipped the balance toward that state being the one with "the most significant relationship." Or perhaps judges simply thought the accident location was the "most significant relationship" in spite of South Dakota Supreme Court precedent to the contrary. Nevertheless, under those interpretations one would still expect that, given random assignment of accident location, whatever factors induced judges to apply Nebraska law 74% of the time when the accident was in Nebraska should have similarly induced them to apply Kansas law 74% of the time when the accident was in Kansas, subject to sampling variation. Yet the judges applied Kansas law only 36% of the time when the accident was in Kansas. A Fisher exact test rejects equality of these two rates with $p=0.011$. So the judges' choices are not consistent with even this looser interpretation of law-following under the Restatement Second's most significant relationship standard.

It is notable that, whether Wyoming or South Dakota was the forum, most deviations from the "correct" legal result were in the direction of *not* applying Kansas law. That is, judges had a tendency to award full compensation (Nebraska law) and *not* to apply the cap (Kansas law). If one looks at the left panel for Wyoming, the judges nearly always applied Nebraska law when they were supposed to (right bar), but

refused to apply Kansas law 40% of the time when that was what the rule required. Similarly, when South Dakota was the forum (right panel), judges applied Kansas law less than half of the time they were supposed to (left bar), but applied Nebraska law most of the time they were supposed to (see the right bar showing that judges applied Kansas law only 26% of the time and thus applied Nebraska law 74% of the time). Thus, as analyzed more rigorously in Section 3.4, there seems to be evidence that judges acted in accordance with a preference for full compensation.

Figure 1 does not differentiate by sympathy. The reader can verify in table 2, however, that judges applied Kansas's cap approximately as often when the plaintiff was sympathetic (9 out of 31 participants) as when the defendant was sympathetic (10 out of 30 participants). Thus, as analyzed more rigorously in the next subsection, there is no evidence supporting the idea that judges were overall influenced by sympathy.

3.2 Pre-Registered Exact Tests

We now report our pre-registered exact statistical tests of the effects discussed in the preceding discussion of figure 1.¹¹

We begin with simple effects of location and sympathy separately within each forum (ignoring possible interactions between location and sympathy). For these effects, we pre-registered the unconditional Boschloo exact test, a conceptually superior generalization of the better-known conditional Fisher exact test (Mehrotra et al. 2003). The Boschloo test is uniformly more powerful than the Fisher test.

Nevertheless, neither the location test for Wyoming nor for South Dakota rejects the null hypothesis that judges were equally likely to apply Kansas law and its damages cap regardless of accident location.¹² As noted above, these are the primary tests of the law effect. If judges applied choice-of-law principles in a legalistic way, varying the accident location should have changed whether Kansas's cap on damages applied. Nevertheless, even the accident location (law) effect in the Wyoming forum attains only a p -value of 0.11, which is not significant at conventional levels, and the p -value for South Dakota is much larger (0.48). While the lack of significance for the Wyoming treatment may appear unsurprising given the small sample of 13 observations in that arm of the experiment, we note again that, if the effect of Wyoming's clear rule had been as strong as we expected and one arguably should expect, then we would have had power to detect it. For example, if at least 11 of the 13 judges in the Wyoming condition had followed Wyoming's clear choice of law rule, the p -value would have been below 0.03. In other words, the lack of significance of the location treatment even in Wyoming reflects our observation above that even a clear rule is disregarded with non-trivial frequency. In addition, the number of observations for the South Dakota treatments is much larger (48), so the inability to detect a statistically significant location (law) effect in that arm of the experiment is notable.

Similarly, neither the sympathy test for Wyoming nor for South Dakota rejects the null hypothesis that judges were equally likely to apply Kansas law and its damage cap regardless of sympathy. That is, there is no evidence that judges were more likely to apply Kansas's cap on damages when the defendant's

¹¹ In addition to the two sets of tests discussed in the main text, we also pre-registered that we would compare the residual variance between the rule and the standard, but this test is redundant where the outcome variable is binary, as here.

¹² In our pre-registration (point 2.1), we wrote that we would code the outcome variable for the location effect test as whether the judge applied the "correct" law. This is obviously nonsensical, however, because the test would then show no effect of the treatment precisely when the law has the strongest effect.

behaviors and characteristics were more sympathetic than the plaintiff's. The p -values are very large – 0.20 for Wyoming and 1.00 for South Dakota – and not even close to statistical significance at conventional levels.

The second set of exact tests we pre-registered are randomization-based Fisher tests¹³ of the *difference* between fora in the sizes of the effects we just tested separately by fora, namely of accident location and sympathy.

For accident location (the strength of the law effect), the test is thus a comparison between the left and the right panel of figure 1, and in particular a comparison of the *difference* between the two bars in each panel. To be more precise, since we are interested in comparing law-following under a rule and a standard, we compare left bar minus right bar in the left panel to right bar minus left bar in the right panel (because, as explained above, under the law, the left bar should have been high in Wyoming and the right bar in South Dakota). This difference of Wyoming ($3/5 - 1/8 = 0.47$) minus South Dakota ($6/23 - 8/23 = -0.09$) equals 0.56.¹⁴ A difference of this magnitude would arise by random chance with a probability less than

¹³ The idea of these tests is simple: sample from the randomization distribution of the chosen test statistic by repeatedly (1) re-randomizing the participants to treatments using our assignment mechanism (i.e., the assignment mechanism we actually used in the experiment), (2) imputing for each participant the outcome that we would expect in the re-randomized treatment condition under the sharp null hypothesis; and (3) calculating the test statistic (see generally, e.g., Athey & Imbens 2016). We take 1,000,000 draws by re-running the actual random assignment script used in the experiment 100,000 times and associating the resulting assignments with with 10 independent random shufflings of the list of participating judges. In this process, we treat one magistrate judge as district judge because that is what we had accidentally done with this judge in the actual experiment as well. As mentioned in section 2.4, we must omit two participants with missing judge type from the randomization-based tests, and we get identical results whether we treat a third judge with missing judge type as district judge or a magistrate judge (the only two possibilities given this judge's treatment assignment).

In our pre-registration, we failed to attend to two important details regarding points (2) and (3) that arise from sample unbalance in a factorial structure. With a balanced sample, one can calculate and then average effects for one factor cell-pair by cell-pair while remaining completely agnostic about other factorial effects and interactions (e.g., Dasgupta et al. 2015). With sample unbalance, however, any other factorial effects do not automatically cancel out, and different cell weightings are plausible and outcome-relevant. One could assume that all factorial effects are zero, that the factorial effects are additive and equal to the observed effects except for the factor being tested, or that the remaining factorial effects are equal to some externally imposed null. (It is not possible to limit the re-randomization to one factor, i.e., only among participants assigned to a particular combination of the levels of the other two factors, because the randomization protocol in the actual experiment by necessity combined randomization for all three factors. The randomization was also stratified, i.e., performed separately for different types of judges, so the only re-randomization possible would be within judge type, of which there are always very few and sometimes only one or none in each cell.) Similarly, one could calculate averages for each of the $2 \times 2 \times 2 = 8$ cells separately and then weight each cell equally, or one could first pool observations for each of the $2 \times 2 = 4$ factor combinations relevant for each of the two interaction tests discussed in the text. With respect to the first point, for lack of a better alternative, we take as the null that neither factor 2 nor 3 has any effect, and in particular that each judge would always choose *the same state's law* regardless of accident location and sympathy (See next subsection). With respect to the second point, we employ the second option (pool observations for each of the $2 \times 2 = 4$ factor combinations) because the small cell sizes for some treatment combinations would add a lot of noise to the statistic. That said, our reported results would be qualitatively similar—and for the law effect, slightly stronger—if we instead calculated separate effect sizes for each level of the other factors and then averaged over them. Finally, we assume that judges' participation decisions as well as the missing judge type information were independent of treatment assignments and hence ignorable.

