What the Court Cannot See

Commitment and Inference in the Civil Law of the Talmud

Torun Dewan* Limor Ressler†

This version: 6 August 2026

Abstract

We study two doctrines of Talmudic civil law as solutions to a commitment-then-inference problem: a court fixes a rule before it can observe the fact the rule turns on, and reads that fact from the conduct the rule shapes. For the goring ox, we show that the doctrine's two liability regimes – partial, capped (*tam*) and full, uncapped (*mu'ad*) – are what a planner minimising expected social cost would choose under Bayesian updating of a hidden type. The *tam*'s equal risk-split is the Borch rule for two equally risk-averse parties; the three-goring threshold calibrates to the prior rarity of danger. For *migo*, the doctrine's three classical bounds – no *migo* against witnesses, against brazen claims, and against extraction – are comparative statics in the claim environment, and the no-extraction default is derived from the structure of the crediting rule rather than assumed. The doctrines differ in one modelling dimension: the ox's liability is a planner's design; *migo*'s burden of proof is legally given. In each case the Talmud's preserved disagreement (*machloket*) acts as the identifying restriction that selects among the mechanisms consistent with the headline rule.

*London School of Economics and Political Science, Department of Government. Comments welcome. Talmudic loci follow the standard Vilna foliation and have been checked against Sefaria; translations are our own unless noted.

†Contributing Lecturer, University of Cambridge, Judge Business School; Mediator and member of the New York and Israel Bars. Formerly Barclays Bank plc, Arnold & Porter LLP.

1 Introduction

What is the optimal liability rule when a court cannot observe whether an animal is dangerous? What should a burden of proof do when a court cannot see whether a debt was paid? These questions share a structure. In each, the court must fix a rule before the fact the rule turns on is in; and the conduct it then observes – a record of past harm, or the claim a litigant chose to press – is itself shaped by that rule. The rule and the inference are jointly determined, and the doctrine that survives is the one that closes this loop. We study both questions through two doctrines of the Talmud's civil law that provide an unusually clean laboratory: the rules are fully specified, their doctrinal rationales are contested in preserved disputes, and those disputes supply the identifying restrictions our analysis needs.

The first doctrine concerns the ox that gores. An ox that has not gored before is *tam*, innocuous in the law's eyes; its owner pays only half the damage, and pays it solely from the body of the animal. Once the ox has gored on three occasions, with warning, it becomes *mu'ad*, forewarned; now its owner pays the full damage and pays from his best property without limit (*Bava Kamma* 1:4, 2:4; 23b–24a). Liability is light before the type is known and heavy after. The escalation tracks what the court has learned.

The second doctrine is *migo* ("credibility by forfeited advantage"). A litigant is believed because a stronger claim was open to him and he declined it. A borrower sued on an unwitnessed loan says "I borrowed and repaid"; he is believed, because he could have said "there was never a loan," been believed for want of witnesses, and owed nothing.[1] Since the better lie was available and unused, the weaker claim is credited. The court reads the truth from the claim not made.

Both doctrines share one structural difficulty: in each, the court must commit to a rule before it knows the fact the rule turns on; yet the conduct it observes is itself shaped by that very rule. The owner guards, or fails to guard, an ox whose danger he is still learning; the litigant chooses his claim knowing how claims are read. The rule and the inference are jointly determined, and the doctrine that survives is the one that closes this loop. Our reading takes each rule as the solution to that fixed point.

These doctrines share a second feature we exploit. The Talmud records not only its rules but its disagreements, and a recorded dispute over a rule's reason can serve as an identifying restriction: the same rule is consistent with two mechanisms, the dispute is over which is at work, and each disputant's auxiliary ruling is the prediction that tells them apart. The argument is economical: two mechanisms agree on the headline rule and part on some further ruling, and where the tradition records which further rulings each side accepts, those side-rulings are the over-identifying restrictions that separate the mechanisms. We state the move here and put it to work twice.[2] The half-damage of

[1] The principle of *migo* is developed in the third chapter of *Bava Batra* (*chezkat ha-batim*, 31a–33b). The loan illustration is a systematisation by the Rishonim (the medieval Talmudic authorities, c. 1000–1500 CE) rather than a verbatim case there: it rests on the rule that one who lends without witnesses need not be repaid before witnesses, so an oral borrower is believed when he says he repaid (*Shevuot* 41b; *Ketubot* 88a; cf. *Bava Batra* 5b–6a).

[2] The method is developed in general in Torun Dewan and Limor Ressler, "Reading the Economic Logic

the *tam* is the subject of a famous *machloket*: is it a fine (*knas*), a penalty imposed on an owner the law presumes blameless, or compensation (*mamon*), owed in principle and softened out of leniency (*Bava Kamma* 15a–b)? The reach of *migo* is likewise disputed: does it merely deflect a charge, or does it positively prove a claim and so allow its holder to extract (Tosafot against Ramban and Rashba on *ein omrim migo l'hotzi*, *Bava Batra* 32b)? These disputes are not noise; in each, the disagreement will identify which mechanism our model is running.[3]

One idea runs through the paper. A court that cannot see the relevant truth commits to a rule before the fact is in and then reads the fact from the conduct the rule shapes; the rule and the inference are fixed together. This is the structure the two doctrines share, unrelated though they are in subject, one a matter of torts and one of evidence. What separates them is the source of the commitment: for the goring ox the rule is a design, the liability a planner facing the inference would choose; for *migo* it is given, a burden of proof the law lays down and the court does not select. We offer this as an organising structure, not a shared theorem. Each doctrine is formalised in its own terms: the ox as a type learned over time, *migo* as a fact inferred at once. The common frame is the lens that makes their shared problem and the point where they part legible; it is not a single model of which the two are cases. What distinguishes these two doctrines from any rule that uses conduct to infer a fact is that the court's instrument is coarse: it cannot price the unobservable directly, but conditions the rule on a public record – a count of gorings or a menu of available claims. The frame identifies a class of doctrines where this coarseness is constitutive, not incidental. A second observation accompanies it, and it is methodological. Where the law records a dispute over a rule's reason, that dispute identifies the mechanism the rule encodes, and we use it so. The first idea is the paper's subject; the second is a tool it affords. The structure recurs because the problem recurs.

We make three sorts of claims. First, for the goring ox, we show that escalating liability – light while the type is unknown, heavy once it is learned – is what a planner would choose, the whole menu of fractions and caps collapsing to the doctrine's two regimes. The leniencies of the *tam* regime, the half and the cap, both lapse at the *mu'ad* threshold, which the court's learning triggers – but for different reasons: the cap because deterring a now-foreseeable danger requires full, uncapped internalisation, the half because the equal risk-split it encodes is superseded once liability passes to the owner in full. What is not keyed to the posterior is the half's *magnitude*: why one-half is a risk-sharing matter, settled before any learning, whereas the cap's content answers directly to what the court has learned.

---

of Law from Its Recorded Disputes: The *Machloket* as Identification" (draft, 2026); here we apply it to two problems of inference.

[3]That a recorded dispute must part on some *further* ruling to carry this information is the tradition's own demand, not one we impose: the Gemara typically treats two views as genuinely distinct only where a case divides them, asking *mai nafka minah* – what practical difference emerges – and *mai beinaihu* – what is between them; it does not always set aside every disagreement that lacks a practical ruling, but in the cases we use here each *machloket* carries a live discriminating ruling – the admission rule for the *knas*/*mamon* dispute, the extraction question for *migo* – that is native to the *sugya* (the Talmudic discussion in which the dispute is embedded), not imposed from outside.