¹⁴ As explained in section 2.4, these numbers and the randomization tests ignore the two participants with unknown judge type because the re-randomization requires stratification by judge type.

0.03, i.e., $p < 0.03$, under the sharp null hypothesis that each judge would have applied the same law regardless of forum and accident location (and sympathy). This suggests that the law effect is stronger under the rule (Wyoming) than under the standard (South Dakota), in accordance with our expectation and most of the literature. The reason why the *difference* can be significant even though the individual effects are not significant is that the South Dakota effect goes in the opposite direction (its sign is negative). That is, when the forum was South Dakota, judges chose Kansas law more often when the accident was in Kansas, even though authoritative interpretation of South Dakota law would suggest the opposite. So it is easier to reject equality of the two effects than to reject that either is different from zero.

The estimated difference in any sympathy effect (i.e., the tendency to apply the cap more often when the defendant is sympathetic) between the rule (Wyoming: $3/6 - 1/7 = 0.36$) and the standard (South Dakota: $6/22 - 8/24 = -0.06$) is 0.42. If anything, according to this estimate, sympathy matters *more* under the rule than the standard, unlike what one would expect if the rule constrained judicial discretion and bias more effectively. Nevertheless, the difference is statistically insignificant, with a two-sided randomization-based p -value of 0.11.

To account for the multiple hypotheses testing, we employ the false discovery rate (FDR) control method of Benjamini and Hochberg (1995) (BH). When running m tests, the method consists of finding the highest p -value p_i among these tests such that $p_i \leq \alpha/m$, where α is the desired FDR. Setting $\alpha = 10\%$, our rule-vs.-standard result ($p < 0.03$) passes the BH threshold if we think of our program as involving three tests (law, sympathy, rule-vs.-standard) but not if we think of it involving four or more tests (e.g., because we test rule-vs.-standard separately for law and sympathy). The BH method also does not account for the fact that our tests statistics are necessarily dependent, in some cases negatively (e.g., law effect under the standard and rule-vs.-standard effect), which would require a further downward adjustment of the threshold (Benjamini and Yekutieli 2001). We skip such subtleties because we think our main result is the absence of a strong law effect, and because our main tests rely on an interpretation of the design that we now revisit in the following subsection.

3.3 An Alternative Interpretation of Our Design

The interpretation of our design underlying the pre-registered tests just reported is not the only possible one. The interpretation that we pre-registered and that was assumed in the previous subsection is that:

- 1) accident location identifies an effect of the law within each forum because each forum's choice-of-law principles require choosing the law by reference to the accident location: in Wyoming, the *lex loci delicti* rule requires choosing the law of the place of accident, and in South Dakota, given our experimental design, the Restatement Second's most significant relationship standard requires choosing Kansas law when the accident was in Nebraska and vice versa, and
- 2) forum identifies the effect of rules vs. standards: the difference in accident location effect between the two forums can be understood as measuring the difference in "bindingness" between Wyoming's rule and South Dakota's standard.

The null hypothesis generating this first interpretation is that, if the law (concretely, applicable choice-of-law principles) did not matter, each judge would always apply a state's law with the same probability

(most likely 0 or 1), be it the law of Kansas or Nebraska, regardless of accident location.¹⁵ It does not matter why a judge would do this. The judge might just choose applicable law randomly. Or the judge might prefer one of the states in general, or might prefer one state's substantive law (caps or no caps on damages), although, of course, such preferences are legally irrelevant. Against this null, any effect of accident location identifies an effect of the law because, by the null hypothesis, accident location would not matter if the law did not matter. Similarly, the only relevant change between the two fora is that one has a choice-of-law rule (Wyoming) while the other has a standard (South Dakota). The change in content of choice-of-law approach (*lex loci delicti* vs. common domicile) flips the expected direction of any law effect but is otherwise orthogonal to judicial behavior.

We have come to think, however, that this interpretation is not the only plausible one because the underlying null hypothesis is not the only plausible description of what all judges would do but for the effect of the law (i.e., but for the applicable choice-of-law principle). Even when given full discretion, at least some judges might be expected to adjust their choice of substantive law to the geographic facts of the case. That is, it is plausible that at least some judges have choice-of-law preferences, i.e., that they prefer applying either the law of the accident state or the law of common domicile. For such judges, law matters if they are willing to disregard their personal choice-of-law preference when instructed to do so by the forum's law. More to the point, for such judges, a change in location confounds a law effect with an effect of the judge's own preferences (e.g. their preference to apply the law of the place of accident, a normative preference, but a preference nonetheless). For such judges, only a change in the applicable choice-of-law principle itself (the forum treatment) cleanly identifies a law effect, and the change in legal form (rules vs. standards) is not separately identified.

In this alternative interpretation of our design, the right test of a law effect is the comparison of the fraction of judges applying the *lex loci* in Wyoming, where it is required by the applicable rule, and in South Dakota, where the applicable choice of law principle applied to our facts demands *not* applying the *lex loci*. The null hypothesis is that the applicable choice of law principle has no effect, i.e., that judges are equally likely to apply the *lex loci* in Wyoming and South Dakota. The one-sided alternative hypothesis of law-following is that deciding the case under Wyoming's *lex loci* rule increases application of the *lex loci* relative to South Dakota's standard favoring the opposite, common domicile. We do indeed find that the fraction of judges applying *lex loci* is higher in Wyoming (10/13=77%) than in South Dakota (26/48=54%). The difference is 23%, and Boschloo exact and randomization-based one-sided *p*-values against the null of no difference are 0.08 and 0.05, respectively.¹⁶ That is, the law effect is small in absolute size—certainly far from 1, which would be implied by perfect law-following—and in fact barely statistically distinguishably from zero. This finding under the alternative interpretation of our design is consistent with that under the pre-registered interpretation in as much as the latter also found evidence of a law effect just at the boundary of conventional statistical significance, albeit on the other side (*p*=0.11).

¹⁵ The sharp null hypothesis necessary for the randomization inference in the previous subsection is slightly more restrictive in that it assumes each judge would always choose the same law, ruling out random choice. But standard inference would be possible against the random choice null.

¹⁶ The randomization tests recalculate the difference in the rate of applying *lex loci* between Wyoming and South Dakota for 1,000,000 re-randomizations of our participants under the sharp null that each participant would have applied or not applied, as the case may be, *lex loci* regardless of treatment assignment. This also imposes the null of no interaction with sympathy. See generally *supra* footnote 13.

3.4 Regressions

Finally, we show the full set of factorial effects in a linear probability regression framework. We present these regressions for those accustomed to reading regression results, as a robustness check to the univariate exact tests to account for any confounding induced by the unbalanced sample, and as a way of testing for effects that we did not preregister, including the preference for full compensation. We caution, however, that the standard errors are of dubious reliability given the small sample and the binary outcome. Although the errors are necessarily heteroskedastic given the data's binary nature, we present standard OLS *t*-statistics to be conservative because the heteroskedasticity-robust standard errors (which rely on large-sample approximations) were smaller in our small sample. To preserve the exact correspondence of the rotated regressions as explained immediately below, we do not include dummies for judge type and the 2017/2018 rounds but have verified that including them does not change the results and that the dummies do not themselves have significant coefficients.¹⁷

The dependent variables and all three factors are contrast coded (-1 or 1), which means that (a) the estimated main effects (for F (forum), A (accident location), and S (sympathy)) represent ½ the difference in means between the respective groups, and (b) each regression equation results from the other by a mere rotation of the coordinate system.¹⁸ The latter point means that, as is appropriate, all three regressions show the same result in a different perspective. In particular, with respect to the effect of law, models 1-3 correspond to the perspective taken in the preceding subsections 3.1-3.3, respectively: model 1 considers application of Kansas's cap as in section 3.1 and figure 1; model 2 considers law following defined by a combination of the applicable choice of law principle and the accident location as in our pre-registered interpretation and section 3.2; and model 3 considers application of *lex loci*, where a difference between fora identifies a law effect under the alternative interpretation offered in section 3.3. Model 1 also corresponds to the analysis of sympathy in sections 3.1 and 3.2.

¹⁷ With bankruptcy or federal claims judges and 2017 as the omitted categories, of the 12 coefficients for three judge categories (circuit, district, or magistrate) and one year (2018) for three dependent variables, only the coefficients on district judge and 2018 with Kansas law as the dependent variables have *t*-statistics large enough that they would attain 10% significance in a single-coefficient test, but this is insignificant if one adjusts for the multi-testing involved.

¹⁸ Specifically, with contrast coding, $K = FAC = -AL$, $C = FAK = -FL$, $L = -AK = -FC$, and each variable multiplied with itself yields 1. (For the meanings of K, C, L, F, A, and S, see Table 3 row and column headings.) This means that one can move from one estimated equation to the next by multiplying by one of the factors, which will associate each coefficient with a different independent variable but leave their magnitude unchanged.

Table 3: Regressions

	(1)	(2)	(3)
	Kansas law (cap) (K)	Correct law (C)	<i>Lex loci</i> (L)
Forum (F) (SD = 1, WY = -1)	-0.09 (-0.60)	-0.32** (-2.16)	-0.22 (-1.49)
Accident location (A) (NE = 1, KS = -1)	-0.32** (-2.16)	-0.09 (-0.60)	0.30* (2.00)
Sympathy (S) (D = 1, P = -1)	0.21 (1.44)	0.19 (1.31)	0.01 (0.09)
F x A	0.22 (1.49)	-0.30* (-2.00)	0.09 (0.60)
F x S	-0.24 (-1.65)	-0.01 (-0.09)	-0.19 (-1.31)
A x S	-0.01 (-0.09)	-0.24 (-1.65)	-0.21 (-1.44)
F x A x S	0.19 (1.31)	0.21 (1.44)	0.24 (1.65)
Constant	-0.30* (-2.00)	0.22 (1.49)	0.32** (2.16)
<i>N</i>	61	61	61

OLS *t* statistics in parentheses. * $p < 0.10$, ** $p < 0.05$

The results are consistent with our univariate exact tests. To begin with the perspective taken in section 3.2 and model 2, judges are more likely to misapply the law under the standard than under the rule, as indicated by the statistically significant coefficient estimate of -0.32 on Forum in model 2. Also as in section 3.2, we can find an effect differential but no significant main effect of law in this perspective: the constant coefficient of 0.22 in model 2 is statistically insignificant.

However, if we take the perspective of section 3.3 and model 3, that same coefficient of 0.22 is now, after rotation, the coefficient for Forum and tests the law effect of Wyoming's and South Dakota's choice of law principles. While the coefficient in the regression table is not statistically significant, that is because the regression implicitly performs a two-sided test. As explained in section 3.3, it is appropriate to perform a one-sided test, in which case the p -value is 0.07 ($\Pr(t_{53} < -1.49) = 0.07$), almost exactly the same as in the univariate tests reported there.¹⁹

¹⁹ When we include judge-type and year dummies, the one-sided p -value rises to 0.13, reflecting the loss of 10 degrees of freedom (whereas the point estimate of -0.19 barely changes).

Lastly, like the univariate tests, none of the regressions provide substantial evidence of a sympathy effect, whether in the main effect or in any interactions.

Beyond our primary tests, one finding that emerges from the regressions is that the judges seem repelled by Kansas's cap, as can be seen, for example, in the negative and (marginally) statistically significantly constant when application of Kansas's cap on noneconomic damages is the dependent variable. That is, our experimental subjects seem to have a preference for full compensation. While we had not considered this in designing the experiment, it is consistent with judicial attitudes more broadly. Damage caps are always legislative in origin and never originate in judicial decisions. In fact, judges in ten states were so hostile to damage caps that they invalidated them on constitutional grounds that seem rather weak, such as equal protection and the right to jury trial (Stein 2019, vol. 3, sections 19:4-10).

4 Discussion

The experiment provides some support for the idea that application of a rule is more predictable than application of a standard and that judges under a rule feel more bound to reach the correct results. The law effect, however, is weak at best. The standard in our experiment seems to have little influence on decisions even though it was authoritatively interpreted by the relevant state supreme court. Even the clear *lex loci delicti* rule did not bind judges completely, and we cannot reject at conventional levels the hypothesis that judges under the rule decided randomly. We also can only weakly reject the hypothesis that judges ignored the applicable choice of law entirely. There is little evidence that judges were motivated by sympathy based on legally irrelevant factors, which is reassuring. Judges do, however, seem to have and to act upon a preference of full compensation.

The results of this experiment leave at least two puzzles. First, why, unlike Wistrich et al. (2014) and Spamann and Klöhn (2016), did we not find a sympathy effect? Second, why did judges follow the law, especially the clear *lex loci delicti* rule, so much less often than we expected? A possible answer to the first question is that the sympathy variation we study here is the relatively minor type of personal difference that judges are trained to ignore, whereas Spamann and Klöhn (2016) studied much more morally charged war crimes and the defendant's attitudes to those crimes. A possible answer to the second question may be that judges were less familiar with choice of law than we had expected and that judges made decisions in our experiment without the law clerks who would ordinarily assist them. While we had thought that most judges would be experienced in choice-of-law analysis, only 31% (17 out of 55 who answered this exit question) had "ever had to choose between the law of the place of the accident or of the parties' common domicile, as in *Parker v. Rogers*." Consistent with this, judges' median prediction of the proportion of other judges that decided like they did was only 60%. This unexpected lack of familiarity and certainty partly reflects the large number of bankruptcy and magistrate judges among our subjects. In addition, when federal judges decide complex legal issues, like those involving choice of law, they ordinarily discuss them with their law clerks, and law clerks usually write the first draft of opinions. It is possible that law clerks, who are usually recent law school graduates, are more legalistic in their orientation, and that their influence causes judges to follow the law more consistently.