Second, for *migo*, we show that crediting the claim a liar would have abandoned is an equilibrium criterion and that the doctrine's three classical bounds – against witnesses, against extraction, and against brazenness – are comparative statics in the claim environment. We show, too, that the no-extraction default, *ein omrim migo l'hotzi*, is derived rather than assumed. A *migo* rests on a forgone *credited* alternative: the court credits a modest claim because a liar would strictly have preferred the credited bolder one. A challenger seeking to extract must affirmatively prove, so the burden of proof refuses every unproven claim of his; with no credited claim on his menu, he has no credited alternative to forgo, and the *migo* criterion cannot credit him. The default follows from the structure of the crediting rule itself, not from an independent assumption about *migo*'s scope.

Third, in each case we read the recorded dispute as an identifying restriction that fixes which mechanism the rule encodes, and we are candid about what the model favours. For the *tam*'s half our account comes down on *mamon*, and two arguments support the reading. The first is about magnitude: the fraction is close to the risk-sharing share two faultless parties would choose, lifted only slightly, by a pull of order μ, toward deterrence – an invariant round half being just what a focal equal-division would fix. This is suggestive but not conclusive, since a *knas* that happens to equal the risk-sharing fraction is consistent with both views. The second argument is structural: the auxiliary ruling on admission, *modeh bi-knas patur*, discriminates between the two readings independently of the fraction's size, and it is this over-identifying restriction that identifies the mechanism. The *knas* reading is the disfavoured but doctrinally preserved alternative, and the *machloket*, through that auxiliary ruling, is what would tell the two apart.

The contribution is one of importation: the models are standard; the source and the use we make of its disagreements are not.

We are not the first to find economics in this law. Aumann and Maschler (1985) showed that the Mishnah's estate divisions encode the nucleolus, and a tradition of comparative and positive analysis runs from Finkelstein (1981) on the goring ox to Levmore (1986) on ancient tort. That work explains why durable rules persist. We do something adjacent and distinct: we treat each doctrine as a rule shaped by the inference it must support, and we use the law's own disputes to identify the mechanism behind it.

The paper proceeds as follows. Section 2 sets out the common structure: a rule committed to before the fact, inference from the conduct it shapes, and the *machloket* as identification. Section 3 treats the goring ox. Section 4 treats *migo*. Section 5 locates the argument in the economics of tort and of evidence, chiefly Shavell (1986) on limited liability, Holmström (1979) on hidden-action moral hazard, and Shin (1994) on the burden of proof, noting in each case where the Talmudic doctrine departs from those benchmarks. Section 6 concludes.

2 Commitment, then inference

We begin by describing the structure the two doctrines share. In each there is a court that must act under a rule fixed before the fact the rule turns on is known. There is a party whose conduct the court observes and a fact about that party the court does not observe.

The unobservable is a type or a state: whether the ox is dangerous, whether the debt was paid. The court forms a belief about the unobservable from the conduct it sees; but the conduct is a response to the rule, so the rule must be right about the very behaviour it induces. The rule and the inference are fixed together, and the doctrine that survives is the one that closes this loop.

The doctrines differ in where the rule comes from. In the goring ox it is chosen – the liability a planner facing the inference would set; in *migo* it is given – a burden of proof the law imposes and the court applies. The difference matters for how we read each: the ox invites the question what rule a designer would choose, *migo* the question what a fixed rule lets the court infer. This distinction is one of modelling strategy, not a claim about the jurisprudential origins of either rule: both are products of rabbinic reasoning, and the Talmud does not itself draw the line this way. The choice of which question to ask – what rule a planner would set, or what a given rule implies – follows from the structure of the doctrine, not from a prior about how the law was made. Two further features make the problem non-trivial in both. Conduct is endogenous – the owner's care and the litigant's claim respond to the rule, so the court cannot read conduct naively – and the court's instrument is coarse: it cannot price the unobservable directly, but conditions the rule on a public record, the count of gorings or the menu of available claims. The doctrines are interesting precisely because the law does well with coarse instruments.

The second feature is methodological. A single rule is typically consistent with more than one account of why it exists: half-damage as a penalty on a blameless owner or as softened compensation, a credited claim as a deflected charge or as a proven right. These accounts agree on the headline rule and disagree elsewhere. Where the law records the disagreement and records which auxiliary rulings each side accepts, the dispute supplies the over-identifying restrictions that separate the accounts. We apply it twice. In the goring ox the dispute is over a number, the half, and the auxiliary ruling is whether the payment can be imposed on a litigant's own admission. In *migo* the dispute is over reach, whether the inference can extract, and the auxiliary ruling is which party may invoke it. In each the side-rulings are exactly the margins on which the two mechanisms part company. We now show this in detail.

3 The ox that learned

We begin our formal analysis by describing a simple model of liability for an animal of unknown disposition. An owner holds an ox of hidden type $\theta \in \{S, D\}$, safe or dangerous. The court and the owner share a prior $\mu_0 = \Pr(\theta = D)$, and μ_0 is small: goring is abnormal behaviour (*meshuneh*), so most oxen are safe. The owner chooses care $e \geq 0$ at convex cost $\psi(e)$ and decides whether to keep the ox, worth V in production, or to cull it for salvage $L < V$. The owner and the prospective victim are risk-averse, each valuing a random payment at its mean plus a premium in its variance; in the paradigm case of one ox goring another they are peer owners, whom we take to be equally risk-tolerant.

An ox of type θ gores with probability $p_\theta(e)$, decreasing in care, with $p_D(e) > p_S(e)$ for every e: the dangerous ox is likelier to gore at any level of guarding. A goring inflicts

harm H on a victim. We assume the type is learned publicly: a goring enters a record that owner and court read alike. This matches the doctrine, in which *mu'ad* status is established by testimony before the court, and it keeps the owner from holding private knowledge the court lacks.[4]

The law sets liability per goring as a function of the public record, and it does two things at once. It *deters*: the larger the share of harm an owner bears, the more care he takes and the readier he is to cull. It also *allocates risk*: a larger share loads more of the harm's variance onto a risk-averse owner. The law has two margins. The first is the *fraction* $f \in [0, 1]$ of the harm H the owner pays. The second is the *cap*: whether payment is drawn only from the body of the ox, bounded by its value Q (the *tam*'s *mi-gufo*), or from the estate without limit (the *mu'ad*'s *min ha-aliyah*). A planner chooses the rule to minimise the sum of expected harm, the cost of care, the productive value lost to inefficient culling, and the risk-bearing premium the rule imposes. The doctrine offers two regimes – *tam*, a half drawn from the body, and *mu'ad*, the full sum from the estate – and assigns one by the record. We show over the course of this section that the planner's choice reduces to these two, rather than ranging over the whole menu of fractions and caps: a kept ox is never optimally placed under full liability (Lemma 1), its efficient share is the interior risk-sharing one (Proposition 4), and inducing a cull requires the full, uncapped rule (Lemma 2). The two doctrinal regimes are the endpoints these results pick out; we assemble the reduction as we go rather than assert it here.

3.1 Why liability escalates

Consider whether liability should depend on the record. Two observations drive the answer. First, care cannot target a danger no one yet perceives: while the ox is probably safe, guarding does little, and liability buys almost no deterrence. Second, full liability is not free; it loads the whole variance of a rare, large harm onto a risk-averse owner. So while the ox is probably safe, the deterrence value of liability is small and its risk cost is real, and the law does better to share the harm than to impose it whole. Once the record marks the ox as dangerous the balance turns. Care and culling now matter, and culling in particular requires that the owner bear the harm he could avoid; the deterrence stakes become first-order and override the case for sharing risk.

Lemma 1 (Full liability is never optimal for a kept ox)**.** *Fix the action at* keep*. At every posterior μ the social-cost-minimising share is interior, $f^*(\mu) < 1$, so full liability is strictly dominated; the optimal share rises with μ but never reaches one.*

Proposition 1 (The foreseeability switch)**.** *There is a threshold belief $\mu^* \in (0, 1)$ such that the cost-minimising rule keeps the ox under partial, capped liability ($f^*(\mu) < 1$) below μ^* and culls it under full, uncapped liability above. By Lemma 1 full liability on a kept ox is*

[4]The symmetric-learning assumption is substantive. If the owner learned his ox's danger privately, before any goring, the problem would become one of signalling, and the optimal rule would screen rather than classify. The Talmud's insistence on three public gorings, witnessed "before the owner and the court" (*Bava Kamma* 24a), is what licenses the symmetric reading.

never optimal, so the optimum's only escalation is from partial-keep to full-cull. Liability escalates with the court's knowledge of the type.

Escalation is not a crude proxy for fault. While the ox is probably safe, full liability would impose the harm's risk for a deterrence gain that care cannot deliver, so partial, risk-sharing liability does better; reserving full liability for beliefs above the threshold improves on imposing it throughout, since below the threshold it would buy deterrence that care cannot deliver. Once the ox is probably dangerous the law must induce culling, and culling requires full internalisation. The *mu'ad* threshold is the belief at which the danger is worth acting on.

Write K_{over} for the net social loss from culling a safe ox: the value of keeping it in service, above its salvage. Write K_{under} for the net social loss from keeping a dangerous ox rather than culling it: its expected harm and the care it warrants, net of the value keeping would add over salvage. These are social stakes, not the owner's private externality; the appendix fixes them as $W_S - L$ and $L - W_D$. Culling becomes efficient once the expected loss from keeping, μK_{under}, exceeds that from culling, $(1-\mu)K_{\text{over}}$, which gives the switch belief

$$\mu^* \;=\; \frac{K_{\text{over}}}{K_{\text{under}} + K_{\text{over}}}. \tag{1}$$

When safe oxen are valuable μ^* rises and the law demands more evidence; when dangerous oxen are very harmful μ^* falls and it switches early.

3.2 The three gorings

The doctrine fixes the evidentiary bar at three. We can say what that bar is calibrated to. Throughout this classification we hold care fixed, so a goring is a signal of type alone; the likelihood ratio λ is then a constant, not the care-dependent object of the deterrence analysis. Each goring multiplies the odds on danger by $\lambda = p_D/p_S > 1$. After k gorings the posterior odds are $\frac{\mu_0}{1-\mu_0}\lambda^k$,[5] and *mu'ad* status should trigger when these cross $\mu^*/(1-\mu^*)$.[6] The required number of gorings is therefore

$$k^* \;=\; \left\lceil \frac{\log\left[\dfrac{\mu^*}{1-\mu^*}\cdot\dfrac{1-\mu_0}{\mu_0}\right]}{\log\lambda} \right\rceil. \tag{2}$$

[5] We condition on the count of gorings, which is what the doctrine itself counts. Strictly the count is a sufficient statistic for the type only at a fixed observation horizon: if gorings arrive as a Poisson process over exposure time t, the posterior odds are $\frac{\mu_0}{1-\mu_0}\lambda^k e^{-(p_D-p_S)t}$, so the sufficient statistic is the count *together with* the elapsed time. The quiet spells between gorings enter through the exposure term $e^{-(p_D-p_S)t}$, which pushes the posterior toward safety and drives the reversion of a quiescent ox to *tam*; the requirement of three gorings on three separate days chimes with it. Holding exposure fixed, the count alone gives the odds used above, and the comparative statics are unaffected.

[6] One illustrative calibration returns three. A prior $\mu_0 = 0.05$ that the ox is dangerous, a likelihood ratio $\lambda = 3$ per goring, and equal classification stakes ($K_{\text{over}} = K_{\text{under}}$, so $\mu^* = \frac{1}{2}$) give $k^* = \lceil \log 19/\log 3 \rceil = \lceil 2.68 \rceil = 3$. We do not claim these are the true primitives; the point is that three is what a rare, informative danger returns, not that the model forces exactly three.

Proposition 2 (The classification threshold)**.** *The* mu'ad *bar* k^* *in* (2) *is increasing in the rarity of danger (as* μ_0 *falls,* k^* *rises) and decreasing in the informativeness of a goring (as* λ *rises,* k^* *falls). For an illustrative rare and informative danger, three gorings is the bar at which the posterior odds first cross* $\mu^*/(1-\mu^*)$*.*

Three is not a magic number; it is what (2) returns for a rare and informative danger. Rarer danger demands more signals; a more telling goring demands fewer. The reversion of a quiescent ox to *tam* status is the same calculation run backwards, as the absence of fresh gorings pulls the posterior back below μ^*.

3.3 Why the cap lifts with the fraction

The *tam* pays from the body of the ox; the *mu'ad* pays without limit. We argue that the cap is a culling instrument and that it must lift exactly when the type becomes known. A cap at Q bounds the owner's loss at the value of the animal. For a known-dangerous ox, a capped owner bears only $\min\{fH, Q\}$ of each goring; when the cap binds his private cost of keeping is muted, so he externalises the rest and keeps a ruinous animal – the judgement-proof problem of Shavell (1986), with the ox's body as the limit of liability. Uncapping restores internalisation: the owner then bears the full harm, his private keep-or-cull problem becomes the planner's, and he culls precisely when a dangerous ox is worth culling – when its salvage exceeds the value keeping would add net of expected harm and care, that is when $W_D < L$.

Proposition 3 (Uncapping turns culling on)**.** *Under the* mi-gufo *cap, an owner keeps a dangerous ox whenever the cap binds and his muted private cost of keeping leaves it worth keeping, even where culling is efficient. Removing the cap at* mu'ad *makes his problem the planner's, so he culls exactly when culling a dangerous ox is efficient,* $W_D < L$*.*

Both leniencies of the *tam* regime lapse at the same threshold, but not for the same reason. The cap lapses because internalisation, and with it the threat of culling, now requires it; the partial fraction lapses because full liability supersedes any split of the harm. Only the first is keyed to what the court has learned. The cap, moreover, does no work as a cap while the ox is *tam*, since a safe ox rarely gores and rarely exhausts Q; its bite is felt only once the ox is known dangerous, which is exactly when the law removes it.

3.4 The half and the *machloket* that identifies the regime

We come to the number itself. Why a half? On the present account the *tam*'s fraction is the share that allocates the harm's risk between the two parties. The paradigm case is one ox goring another: two peer owners, neither at fault, facing a loss neither could foresee. The efficient division of a risk loads each party in inverse proportion to his risk aversion – the Borch rule – and for two equally risk-averse peers that division is an equal split. The half is that split.

Proposition 4 (The half is the risk-sharing share)**.** *As* $\mu \to 0$*, where care has no product, the optimal owner's share is the risk-sharing (Borch) share* $f^* = \rho_v/(\rho_o + \rho_v)$*, which equals*

one-half when the two parties are equally risk-averse. Near that limit the share is raised above f_B only by a deterrence pull that vanishes with μ; the departure grows with μ but stays small through the tam *region when care is weakly productive relative to the risk borne. The share is interior for any positive risk aversions and departs from one-half only as those aversions differ.*

So the half is not an unexplained number but the equal division of a shared risk, exact under the symmetry the paradigm case supplies – two peer owners. On this account the half is *mamon*, compensation for a loss shared between faultless parties, and not *knas*, a penalty: the deterrence the fraction provides is of order μ and fades as the ox is presumed safe, so a penalty is not what the model delivers. This magnitude argument is suggestive but not alone sufficient, since the numerical coincidence is consistent with either reading; the identification rests on the structural argument that follows. Asymmetric risk tolerance moves the share off one-half, so we read the constant *chatzi nezek* as the focal equal-sharing rule. The model explains that an equal division is chosen, not its fine magnitude. The same half governs *tam* cases that lack two faultless peers to share between – an ox goring another's chattel, say – which a focal equal-division reading explains but a tailored risk-share would not.