Reference List

- Athey, Susan, and Guido Imbens. 2016. The Econometrics of Randomized Experiments. Working paper.
- Benjamini, Yoav, and Yosef Hochberg. 1995. Controlling the False Discovery Rate: A Practical and Powerful Approach to Multiple Testing. *Journal of the Royal Statistical Society B* 57:289–300.
- Benjamini, Yoav, and Daniel Yekutieli. 2001. The Control of the False Discovery Rate in Multiple Testing Under Dependency. *Annals of Statistics* 29:1165-1188.
- Dasgupta, Tirthankar, Natesh S. Pillai, and Donald B. Rubin. 2015. Causal inference from 2^k factorial designs by using potential outcomes. *Journal of the Royal Statistical Society B* 77:727–753.
- Frank, Jerome. 1930 [2009]. *Law & The Modern Mind*. New Brunswick, NJ: Transaction Publishers.
- Kahan, Dan M., David Hoffman, Danieli Evans, Neal Devins, Eugene Lucci, and Katherine Cheng. 2015. “Ideology” or “Situation Sense”? An Experimental Investigation of Motivated Reasoning and Professional Judgment. *University of Pennsylvania Law Review* 164:349-439.
- Kantorowicz, Hermann [Gnaeus Flavius]. 1906. *Der Kampf um die Rechtswissenschaft*. Carl Winter’s Universitätsbuchhandlung.
- Kaplow, Louis. 1992. Rules versus Standards: An Economic Analysis. *Duke Law Journal* 42: 557-629.
- Kennedy, Duncan. 1998. *A Critique of Adjudication (fin de siècle)*. Harvard University Press.
- Llewellyn, Karl N. 1940. On Reading and Using the Newer Jurisprudence. *Columbia Law Review* 40:581-614.
- Mead, R., Gilmour, S.G., and A. Mead. 2012. *Statistical Principles for the Design of Experiments*. Cambridge University Press.
- Mehrotra, Devan V., Ivan S. F. Chan, and Roger L. Berger. 2003. A Cautionary Note on Exact Unconditional Inference for a Difference between Two Independent Binomial Proportions. *Biometrics* 59:441-450.
- Rachlinski, Jeffrey J., and Andrew J. Wistrich. 2017. Judging the Judiciary by the Numbers: Empirical Research on Judges. *Annual Review of Law and Social Science* 13:203-229.
- Schauer, Frederick. 2009. *Thinking Like a Lawyer*. Harvard University Press.
- Segal, Jeffrey A., and Harold J. Spaeth. 2002. *The Supreme Court and the Attitudinal Model Revisited*. Cambridge University Press.
- Spamann, Holger, and Lars Klöhn. 2016. Justice is Less Blind, and Less Legalistic, Than We Thought: Evidence from an Experiment with Real Judges. *Journal of Legal Studies* 45:255-280.
- Stein, Jacob. 2019. *Stein on Personal Injury* (3d. ed, April 2019 update).
- Symeonides, Symeon C. 2006. *The American Choice-of-Law Revolution: Past, Present and Future*. Leiden: Martinus Nijhoff.

Wistrich, Andrew J., Jeffrey J. Rachlinski, and Chris Guthrie. 2014. Heart Versus Head: Do Judges Follow the Law or Follow Their Feelings? *Texas Law Review* 93:855-923.

Appendix 1: 2017 Invitation Letter

Re: Invitation to Participate in an Experiment at the Harvard/FJC Law & Society Program

Dear Judge:

I look forward to welcoming you to the session on “Theories and Empirics of Judicial Decision-Making” at the Harvard/FJC Law & Society program on April 10 at 10:45am.

In agreement with the FJC, I will use the first 50 minutes of my session for an experimental study of judicial decision-making. I hope you will participate in the study, but participation is entirely voluntary.

If you do participate, you will be asked to judge a fictitious case, and provide brief reasons for your decision on a piece of paper. The goal of the study is to learn about legal reasoning and the role of various factors therein. I am not testing your knowledge of, or opinions about, particular legal issues.

Your answers will be anonymous. All relevant legal materials will be provided to you.

The attached form describes the study as required by Harvard’s Institutional Review Board (IRB). If you choose to participate, the IRB also requires that I obtain your signature on this form to document your consent. I will distribute paper copies of the form at the event, but if you wish, you may also bring a signed form with you. Your name and signature will be kept entirely separate from the study document, and will not be linked to your study answers in any way.

If you have any questions about the study, please do not hesitate to contact me. I look forward to meeting you on April 10.

Appendix 2: 2017 Email to Repeat Participants

[Subject: FJC/HLS Seminar, April 10]

Dear Judge ____:

Denise Neary (FJC) told me you were planning to attend the FJC/Harvard Law & Society Seminar again on April 10. I look forward to welcoming you back.

I am reaching out to you because I will again offer a study as I did two years ago, and discussed last year. If you were in attendance at those sessions, you are still very welcome to participate in this year's study, which is different from the last one. That said, to minimize interference of the two studies with one another, it would be great if you did not discuss the last study with other seminar participants until after my session. I would really appreciate your cooperation.

Sincerely yours,

Holger Spamann

Appendix 3: 2017 Opening Remarks

Good morning, your honors.

My name is Holger Spamann, and I am a professor here at the law school.

It is truly a great pleasure and a privilege to be able to talk to you today.

As you know from Denise and from my letter, I have prepared an exercise for you, and I suggest we get started on that right away so that we have enough time for discussion later.

You will find in front of you a binder and an envelope with your name on it. If the name is not yours, please let my assistants know and we will change it.

The name is on the envelope only so that we could keep track of the materials as we distributed them. Your name is not on any other document and will not be linked to your answers in any way. Your answers will be completely anonymous.

The only document that is not anonymous is the consent form, which is the first document you will find in the envelope. The form is required by Harvard's rules for research. Please sign the consent form if you want to participate. When we collect the materials, we will separate the consent form from the other materials, so that your name will not be linked to your answers.

Everything else is explained in the instructions in the envelope. You have up to 50 minutes to complete the exercise. When you are done, please return the full stack of materials to me or one of my two assistants, Brooke and Jeremy. We will then give you an exit survey to complete. Once everyone is done, I will debrief.

Please begin. I hope you will find this interesting and educational.

Appendix 4: 2018 Invitation Letter

Re: Invitation to Participate in a Study in Preparation for the Harvard/FJC Law & Society Program

Dear Judge X:

I look forward to welcoming you to the session on “Empirical Studies of Judicial Decision-Making in the U.S. and Abroad” at the Harvard/FJC Law & Society Program on April 25.