This is where the recorded dispute identifies the reading. The Talmud asks whether the *tam*'s half-payment is *knas* or *mamon* (*Bava Kamma* 15a–b): a penalty imposed to deter an owner the law presumes blameless, or compensation for a loss shared between two faultless parties. These are the two roles the fraction plays in the model – a deterrence margin and a risk-sharing margin. They agree on the number and disagree on its function. They also disagree on an auxiliary ruling: a *knas* cannot be imposed on a defendant's own admission (*modeh bi-knas patur*), whereas *mamon* can. That ruling does not bear on the half; it bears on whether the half is a penalty or a compensation and so on which work the *tam*'s fraction does.

We now ask what each reading predicts about that ruling, rather than take the ruling as a label. Read as compensation, the half undoes a loss the two parties share. A defendant's admission confirms that the loss occurred; it can only strengthen the case for the transfer, so compensation is collected on admission as it is against a stranger. Read as a penalty, the half is a wedge above compensation, justified by the care it induces before the fact. Here self-reporting bites. An owner deciding whether to come forward weighs the sanction he would bear on admission against the expected sanction if he stays silent and is later caught. Set the admission sanction just below that expected detected sanction and he discloses; because the deterrent works through the expected sanction he faces *ex ante*, which the concession leaves in place, care is undiminished. So the penal part is relaxed on admission – set below the sanction on detection – while the compensatory part is not. This is the self-reporting logic of Kaplow and Shavell (1994). The two readings move in opposite directions on the same ruling, and *modeh bi-knas patur* records which way the doctrine moves. The ruling is thus the over-identifying restriction – a prediction the model makes, not a fact it borrows.

Two caveats bound the claim. The model pins the direction of the ruling, not its size. It says a penalty is treated more leniently on admission than a compensation; it does not

say the penalty is extinguished outright. The doctrine's total exemption is the corner of that prediction, not an optimum we claim – a partial relief would carry the same sign. The distinction has force only where the court may fail to see the deed, the imperfect detection on which this whole reading turns. Where every goring is witnessed there is no admission to reward, and the ruling has nothing to discriminate.

Proposition 5 (The *machloket* corresponds to the role of the half)**.** *The* knas/mamon *dispute corresponds to the two roles the fraction plays in the model:* knas *to the deterrence margin, on which the half is a penalty that induces care, and* mamon *to the risk-sharing margin, on which the half is the efficient division of a shared loss. The model predicts that the two respond oppositely to a defendant's own admission: the compensation is collected, the penalty is relaxed. The auxiliary ruling* modeh bi-knas patur *records that opposition and so over-identifies the regime rather than merely labelling it. The prediction fixes the direction of the ruling, not its magnitude. Here the model supplies the prediction the dispute confirms.*

The goring ox, read this way, is a law in transition. The *tam* regime treats a goring as an accident between blameless parties and shares the loss; the *mu'ad* regime treats it as the foreseeable harm of a known danger and places it on the owner. The doctrine moves from sharing to internalisation as a hidden type is learned. Our model comes down on the *mamon* side of the *tam*'s half – the half as shared compensation – but the *machloket* preserves the *knas* reading and, in *modeh bi-knas patur*, the auxiliary ruling that would tell the two apart.

4 *Migo*: credibility from the option set

We now turn from torts to evidence and from a type learned over time to a fact inferred at once. The structure, however, is the same. A court that cannot see what passed between two litigants reads the truth from a litigant's choice among the claims he might have made. We begin by describing the menu of claims and the chance that each survives scrutiny.

A claimant privately knows the facts. He is one of two types: an honest type, for whom a particular modest claim c_0 is true, and a lying type, for whom it is false; the court's prior that he is honest is q. He chooses a claim c from a menu ordered by *boldness*, where a bolder claim concedes less and keeps more. Claim c carries a benefit if credited, $b(c)$, increasing in boldness, and a survival probability $s(c)$, the chance the claim is not refuted by evidence that may come to light. Survival is truth-dependent. A true claim survives for certain, $s(c) = 1$. A false claim survives with probability $\sigma(c) \in [0, 1)$, set by how much contestable ground the claim must hold and by what evidence stands against it. Refutation is one thing; crediting is another. The court is bound by a burden-of-proof rule, not free to act on a hunch: confronting an unrefuted claim it credits a defending possessor and refuses a challenger, who must affirmatively prove. Write $\chi(c) \in \{0, 1\}$ for whether the rule credits an unrefuted c. The rule is *consistent* when each crediting decision is correct given how claimants choose against it, and the question *migo* answers is what the consistent rule does with a modest claim, one that concedes part.

A remark on the sign of σ forestalls a natural error. It is tempting to think that a bolder claim is always likelier to be refuted, so that σ falls in boldness. This, however, is not our interpretation. A clean denial that concedes nothing may be the least contestable claim of all: the borrower who says "there was never a loan," with no witnesses against him, leaves nothing to be picked apart. What lowers a false claim's survival is the contestable ground it must defend, not the boldness of what it asserts. In the canonical case the bold denial concedes no contestable fact, so σ is higher for the bold claim than for the modest one, and this is exactly what makes the inference run.

4.1 *Migo* as a crediting rule

A lying claimant chooses the claim that maximises his expected benefit, $\chi(c)\, b(c)\, \sigma(c)$, over the claims the rule credits ($\chi(c) = 1$) and evidence leaves standing ($\sigma(c) > 0$). Let c^* denote his best lie. Now consider a claimant who has made some modest claim c_0. If a liar would strictly have preferred an available credited claim c^*, then no liar would have stopped at c_0; a claimant who did stop there is telling the truth.

Proposition 6 (The *migo* criterion)**.** *The consistent burden-of-proof rule credits the modest claim c_0 if and only if a lying claimant would strictly have preferred an available, credited, bolder claim c^*, that is if $b(c^*)\,\sigma(c^*) > b(c_0)\,\sigma(c_0)$. The honest type then separates onto c_0, provided he does not himself reach for the bold lie, $b(c_0) \geq b(c^*)\,\sigma(c^*)$; the two conditions hold together over a non-empty range of benefits. Where they hold, a claimant who chooses c_0 is truthful, and crediting him is correct.*

The criterion is the doctrine's "since he could have claimed c^*," stated as a sorting condition. The bold claim is the suspect one, the claim a liar would have made; *migo* transfers its standing to the modest claim a liar would have abandoned. The court does not credit c_0 because it believes it; it is committed to credit an unrefuted possessor, and *migo* extends that rule to the modest claim exactly when a liar would have gone bold.[7]

Crediting is thus a certification, not a wager on the odds. The rule credits c_0 only when doing so is correct with certainty – when the modest claim fully separates the honest type, so no lying type chooses it. That is why the prior q does not enter the criterion. It fixes only how doubtful the modest claim would be with no *migo* to vouch for it; where separation fails and the claim pools, as under refuting witnesses below, crediting is simply withdrawn, not traded off against q.

4.2 The three bounds as comparative statics

The power of the reading is that the doctrine's classical bounds, which the tradition states as separate rules, fall out as comparative statics in the evidence environment. We take them in turn.

[7]The two conditions read $b(c_0)\sigma(c_0) < b(c^*)\sigma(c^*) \leq b(c_0)$, a non-empty interval since a false modest claim has $\sigma(c_0) < 1$. They bind together only when the bold alternative keeps strictly much more than the modest one; in the canonical equal-benefit cases the honest condition is slack.