The most interesting data are those that we generate ourselves. I invite you to participate in the enclosed study so that we can discuss it at the session. The maximum time you should spend on the study is 50 minutes, and you can do it at any time before you arrive at Harvard. You can submit your answers online at [link specific to treatment condition], or write them on the enclosed form [which contained a file number that encoded the treatment condition] and hand the form to me on the first morning of the workshop (April 23).

Your answers will be anonymous. The study is designed to generate information about judicial decision-making in general. I am not testing your knowledge of, or opinions about, particular legal issues.

I hope you will participate in the study, but participation is entirely voluntary. The attached form describes the study as required by Harvard’s Institutional Review Board (IRB).

If you have previously participated in a session that I offered at the Harvard/FJC Program or read about it, you may still participate in the study if you wish, but you should know that I will not use your answers in analyzing and discussing the results.

If you have any questions about the study, please do not hesitate to contact me. I look forward to meeting you on April 25.

Appendix 5: Instructions for Participants

Instructions

Please imagine you are sitting by designation in the District of [Wyoming / South Dakota]. One of the judges in the District of [Wyoming / South Dakota], Judge Frederick Simmons, had a stroke and is on disability leave. You have been assigned his cases.

One of Judge Simmons's cases is Parker v. Rogers. Judge Simmons held a bench trial in that case and drafted Findings of Fact and Conclusions of Law. Damages, however, depended on whether Kansas's statutory cap on non-economic damages applied. Judge Simmons, therefore, ordered both parties to brief whether Kansas or Nebraska law should apply. Unlike Kansas, Nebraska does not cap non-economic damages. Judge Simmons asked both parties to submit their briefs simultaneously, without response or reply.

Your packet of materials includes:

- *Judge Simmons's Findings of Fact and Conclusions of Law*
- *Defendant's Memorandum for Application of Kansas Law*
- *Plaintiff's Memorandum for Application of Nebraska Law*
- *Full texts of all cases, statutes, and other materials cited in the parties' memoranda*
- *A form for recording your decision and reasoning*

Your task is to decide whether Kansas or Nebraska law applies to the calculation of damages.

[2017 version:]

Please do NOT access any information other than the information in your packet, and please do NOT talk to your neighbors until the study is completed.

You have 50 minutes to reach a decision and submit a brief summary of your reasoning on the paper provided.

[2018 version:]

While working on the task, please do NOT access any information other than the information in your packet, and please do NOT discuss the task or issues it raises with anyone until you have finished.

In your own interest and that of the study, please spend no more than 50 minutes on this task, including time reading the materials and writing the brief summary of the reasons for your decision.

Please write and submit your decision and the brief summary of your reasoning either

online at [treatment specific link]

or

on the paper provided and hand it to me the first morning of the Program (April 23).

Appendix 6: Judge Simmons’s Findings of Fact and Conclusions of Law

[NB: Words in square brackets indicate words, phrases, or sentences that differ by treatment condition, as explained in section 2.3. Some formatting and spacing has been removed for brevity.]

**IN THE UNITED STATES DISTRICT COURT
FOR THE DISTRICT OF [FORUM STATE]**

BEATRICE PARKER,)
)
 Plaintiff,)
)
 vs.) CIV 14-4149-[3-letter treatment code]
)
 GARY ROGERS,)
)
 Defendant)

FINDINGS OF FACT AND CONCLUSIONS OF LAW

After a non-jury trial in the above captioned matter, and review of the pleadings filed by the parties, the Court makes the following Findings of Fact and Conclusions of Law.

FINDINGS OF FACT

1. Plaintiff Beatrice Parker (“Parker”) is a citizen of [Domicile state]. She has lived in [City in domicile state, Domicile state] all her life, works there as a [Parker’s job], and plans to reside there indefinitely.
2. Defendant Gary Rogers (“Rogers”) is a citizen of [Forum state]. Before the accident, he lived all his life in [City in domicile state, Domicile state], but, after the accident, he moved to [City in forum state, Forum state, reason Rogers moved to Forum state]. He plans to reside in [Forum state] indefinitely.
3. The amount in controversy between Parker and Rogers exceeds \$75,000.
4. Parker and Rogers grew up in [City in domicile state, Domicile state]. They met in high school. [Description of post-high school lives]
5. In April 2014, Rogers’s grandfather Edwin Rogers died. He left Rogers [description of inheritance] in [Accident state]. Rogers wanted to inspect the property but [reason he didn’t have a car]. Parker offered to drive him in her [Parker’s vehicle].
6. Parker and Rogers started to drive to [Accident state] in the early afternoon of May 16, 2014. They planned to reach their destination by 9 p.m. Near [City in accident state], however, the engine of the [Parker’s vehicle] overheated. Parker and Rogers had to interrupt their drive and leave the [Parker’s vehicle] in a garage in [City in accident state].
7. Parker and Rogers attempted to rent a replacement vehicle from Avis Car Rental in downtown [City in accident state]. However, [Paker/Rogers] got into an argument with the manager, Arjun Gupta, about the price, which [Parker/Rogers] found abusively high. When [Parker/Rogers] used a racial slur, Gupta refused to rent to Parker and Rogers. As it was then already 6 p.m., other car rentals in downtown [City in accident state] were closed, and Parker and Rogers had to go to the airport to find an open car rental. Rogers ultimately rented a [rental car type] from Hertz Rent a Car. As part of the rental, he purchased the Liability Insurance Supplement (LIS), which provides liability coverage up to \$1,000,000.
8. Parker and Rogers finally continued their drive at around 8 p.m. Rogers’s inheritance was located in [Small city in accident state], a small town about four hours away from [City in accident state]. Rogers drove the rented [rental car type]. Parker sat in the front passenger seat.
9. Around 11:30 p.m., Parker noticed that Rogers was tired and suggested that he pull over and take a nap. Rogers declined. After driving a little longer, he fell asleep. The car veered off the

road and hit a telephone pole while traveling approximately sixty miles per hour. The car struck the pole near the right front corner of the car such that the biggest impact was on the passenger side.

10. Rogers suffered minor injuries.

11. Parker suffered major injuries to her head, back, and right leg, including a shattered jaw and fracture-dislocated vertebrae that compressed her spinal cord.

12. Parker was trapped in the car for several hours. Emergency workers did not remove her immediately, because they thought moving her would aggravate her spinal injury. During this period, Parker endured enormous pain in her jaw, and her lower body was numb from the spinal injury.

13. Parker was hospitalized for ten weeks at the University of [Accident state] Medical Center. The injuries to her jaw, spine, and leg required multiple surgeries, including reconstructive surgery for her shattered jaw. She subsequently required extensive physical therapy to regain partial use of her leg. Her medical expenses from the surgeries, hospitalization, and subsequent treatment, including physical therapy, were \$624,257.34. The treatments were successful in that she regained sufficient use of her legs to perform ordinary tasks of daily living and to pursue her chosen career as a [Parker's job].

14. Parker's injuries made it impossible for her to work for five months, until October 22, 2014. Her [salary or average commission] is \$3000 per month. So the injury caused her to lose five months of income or \$15,000.

15. In the weeks after the accident, Parker suffered from considerable pain in her back, legs and face. The face and leg pain has been cured and the back pain has lessened, but her doctors predict that she will continue to endure mild back pain for the rest of her life.

16. As a result of the accident, Parker has a large scar across her right cheek, in spite of extensive reconstructive surgery. Her doctors predict that the scar will remain visible for the rest of her life.

17. Although the surgeries and physical therapy allowed her to regain sufficient use of her right leg to work and to perform ordinary tasks of daily living, Parker was not able to regain full motion in her right leg, and doctors predict she never will. As a result, Parker walks with a limp, and she is unable to perform strenuous activities or play most sports. Prior to the accident, she went on eight-mile runs several nights a week and spent virtually all her vacations hiking. She is no longer able to hike or run.

Conclusions of Law

1. The court has jurisdiction over the case pursuant to 28 U.S.C. § 1332.

2. Rogers's conduct, falling asleep while driving, constitutes negligence under both Nebraska and Kansas law.

3. Damages depend on whether Nebraska or Kansas law applies. The plaintiff and defendant shall each, therefore, submit to the court, with a copy to the other party, no later than 21 days from the date of this ruling, a memorandum of points and authorities on whether Nebraska or Kansas law applies. Neither party shall file a response, opposition, or reply memorandum unless the court so orders.

4. **If Nebraska law applies to damages**, Rogers is liable to Parker for the full amount of both the economic and non-economic damages. Economic damages are \$639,257.34, which is \$624,257.34 in medical costs plus \$15,000 in lost income. The Court finds the following amounts in non-economic damages to be fair and just considering the nature and extent of Parker's injuries: \$300,000 for past pain and suffering, \$200,000 for future pain and suffering, \$50,000 for scarring and disfigurement, and \$200,000 for loss of enjoyment of life. Non-economic damages, therefore, total \$750,000. Therefore, total damages are \$1,389,257.34.

5. **If Kansas law applies to damages**, Rogers is liable to Parker for the entirety of the economic damages, but non-economic damages are capped at \$250,000. Kan. Stat. Ann. § 60-19a02 (2014). As noted in the preceding paragraph, economic damages are \$639,257.34 and uncapped non-economic damages would be \$750,000. With Kansas's cap on non-economic damages, total damages are, therefore, \$889,257.34.

rd

DATED this 3 day of March, 2018.

BY THE COURT

A handwritten signature in black ink, appearing to read "Frederick W. Simmons", written over a horizontal line.

FREDERICK W.
United State District Judge

SIMMONS

Appendix 7: Sample Legal Briefs

Jeffrey G. Klosterman
P. O. Box 10701
Casper, WY 82602
(307) 265-0732
ATTORNEY FOR PLAINTIFF

**IN THE UNITED STATES DISTRICT COURT
FOR THE DISTRICT OF SOUTH DAKOTA**

BEATRICE PARKER,)
)
 Plaintiff,)
)
vs.) CIV 14-4149-SKS
)
GARY ROGERS,)
)
 Defendant)

**PLAINTIFF’S MEMORANDUM OF POINTS AND AUTHORITIES IN SUPPORT OF
APPLICATION OF NEBRASKA LAW**

Under *Klaxon Co. v. Stentor Elec. Mfg. Co.*, 313 U.S. 487 (1941), this Court must apply South Dakota’s choice of law rules. The South Dakota Supreme Court has made clear on numerous occasions that South Dakota employs the *Restatement (Second)* approach to choice-of-law issues. *See Andrews v. Ridco, Inc.*, 863 N.W.2d 540, 554 (S.D. 2015); *Chambers v. Dakotah Charter*, 488 N.W.2d 63, 67 (S.D. 1992). In tort cases, the *Restatement (Second) of Conflict of Laws* § 145(1) states that the law of the state with “the most significant relationship to the occurrence and the parties” applies. “Contacts to be taken into account ... include (a) the place where the injury occurred, (b) the place where conduct causing the injury occurred, (c) the domicile [and] residence ... of the parties, and (d) the place where the relationship, if any, between the parties is centered.” *Id* at § 145(2). These provisions govern “the measure of damages,” *id* at § 171, including “what limitations, if any, are imposed upon the amount of recovery.” *Id* at § 171 cmt. a.

Contacts (c) and (d) mentioned in *Restatement (Second) of Conflict of Laws* § 145(2) both suggest that Nebraska is the state with “the most significant relationship” to the dispute. As to contact (c), both Beatrice Parker and Gary Rogers resided and were domiciled in Nebraska at the time of the accident, for most of their lives, and immediately after the accident. Beatrice Parker continues to live in Nebraska. Although Gary Rogers moved sometime after the accident, his unilateral, post-accident actions should not influence applicable law. As to contact (d), Nebraska is clearly where the relationship between the parties was centered. Nebraska is where Beatrice Parker

and Gary Rogers met, where they planned their trip to Kansas, and where they returned after the accident.

Contacts (a) and (b) mentioned in *Restatement (Second) Conflict of Laws* § 145(2) both point to Kansas, because that is where the accident took place. Nevertheless, the *Restatement (Second)* makes clear that in “common domicile” cases like this one, the state where both parties were domiciled and where their relationship was centered is the state with the most significant relationships:

When the plaintiff and defendant are domiciled in the same state, and particularly if in addition there is a special relationship between them which is centered in this state, it would seem that this state is likely to be the state with the most significant relationship with respect to the issue of damages.

Restatement (Second) Conflict of Laws § 171 cmt. b.

The Supreme Court of South Dakota has decided two common domicile tort cases since adopting the *Restatement (Second)* in 1992. In both cases, the Court decided that the state where plaintiff and most defendants resided was the state with the most significant relationship and applied the law of that state, even though it was not the state where the injury took place.

Chambers v. Dakotah Charter, 488 N.W.2d 63, 67 (S.D. 1992) is the case in which the South Dakota Supreme Court adopted the *Restatement (Second)*'s most significant relationship approach to choice of law. The suit involved a South Dakota plaintiff who sued a South Dakota corporation for injuries the plaintiff sustained in Missouri while falling from the steps of defendant's bus. Although Missouri was the place of injury, the Supreme Court of South Dakota applied South Dakota's law of comparative negligence. The court emphasized that “South Dakota was the domicile, residence, place of incorporation, and place of business of the parties, as well as the place where the relationship of the parties was centered.” *Id.* at 68.

In *Selle v. Pierce*, 494 N.W.2d 634 (S.D. 1993), a Nebraska resident filed a defamation action against another Nebraska resident regarding a defamatory letter mailed to South Dakota. Even though the plaintiff's business and social reputation in South Dakota was injured by the defamation, the Court applied Nebraska law, which barred punitive damages, because the parties resided in Nebraska and the relationship between plaintiff and defendant was “centered” in Nebraska. *Id.* at 637.