No *migo* against witnesses. The first bound holds that there is no *migo* where witnesses refute a litigant's claim (*migo b'makom edim lo amrinan*). The doctrine's usual statement bars *migo* where witnesses contradict the claim the litigant *made*; that case is a direct evidentiary override – the witnesses refute c_0 itself, and crediting simply fails – distinct from and simpler than the channel we model. We model the cognate case in which the witnesses bear on the bolder claim he *forwent*, which is what the revealed-preference logic turns on; the two coincide when the witnessed fact is what separates the claims. In the model, witnesses are standing evidence against the bold denial; they drive its survival $\sigma(c^*)$ toward zero. When $\sigma(c^*)$ is small enough, the liar no longer prefers c^*; he too would have stayed modest, the modest claim pools, and its credit disappears.

No *migo* off a brazen claim. The second bound holds that the forgone claim must be one a liar would actually have made; a claim too brazen to utter, even when true, certifies nothing (*migo d'he'aza*). The Talmud's own *chazaka* that "a man will not brazenly deny a debt to his creditor's face" (*Bava Metzia* 3a) supplies the friction. We add a brazenness cost $\kappa(c)$, rising in boldness and incurred whether or not the claim is true. The liar now compares $b(c^*)\,\sigma(c^*) - \kappa(c^*)$ with $b(c_0)\,\sigma(c_0) - \kappa(c_0)$.
The cost κ is a private cost of brazen speech, not a truth-dependent penalty on lying: it falls on anyone who utters the marked claim, honest or lying. Its work is not to punish but to make the bold claim informative – uttering it is costly, so a liar with a cheaper credible option declines it, and the modest claim separates. The cost is invoked selectively, for the particular claims the doctrine marks as brazen, just as the *chazaka* is. That selectivity is not arbitrary: the *chazaka* turns on a concrete feature – denial *to the creditor's face* – so κ is high exactly for a claim that brazenly contradicts a party present to gainsay it and negligible otherwise. The model thus does not tax boldness in general, only where the tradition names a brazenness worth pricing; where it does not, a claimant bold and truthful speaks freely.

No *migo* to extract. The third bound holds that *migo* shields a possessor but cannot dispossess another (*ein omrim migo l'hotzi*). It rests on the burden default that "the one who extracts from his fellow bears the proof" (*ha-motzi me-chavero alav ha-re'aya*, *Bava Kamma* 46b). Here the distinction between refutation and crediting earns its keep. A challenger's bold claim may be perfectly unrefuted, with a high σ, and still the burden rule refuses it: $\chi = 0$ across his menu, because he must prove and cannot. He therefore has no *credited* better alternative to forgo, and the *migo* has nothing to transfer. A defending possessor, whose unrefuted claims the rule credits, keeps a credited c^* and with it a live *migo*.

Proposition 7 (The three bounds as comparative statics)**.** *The doctrine's three classical bounds are comparative statics in the claim environment. (i)* Witnesses*: there is a threshold* $\underline{\sigma} > 0$ *such that the criterion credits* c_0 *only if* $\sigma(c^*) > \underline{\sigma}$*; evidence that drives* $\sigma(c^*)$ *below* $\underline{\sigma}$ *extinguishes the* migo*. (ii)* Brazenness*: if* $b(c^*)\,\sigma(c^*) - \kappa(c^*) \leq b(c_0)\,\sigma(c_0) - \kappa(c_0)$*, forgoing* c^* *certifies nothing and the* migo *fails; the bound applies only to claims the named* chazaka *marks as brazen. (iii)* Extraction*: if the burden rule sets* $\chi(c) = 0$ *on every claim of a*

party's menu, he has no credited alternative to forgo and the criterion never credits him, however unrefuted his claims.

4.3 The *machloket* that identifies the inference

The reach of *migo* is itself disputed – the principle *ein omrim migo l'hotzi*, that *migo* will not dispossess (*Bava Batra* 32b) – and the dispute identifies what the inference is. Tosafot hold that *migo* is mere leverage to deflect a charge; it cannot extract. Ramban and Rashba treat *migo* as positive proof of a claim's legitimacy, so that in principle it could extract, leaving the question on their view genuinely open. The two readings agree that *migo* credits the possessor and disagree on whether it can also dispossess. They correspond to two readings of the inference in our model: whether the revealed-preference argument runs only defensively, lowering the bar a possessor must clear, or runs as affirmative evidence that can shift the burden itself.

Here the model does more than house the two readings; it derives the baseline. *Migo* is a revealed-preference device: it credits a modest claim because a bolder *credited* claim was forgone. Its whole force rests on there having been a credited alternative to abandon. A defending possessor has one – the burden credits his unrefuted claims, so a bolder unrefuted claim would have been credited too, and declining it earns the modest claim credit. A challenger has none: by the burden of proof he must affirmatively prove, so no unrefuted claim of his is credited (Proposition 7(iii)), and there is no credited bolder claim for a *migo* to rest on. So *migo* can relay the crediting the burden already grants, but cannot manufacture crediting where the burden grants none. This is *ein omrim migo l'hotzi*, derived rather than imposed: *migo* credits, but does not by itself extract. The *machloket* is then over whether *migo* can ever do more than relay – whether a compelling enough *migo* can itself confer crediting and move the burden – not over whether *migo* is inference at all, which both sides grant.

Proposition 8 (The *machloket* corresponds to the mechanism)**.** *The model derives* ein omrim migo l'hotzi *as the default.* Migo *rests on a forgone* credited *alternative; by the burden of proof a challenger has none (Proposition 7(iii)), so* migo *credits a defending possessor but cannot by itself extract. The Tosafot–Ramban dispute is over the exception. On the Tosafot reading* migo *only relays the crediting the burden already allows: it credits a possessor's modest claim but leaves* $\chi = 0$ *on extraction. On the Ramban and Rashba reading a sufficiently compelling* migo *confers crediting itself, moving* χ *from* 0 *to* 1 *and carrying a party past the burden to extract. The model derives the default and houses both exceptions; which party may invoke* migo *is the auxiliary ruling that records how the doctrine settles the exception.*

As with the goring ox, the doctrine's disagreement is the datum that tells us which mechanism the rule encodes. Tosafot read *migo* as a procedural relaxation of the burden; Ramban reads it as evidence that can meet the burden. The model derives the default and houses both exceptions; the recorded dispute selects between them.

5 Related literature

We locate the argument in the economic literatures it draws on. The reading of the goring ox belongs to the economic analysis of tort. The bilateral-care model is Brown (1973); the comparison of strict liability and negligence, and the distinction between care and activity levels, are Shavell (1980); the limited-liability or judgement-proof problem, which our cap instantiates, is Shavell (1986). The doctrine is a hybrid of the two classical regimes: lighter before *mu'ad*, where the owner bears only a partial, capped liability for a danger he could not foresee, and full strict liability after. The care problem embedded in the *tam* regime is a standard hidden-action moral hazard (Holmström 1979): the owner's incentive to guard responds to the fraction of harm he bears. What is new is the learning. The same owner moves between regimes as evidence about a fixed hidden type accrues, and the *mu'ad* bar is the posterior at which the move is efficient. The common-law analogue, the dangerous-propensity or one-bite rule of liability for animals (Landes and Posner 1987), shares this structure and has, to our knowledge, not been modelled as a posterior crossing a classification threshold.