South Dakota's application of the law of the place of common domicile is in accord with the decisions of nearly all states that have adopted the *Restatement (Second)* and other modern choice-of-law methods. So, for example, Symeon Symeonides, a leading choice-of-law scholar, concluded:

When both the tortfeasor and the injured party are domiciled in the same state, judicial opinions converge on the proposition that the state of common domicile has a better claim to apply its law... All together, a total of 50 common-domicile cases have reached 34 state supreme courts with, and since, the abandonment of the *lex loci delicti* rule. Forty-four of these cases (or 88%) applied the law of the common domicile, and six cases (12%) did not.... Of the latter cases, two were factually exceptional, one was overruled, and the remaining three are probably discredited.

Symeon Symeonides, *The American Choice of Law Revolution: Past, Present and Future* 155 (2006).

Symeonides does mention one case that is superficially similar to this case and which applied the law of the place of accident. *Peters v. Peters*, 634 P.2d 586 (Haw. 1981); Symeonides, *supra* at 155. In that case, a husband and wife, who resided in New York, rented a car in Hawaii and got into an accident there. The wife sued the husband, and the court applied Hawaii law, which barred suit between spouses. The court applied Hawaii law because the defendant had purchased insurance from the rental car agency, and the court thought the rental car agency would have calculated premiums based on the assumption that Hawaii law would apply. The court therefore concluded that it would be improper for the rental car agency to have to pay a claim that would have been barred by Hawaii law. *Peters* does not provide justification to deviate from the law of the place of common domicile for several reasons. First, *Peters* was decided by the Hawaii Supreme Court, and so, unlike *Chambers* and *Selle*, the South Dakota cases discussed above, it is not binding authority. Second, unlike South Dakota, Hawaii does not apply the *Restatement (Second)*'s "most significant relationship test," but instead performs "an assessment of the interests and policy factors involved." *Peters*, 634 P.2d at 593. Third, when a car is rented in Hawaii, it is absolutely clear that the car will only be driven in Hawaii, because Hawaii is an island. It thus makes sense that both the rental car agency and renter would expect Hawaii law to apply. In contrast, when, as in this case, a car is rented in Wichita, which is less than 60 miles from the border with Oklahoma and easy driving distance to Nebraska and Missouri, neither the rental car agency nor the renter has a reasonable expectation that Kansas law will apply.

Thus, the overwhelming weight of authority, both in South Dakota and elsewhere, supports the application of Nebraska law, because both plaintiff and defendant were domiciled in Nebraska before and after the accident, and because their relationship was centered there.

Respectfully submitted, March 24, 2017.

By: /s/ Jeffrey G. Klosterman

Eric K. Rutledge
P. O. Box 4086
Cheyenne, WY 82003
(307) 632-0545
ATTORNEY FOR DEFENDANT

**IN THE UNITED STATES DISTRICT COURT
FOR THE DISTRICT OF SOUTH DAKOTA**

BEATRICE PARKER,)	
)	
Plaintiff,)	
)	
vs.)	CIV 14-4149-SKS
)	
GARY ROGERS,)	
)	
Defendant)	

**DEFENDANT’S MEMORANDUM OF POINTS AND AUTHORITIES IN SUPPORT OF
APPLICATION OF KANSAS LAW**

Klaxon Co. v. Stentor Elec. Mfg. Co., 313 U.S. 487 (1941) requires a district court to apply the same choice-of-law methods as the state in which the district court is sitting. In resolving choice-of-law issues in tort cases, South Dakota applies the *Restatement’s* “most significant relationship test.” *Burhenn v. Dennis Supply Co.*, 2004 S.D. 91, ¶ 24, 685 N.W.2d 778, 784 (S.D. 2004); *Selle v. Pierce*, 494 N.W.2d 634, 635-36 (S.D. 1993); *Restatement (Second) of Conflict of Laws* § 145(1).

Kansas is the state with the most significant relationship to this dispute, because that is where the accident took place. In addition, Kansas was the place where Gary Rogers rented the car, where Beatrice Parker and Gary Rogers’s journey in the rental car began, where they planned to inspect Rogers’s inheritance, and where their journey in the rental car was anticipated to end.

Under the *Restatement (Second)* there are four contacts that courts analyze: “(a) the place where the injury occurred, (b) the place where conduct causing the injury occurred, (c) the domicile [and] residence ... of the parties, and (d) the place where the relationship, if any, between the parties is centered.” *Restatement (Second) of Conflict of Laws* § 145(2). The first two contacts point unambiguously toward Kansas, because that is where the accident occurred. The third contact, the parties’ residence and domicile, is unhelpful, because it does not point to a single state. While Beatrice Parker resides and is domiciled in Nebraska, Gary Rogers resides and is domiciled in South Dakota.

While it is true that the South Dakota Supreme Court has twice decided cases in favor of the law of the place of common domicile, *Chambers v. Dakotah Charter*, 488 N.W.2d 63 (S.D. 1992); *Selle v. Pierce*, 494 N.W.2d 634 (S.D. 1993), those cases do not support application of Nebraska law in this case for two reasons. First, as noted above, this is not a “common domicile” case, because Beatrice Parker is domiciled in Kansas, and Gary Rogers is domiciled in South Dakota. Second, neither *Chambers* nor *Selle* involved a situation, like that presented here, in which

the injury and injurious conduct were both entirely in a single state that was different from the state where the parties resided or were domiciled. When injury and injurious conduct are in the same state, as in this case, the first two of the *Restatement (Second)* factors point toward the same state, and that state has a powerful interest in applying its own law. Neither of the South Dakota Supreme Court cases was of this kind.

Chambers v. Dakotah Charter, 488 N.W.2d 63 (S.D. 1992), is very different from this case. First, as noted above, “the principal conduct which allegedly caused the injury” (the distribution of candy on the bus) and the injury (slipping on candy and falling off a step of the bus) did not happen in the same state. It is unknown where the wrongful conduct occurred, but the court is clear that it did not happen in Missouri, where the injury occurred. *Chambers*, 488 N.W.2d at 68. This bifurcation of wrongful act and injury between two states made it less plausible that either would be the state with “the most significant relationship” and heightened the significance of the common domicile. Second, the Court in *Chambers* noted that the bus involved in the accident was “on an interstate journey from South Dakota to Arkansas. It was merely fortuitous that [the plaintiff] slipped while the bus was passing through Missouri.” *Id.* In contrast, the fateful journey in this case took place entirely in Kansas. Gary Rogers rented the car in Wichita, and Beatrice Parker and Gary Rogers intended to drive from there to a place in Kansas and then back again to Wichita, where they would have to return their rental car.

Selle v. Pierce, 494 N.W.2d 634 (S.D. 1993) also points towards application of Kansas rather than Nebraska law. In *Selle*, the injury was spread over several states. The plaintiff’s reputation was injured in his home state, Nebraska, as well as in South Dakota. In fact, because the defamatory words primarily concerned the plaintiff’s activities in Nebraska, that state was not only the state of common domicile, but also the place where the brunt of the injury occurred. *Id.* at 637. In contrast, in this case, the injury to Beatrice Parker, as well as Gary Rogers’s conduct, both occurred entirely in Kansas.