A second connection runs to the economics of strike rules in the cabinet. In Dewan and Myatt (2007) a minister faces a fixed two-strike rule – dismissal follows a second scandal – and the leader chooses how hard to come down on the first: the hazard rate of dismissal after one scandal. That rate is a pure incentive instrument; there are no types and nothing is learned. Raising it cuts two ways. It sharpens effort directly, a substitution effect, but it also makes a ministerial career more precarious and so worth less, an income effect. Our goring ox is the mirror image: a strike rule in which the count carries *learning* about a fixed hidden type, with care held fixed in the classification (Section 3.2). Joining the two is the natural next step: a threshold at which the accumulating count both updates the type and sets the incentive, the substitution and income effects playing out against the Bayesian bar. Such a model would bridge these Talmudic rulings to the government of cabinets; we do not attempt it here.

Our reading of the goring ox is nearest, in its subject, to Levmore (1986). He surveys the ox of Exodus and of the Mesopotamian codes and asks why a system splits the loss between two owners when neither ox was at fault. His answer is uncertain causation: when no ox has a goring history, the law cannot see which animal began the fight, "or whether the survivor presents a future danger." That last clause is the seed of our hidden type. Levmore names, too, the cap that holds the owner's payment to the value of the surviving ox and likens it to the Roman noxal surrender. Yet his aim is comparative, not optimal: "I do not argue that any one of these many rules is superior or inferior," he writes; systems converge where a rule bears on behaviour and diverge where it does not, and the variety is on his account harmless. We depart here. We take the survivor's future danger as a type the court learns through harm, and we let the law's own disagreement, the *knas/mamon machloket*, identify the regime the rule belongs to. By contrast with Levmore, who reads the law's variety as behaviourally idle, we read its disagreements as data.

The reading of *migo* belongs to the economics of evidence and persuasion. Its nearest antecedent is Shin (1994), who studies an arbitrator who infers from the evidence a litigant did not present. We share the inference from a forgone stronger option. We differ in three

ways. First, our verification is soft: a false claim survives with a probability set by its contestable ground, rather than being either hard or absent. Second, the single-crossing that drives separation is derived from that truth-dependent survival, not assumed as a cost. Third, we recover a body of named doctrinal bounds as comparative statics in the evidence environment. The broader setting is the literature on verifiable disclosure and unravelling (Grossman 1981; Milgrom 1981; Milgrom and Roberts 1986) and on evidence games with state-dependent message sets (Glazer and Rubinstein 2004; Bull and Watson 2007; Hart, Kremer and Perry 2017). Our court's commitment, however, is exogenous – the burden of proof is a legal primitive, not a mechanism the court chooses – which separates the setting from the optimal-commitment results of that last literature. The contrast is sharpest with Bayesian persuasion (Kamenica and Gentzkow 2011), where a sender commits to a signal she designs to move a receiver's belief. Our court commits too, but to a burden the law lays down; the commitment is legal, not chosen. *Migo* is partial unravelling halted by refutation risk: not the full cascade to the maximal claim, but a separating step in which the bold claim's exposure keeps the honest modest claim credible. Because the survival probability is truth-dependent, the model is one of communication with state-dependent verifiability rather than cheap talk (Crawford and Sobel 1982); the nearest cost-based model is Kartik (2009). Our *exogenous* burden of proof also connects the model to the economics of the standard of proof (Demougin and Fluet 2006) and of adversarial evidence provision (Daughety and Reinganum 2000); we hold the standard fixed and study the inference a litigant's choice of claim permits.

The methodological move has antecedents too. Aumann and Maschler (1985) read a sugya as encoding a solution concept, and Aumann (2003) finds the economics of risk-bearing in the sages' own reasoning; a positive tradition reads ancient law as economically intelligible – though Finkelstein (1981), often enlisted here, in fact reads the goring ox more as a quasi-person of ritual law than in economic terms; and an informal law-and-economics literature treats the burden of proof and *migo* directly (Sinai 2008). A separate tradition in the economics of religion treats religious legal rules as institutional mechanisms solving community commitment and collective-action problems; our focus is on the inference structure of individual doctrinal rules rather than on community-level institutional design, and our tools are the models of tort and evidence rather than club-good theory. Our use of the law differs in one respect. We do not only ask what economic logic a rule serves; we use the law's recorded disputes, the *machloket*, as the identifying restrictions that select among the logics a rule admits. The disagreements the tradition preserves are, on this reading, data.

6 Conclusion

We have read two doctrines of Talmudic civil law as solutions to a court's problem of inference. The goring ox makes liability escalate as a hidden type is learned through harm and at the *mu'ad* threshold lapses both leniencies – the cap because deterring a now-foreseeable danger requires full internalisation, the half because the equal risk-split it encodes is then superseded. *Migo* credits the claim a liar would have abandoned, and

its classical bounds – against witnesses, against brazenness, and against extraction – are comparative statics in the claim environment: in the evidence that can refute a bold claim, the speech cost of a brazen one, and the burden of proof that fixes what crediting can achieve. The doctrines share no subject matter. They share a structure: a court commits to a rule before it can see the relevant truth, then infers that truth from the conduct the rule shapes – by design in the ox, where the rule is the liability a planner would set, and by law in *migo*, where the burden of proof is given.

The commitment structure suggests a design principle. Where a court cannot observe the relevant fact, the optimal public-record rule is coarse – a threshold at the posterior μ^*, not a continuous tariff – because the informational constraint that motivates coarseness (the court conditions only on a public count or a menu of claims) is precisely the constraint that makes the two-regime solution efficient. The *migo* criterion points in the same direction: it credits when separation is certain, so the prior probability of honesty, q, drops out of the equilibrium outcome. Neither doctrine requires the court to calibrate a continuous schedule to unobservable primitives; both achieve first-best inference from a coarse instrument.

Running alongside this is a method. In each doctrine the law preserves a dispute over the rule's reason, and the dispute is an identifying restriction. The *knas*/*mamon machloket* tells us which baseline the half departs from; the Tosafot–Ramban dispute tells us whether *migo* relaxes a burden or meets it. The tradition records not only its rules but its disagreements, and the disagreements pin the mechanisms the rules encode. We have tried to take the disputes seriously as evidence rather than to resolve them.

The reading has limits. Our account of the half comes down on *mamon*, compensation for a shared loss; it rationalises an interior, near-equal share but does not uniquely pin one-half, and the *knas* reading survives as the doctrinally preserved alternative. The learning in the goring ox is public, and a model of private warning would look different. *Migo*'s bounds are recovered under a particular account of survival, and other accounts are possible. The unity we claim is of structure, not of a shared theorem: the two models are built in their own terms, and the common frame is what makes their problem, and its two solutions, one subject. In sum we hope that the paper offers a small step toward reading the civil law of the Talmud as a body of designed responses to inference, with its own disputes as the data that identify them.

A Proofs

Throughout the appendix we work with a tractable specialisation of the model in the text and state each simplifying assumption as we use it. Owner and victim are risk-averse, with mean-variance preferences; the planner minimises expected social cost inclusive of the risk premium each party bears. The analysis is of a single ox, and of a single claimant, static given the posterior the public record induces.

A.1 The goring ox

Fix the per-period primitives. An ox of type $\theta \in \{S, D\}$ gores with probability $p_\theta(e)$ under care e at convex cost $\psi(e)$, with $p_D(e) > p_S(e)$ and $|p'_D(e)| > |p'_S(e)|$, the latter vanishing for a safe ox: a dangerous ox is both likelier to gore and more responsive to care, while a safe ox barely gores however it is guarded. A goring inflicts harm H, so the loss is random with variance Σ at the prevailing goring probability. A kept ox of type θ yields net social value $W_\theta \coloneqq \max_{e\geq 0}[\, V - \psi(e) - p_\theta(e)H \,]$, attained at e^*_θ; culled, it yields salvage L. Since $p_D > p_S$, $W_D < W_S$. We study the decision-relevant case

$$W_S > L > W_D, \tag{A1}$$

so a safe ox is worth keeping and a dangerous one worth culling, and write $K_{\text{over}} \coloneqq W_S - L > 0$ and $K_{\text{under}} \coloneqq L - W_D > 0$ for the costs of culling a safe ox and of keeping a dangerous one.