Cases from outside South Dakota support the application of Kansas law in this case. A Hawaii Supreme Court case, *Peters v. Peters*, 634 P.2d 586 (Haw. 1981), is very similar to this case. In that case, the plaintiff and defendant, Mr. and Mrs. Peters, both resided in New York. They flew to Hawaii, where they rented a car. While Mr. Peters was driving, their car collided with a truck and Mrs. Peters was injured. The issue was whether the court should apply Hawaii law, which barred suits between spouses, or New York law, which did not. The Court held that Hawaii law should apply because the journey that resulted in the accident took place exclusively in Hawaii, even though both Mr. and Mrs. Peters were domiciled in New York and even though their relationship was centered there. The Court emphasized the importance of the purchase of insurance in Hawaii for the rental car. It noted that the “the insurance policies covering them undoubtedly were written with the laws of Hawaii in mind. To have New York law govern a tort action arising from the operation of such a vehicle would, of course, contravene the expectations of both insurer and lessor.” *Id.* at 594. Similarly, in this case, Gary Rogers purchased insurance for the rental car in Kansas, so it was the expectation of both the rental car agency and Gary Rogers that Kansas law would apply.

A leading authority on choice of law, Symeon Symeonides, acknowledges that *Peters*, *supra*, makes a justified departure from the pattern of common domicile cases. Unlike other cases deviating from that pattern, which he describes as “overruled” or “probably discredited,” Symeonides classifies *Peters* as “factually exceptional.” Symeon Symeonides, *The American Choice of Law Revolution: Past, Present and Future* 155 (2006). See also non-critical discussion of *Peters* in Symeonides, *supra* at 145 n. 10. This case, like *Peters*, differs factually from the common

domicile cases decided by the South Dakota Supreme Court in that both injury and injurious act occurred in the same state (Kansas). In addition, as noted above, this case is not really a common domicile case at all, because Gary Rogers is domiciled in South Dakota.

The state with the most significant relationship with the dispute is therefore Kansas, where Gary Rogers's injurious conduct occurred, where Beatrice Parker suffered injury, where Gary Rogers rented the car and purchased insurance, and where their entire journey in the rental car both occurred and was intended to occur. As a result, Kansas law, which caps noneconomic damages, should apply.

Respectfully submitted, March 24, 2017.

By: /s/ Eric K. Rutledge

Appendix 9: Exit Survey (differences between 2017 and 2018 versions in square brackets)

Thank you very much for completing the main study! Please answer the following six short follow-up questions:

1. *What proportion of your colleagues do you think decided the case as you did?*
2. *What type of judge are you: circuit, district, bankruptcy, [2018 only: federal claims,] or magistrate?*
3. *In the cases you have decided as a judge, have you ever had to choose between the law of the place of the accident or of the parties' common domicile, as in Parker v. Rogers?*
4. *Did the case seem realistic to you? If not, why not?*
5. *What do you think this study was about?*
6. *I administered a similar [2017: experiment; 2018: study] at this workshop [2017: two; 2018: one and three] years ago, and talked about it [2017: last year; 2018: two years ago]. Did you take part in [2017: this workshop last year or the year before; 2018: those workshops], or did you otherwise hear or read about my prior [2017: experiment; 2018: studies]?*

Appendix 8: Pre-Registration



You are logged in as: [hspamann@law.harvard.edu](#) (Log out [logout.php](#))

[HOME](#) ([index.php](#)) [BACK](#) ([see_list.php](#))

[Make Suggestion](#) ([mailto:larry@AsPredicted.org?Subject=I have a suggestion for AsPredicted](mailto:larry@AsPredicted.org?Subject=I+have+a+suggestion+for+AsPredicted)) [Change your email](#) ([update_email1.php](#))

As Predicted:"FJC/Harvard 2017 Experiment" (#3686)

Created: 04/09/2017 10:21 PM (PT)

Author(s)

Holger Spamann (Harvard Law School) - hspamann@law.harvard.edu
Dan Klerman (USC Gould School of Law) - dklerman@law.usc.edu

1) What's the main question being asked or hypothesis being tested in this study?

Judges' choice of law decision will be influenced by (1) the relevant location parameters identified by the applicable conflicts law, and (2) by its impact on the more sympathetic party. These effects will be mitigated by the type of conflicts law: (1) will be (much) stronger under a rule than a standard, to the point of eliminating (2) altogether.

2) Describe the key dependent variable(s) specifying how they will be measured.

The key dependent variable is the law chosen by the judge. However, the way to code the law differs by effect:

- 1) For measuring the impact of location parameters, the coding is 1 if the judge chose the law of the state of the accident and the forum is Wyoming, or if the judge chose the law of the state of common domicile and the forum is South Dakota.
- 2) For measuring the impact of sympathy, the coding is 1 if the judge chose the law of Nebraska and the defendant (Rogers) is sympathetic, or if the judge chose the law of Kansas and the defendant is not sympathetic.

3) How many and which conditions will participants be assigned to?

8 conditions: 2x2x2 factorial.

Factor 1: Forum: Wyoming or South Dakota. Wyoming follows *lex loci delicti* (law of place of accident applies). South Dakota follows 2nd restatement (which points towards the law of common domicile).

Factor 2: Accident state: Kansas or Nebraska. The parties reside in the respective other state.

Factor 3: Sympathies: Either the defendant or the plaintiff is sympathetic, the other is not.

We block the randomization on judge type (circuit, district, bankruptcy, or magistrate), and randomize only 2 judges from each type into the Wyoming level of factor 1 (because we expect the residual variance under that factor to be zero or close to zero).

4) Specify exactly which analyses you will conduct to examine the main question/hypothesis.

- 1) Main effects of factors 2 and 3, separately for both levels of factor 1. We will use the unconditional Boschloo exact test for these tests because the usual randomization based Fisher test does not have enough power in very small samples (here, 8 observations for the Wyoming level of factor 1) due to its discrete nature.
- 2) Differences in effect sizes for factors 2 and 3 between both levels of factor 1. We will use randomization based Fisher tests for these.

5) Any secondary analyses?

Variance: we will test for differences in residual variance under the standard and the rule -- it should be lower under the rule.

Robustness: Our main tests use all observations, but we will also show results without participants who indicated that they had participated in, or heard about, a similar study that one of us conducted in the past (Spamann and Kloehn, *Journal of Legal Studies* 2016), and/or who indicated that they did not think the case seemed realistic.

6) How many observations will be collected or what will determine sample size?

No need to justify decision, but be precise about exactly how the number will be determined.

All judges who agree to participate at a workshop at Harvard Law School on April 10, 2017. If less than 40, we will try to find more at another workshop.

7) Anything else you would like to pre-register?

(e.g., data exclusions, variables collected for exploratory purposes, unusual analyses planned?)

We also ask participants (1) what percentage of other participants they think decided the case they did, and (2) what they think the study was about.

8) Have any data been collected for this study already?

No, no data have been collected for this study yet