Owner and victim have mean-variance preferences with risk-aversion coefficients $\rho_o, \rho_v > 0$: a party bearing a random loss X incurs $\mathbb{E}[X] + \frac{\rho}{2}\mathrm{Var}[X]$. Under a rule giving the owner a fraction f of each harm, the owner sets care to minimise his private cost; his marginal return to care rises in f, so his care $e(f;\mu)$ increases in f and is first-best at $f = 1$.

Proof of Proposition 4. Hold the action at *keep* and care at the owner's response. The rule's risk-bearing cost is $R(f) = \frac{\rho_o}{2} f^2\Sigma + \frac{\rho_v}{2}(1-f)^2\Sigma$, convex in f and minimised at the Borch share $f_B = \rho_v/(\rho_o + \rho_v)$, which is $\frac{1}{2}$ iff $\rho_o = \rho_v$. The deterrence cost $G(f;\mu) = \mathbb{E}_\mu[p(e(f;\mu))]H + \psi(e(f;\mu))$ is minimised at $f = 1$ and, as $\mu \to 0$, is flat in f, since then care has no product ($|p'_S| \to 0$). Hence as $\mu \to 0$ the optimal share solves $R'(f) = 0$, giving $f^* = f_B$; it is interior for any $\rho_o, \rho_v > 0$ and equals one-half exactly when $\rho_o = \rho_v$. Away from the limit, the optimal share solves $R'(f) + G_f(f;\mu) = 0$. By the owner's care first-order condition, $\psi'(e) = -f\,\mathbb{E}_\mu[p'(e)]H$, so the deterrence gradient reduces to $G_f = (1-f)\,\mathbb{E}_\mu[p'(e)]H\, e_f$. Only the dangerous type's goring probability responds to care, so $\mathbb{E}_\mu[p'(e)] = O(\mu)$ as $\mu \to 0$; the care response $e_f = -\mathbb{E}_\mu[p'(e)]H/\psi''(e)$ is then $O(\mu)$ too, so $G_f(f_B;\mu) = O(\mu^2)$ to leading order (and $O(\mu)$ as a loose bound). Since $R'' = (\rho_o + \rho_v)\Sigma$ is bounded away from zero, the shift near the limit is $f^* - f_B = -G_f/R'' = O(\mu^2)$. The departure grows with μ as $\mathbb{E}_\mu[p']$ does; through the *tam* region it is bounded by $(1-f)\,\mathbb{E}_\mu[p'(e)]H\, e_f/[(\rho_o + \rho_v)\Sigma]$ and stays small whenever care is weakly productive relative to the risk borne – when the care-responsive harm $\mathbb{E}_\mu[p'(e)]H$ is small beside the risk curvature $(\rho_o + \rho_v)\Sigma$. Under that condition the optimal share stays near the risk-sharing share across the region below μ^*. □

Proof of Lemma 1. Hold the action at *keep*, and take the harm's variance Σ as fixed by its size (constant in care).[8] Bearing fraction f, the owner chooses care to minimise $\psi(e) + f\,\mathbb{E}_\mu[p(e)]H + \frac{\rho_o}{2} f^2\Sigma$, with first-order condition $\psi'(e) + f\,\mathbb{E}_\mu[p'(e)]H = 0$; this defines

[8] If instead Σ fell in care, the owner at $f = 1$ would internalise the variance as well as the expected harm. By the envelope theorem the induced-care effect on total social cost still vanishes at $f = 1$: the harm-plus-care term's rise is offset by the risk-bearing term's fall, so $C'(1;\mu) = \rho_o\Sigma > 0$ continues to hold and the interior conclusion is unchanged – though not, as one might think, strengthened. We keep Σ constant for simplicity.

$e(f;\mu)$, increasing in f. Social cost given keeping is $C(f;\mu) = G(f;\mu) + R(f)$, where $G(f;\mu) = \mathbb{E}_\mu[p(e(f;\mu))]H + \psi(e(f;\mu))$ and $R(f) = \frac{\rho_o}{2} f^2\Sigma + \frac{\rho_v}{2}(1-f)^2\Sigma$. Differentiating G and substituting the owner's first-order condition,

$$G'(f;\mu) = \left[\mathbb{E}_\mu[p'(e)]H + \psi'(e)\right] e'(f) = (1-f)\,\mathbb{E}_\mu[p'(e)]\,H\,e'(f),$$

which vanishes at $f = 1$: under full liability the owner already internalises the harm, so a marginal cut in his share forgoes no deterrence to first order. But $R'(1) = \rho_o\Sigma > 0$, so $C'(1;\mu) = \rho_o\Sigma > 0$ for every μ; lowering f below one strictly reduces social cost, and the optimum is interior, $f^*(\mu) < 1$. A higher μ puts more weight on the care-responsive dangerous type, making $\mathbb{E}_\mu[p'(e)]$ more negative, so $C_{f\mu} = G_{f\mu} < 0$ and $f^*(\mu)$ rises with μ; yet $C'(1;\mu) = \rho_o\Sigma > 0$ throughout, so it never reaches one. □

Lemma 2 (Judgement-proof keeping)**.** *Let $\Psi_D^{\text{cap}} = \min_{e\geq 0}[\,\psi(e) + p_D(e)\min\{fH, Q\}\,]$ be a dangerous ox's owner's least private cost of keeping under a capped partial-liability rule. When the cap binds and $V - \Psi_D^{\text{cap}} > L$ the owner keeps the ox; under full uncapped liability he culls it, since by (A1) $W_D < L$.*

Proof. Facing the capped rule, the owner chooses care to minimise $\psi(e)+p_D(e)\min\{fH, Q\}$, attaining Ψ_D^{cap} at some $\hat{e}_D$; he keeps when the ox's value net of this cost exceeds salvage, $V - \Psi_D^{\text{cap}} > L$. A binding cap, $Q < fH$, mutes both his liability and the care it would buy, so Ψ_D^{cap} is small and the inequality holds for a productive ox: he keeps a dangerous animal whose harm he does not fully bear. (A risk premium on the capped stake adds a bounded amount to Ψ_D^{cap} and does not overturn this while the cap is tight.) Under full uncapped liability his private cost is the social one, so his value from keeping is W_D; by (A1) $W_D < L$, and he culls. □

The cap, not the fraction, governs culling: the half can sit alongside it without inducing a cull, and only removing the cap turns culling on. This is Proposition 3.

Proof of Proposition 1. By Lemma 1 a kept ox optimally bears an interior, capped share, and by Lemma 2 inducing a cull requires full uncapped liability (a cap is moot once the ox is culled, since a culled ox does no harm); so at posterior μ the planner's choice is between keep under partial capped liability (*tam*) and cull under full uncapped liability (*mu'ad*). At the first-best keeping benchmark, keeping is efficient when its expected net loss falls below culling's, $\mu K_{\text{under}} < (1-\mu)K_{\text{over}}$, that is for $\mu < \mu^*$ with $\mu^* = K_{\text{over}}/(K_{\text{under}} + K_{\text{over}})$. The *tam* regime does not quite attain that benchmark: partial capped liability leaves a risk premium of order $\rho_o\Sigma$ and a residual care wedge, both of which lower the value of keeping. Taken small relative to the culling stakes $K_{\text{over}}, K_{\text{under}}$, they shift the exact threshold by a like-bounded amount and leave its comparative statics – rising in the value of a safe ox, falling in the harm of a dangerous one – intact; we read μ^* as this benchmark switch. Below μ^* the optimum keeps, at the interior share of Lemma 1 (near the Borch share of Proposition 4); above μ^* it culls, which by Lemma 2 takes full uncapped liability, the risk-sharing motive of order $\rho_o\Sigma$ being dominated near the switch by the first-order culling stake $\mu K_{\text{under}} - (1-\mu)K_{\text{over}}$. Hence liability escalates from partial-keep to full-cull at μ^*, and the escalation is over the full instrument set: keeping never optimally uses full liability, and only culling does. □

Proof of Proposition 2. A single goring has likelihood $p_D(e)$ under D and $p_S(e)$ under S, so it multiplies the posterior odds on D by $\lambda = p_D/p_S > 1$. After k independent gorings the odds are $\frac{\mu_0}{1-\mu_0}\lambda^k$. The *mu'ad* switch of Proposition 1 requires these odds to reach $\mu^*/(1-\mu^*)$, that is $\frac{\mu_0}{1-\mu_0}\lambda^k \geq \frac{\mu^*}{1-\mu^*}$. Taking logarithms and solving for the least integer k gives k^* as in (2). Since $\log\lambda > 0$, k^* is decreasing in λ; and since the numerator $\log[\frac{\mu^*}{1-\mu^*} \cdot \frac{1-\mu_0}{\mu_0}]$ is decreasing in μ_0, k^* is increasing as μ_0 falls. Reversion is the same threshold crossed from above as the absence of fresh gorings lowers the odds. □

Proof of Proposition 5. The two roles are the two margins through which the share enters social cost in the proof of Proposition 4: the risk-sharing cost $R(f)$, on which the share is the efficient division of a shared loss (*mamon*), and the deterrence cost $G(f;\mu)$, on which it is a wedge that induces care (*knas*). Call the associated payments the compensatory transfer and the penal wedge, to distinguish them from these cost terms. We show the two payments respond with opposite sign to a defendant's own admission, in an imperfect-detection setting where the deed is witnessed with probability $p < 1$.

Take the compensatory transfer first. It undoes a loss the two faultless parties share and is owed on the state in which the loss occurred. An admission confirms that state, so it cannot lower the risk-sharing case for the transfer: compensation is collected on admission as it is on detection. Now the penal wedge. Its whole justification is the care it induces before the fact, and care responds to the *expected* sanction the owner faces *ex ante*. Let s_c be the sanction on conviction (probability p) and s_a the sanction on self-report. An owner discloses only if s_a is no greater than the sanction he expects from staying silent; setting s_a just below that expected detected sanction secures disclosure. Because deterrence is carried by the ex-ante expected sanction, and the report is made after the fact, lowering s_a in this way leaves the deterrent in place while gaining the disclosure the planner values: raising s_a toward s_c would only deter reporting without sharpening care. So the optimal penal sanction on self-report satisfies $s_a < s_c$: the penal wedge is relaxed on admission, the compensatory transfer is not.

The two payments therefore move in opposite directions on the same ruling – the compensation collected, the penalty relaxed. *Modeh bi-knas patur* records that opposition and so over-identifies the regime. Two limits bound the claim. The argument signs the response, not its size: it delivers $s_a < s_c$, of which the doctrine's total exemption ($s_a = 0$) is the corner, not an optimum we claim – a partial relief carries the same sign. It has force only where detection is imperfect, $p < 1$; where every goring is witnessed there is no admission to condition on, and the ruling has nothing to discriminate. □

A.2 *Migo*

We cast the doctrine as a verification problem with a committed court. The claimant is honest with probability q and lying with $1-q$; the honest type's modest claim c_0 is true, the lying type's is false. A claim c delivers benefit $b(c)$ if credited, b increasing in boldness, and is refuted by evidence with probability $1-s(c)$, where a true claim has $s(c) = 1$ and a false claim $s(c) = \sigma(c) \in [0,1)$. The court is committed to a crediting rule $\chi : c \mapsto \{0,1\}$ that embodies the burden of proof: it credits an unrefuted claim by a possessor and refuses

an unproven claim by a challenger. Refutation σ and crediting χ are distinct – evidence and decision rule. A claimant choosing c obtains $\chi(c)\, b(c)$ if true, since it survives, and $\chi(c)\, \sigma(c)\, b(c)$ if false. The burden of proof is exogenous: it commands $\chi(c) = 1$ on an unrefuted claim by a possessor and $\chi(c) = 0$ on a challenger's, so the only latitude is over a modest claim that concedes part. The rule is *consistent* when it credits that modest claim only if doing so is correct given the claimant's best response; we characterise the unique consistent rule. Since the burden fixes χ on every claim but c_0, and a liar's best credited option over the whole menu is c^*, restricting attention to $\{c_0, c^*\}$ is without loss.

Proof of Proposition 6. By the exogenous burden, $\chi(c^*) = 1$ for the believable bold claim $c^* = \arg\max_c b(c)\, \sigma(c)$ a possessor may make; a rule crediting nothing would violate the burden, so $\chi(c^*) = 1$ is forced and the only open choice is $\chi(c_0)$. Against the rule a lying claimant's best option is c^*, worth $b(c^*)\sigma(c^*)$. Consider crediting the modest claim, $\chi(c_0) = 1$. The lying type would abandon c_0 for c^* exactly when $b(c^*)\sigma(c^*) > b(c_0)\sigma(c_0)$, his payoff from the false modest claim. When this holds, only the honest type chooses c_0, so crediting it is correct and the rule is consistent. When it fails, the lying type is content with c_0; crediting c_0 would credit a liar, so the consistent rule sets $\chi(c_0) = 0$ and there is no *migo*. The honest type indeed plays c_0 rather than his own best lie c^* provided $b(c_0) \geq b(c^*)\sigma(c^*)$; with the *migo* condition this gives $b(c_0)\sigma(c_0) < b(c^*)\sigma(c^*) \leq b(c_0)$, a non-empty range since $\sigma(c_0) < 1$. Any false claim c pays $\chi(c)b(c)\sigma(c)$, maximised at c^* – for the lying type on every claim, and for the honest type on every claim but his true c_0 – so the two comparisons cover all deviations. The court never credits on the strength of belief; it is bound to credit an unrefuted possessor, and *migo* is the unique consistent extension of that rule to the modest claim. □

Proof of Proposition 7(i). Let $\underline{\sigma} = b(c_0)\sigma(c_0)/b(c^*)$. By Proposition 6 the modest claim is credited if and only if $b(c^*)\sigma(c^*) > b(c_0)\sigma(c_0)$, that is if and only if $\sigma(c^*) > \underline{\sigma}$. Witnesses to the transaction are standing evidence against the bolder claim and lower its survival $\sigma(c^*)$. Once they drive $\sigma(c^*)$ below $\underline{\sigma}$, the inequality fails, the lying type no longer prefers c^*, and the *migo* is extinguished. The bound is the comparative static in $\sigma(c^*)$. □

Proof of Proposition 7(ii). Add a brazenness cost $\kappa(c) \geq 0$, increasing in boldness and borne whether or not the claim is true. The lying type now compares $b(c^*)\sigma(c^*) - \kappa(c^*)$ with $b(c_0)\sigma(c_0) - \kappa(c_0)$. If the former is no greater, the separating condition of Proposition 6 fails with the cost included, the liar does not reach for c^*, and forgoing it certifies nothing. Since κ is incurred by the truthful type as well, it does not by itself impeach a true bold claim; it blocks the inference only for claims to which the named *chazaka* assigns a high enough κ, which is the doctrine's restriction of the bound to brazen alternatives. □

Proof of Proposition 7(iii). For a challenger seeking to extract, the burden rule sets $\chi(c) = 0$ on every claim of his menu: an unrefuted claim is not credited, since he must affirmatively prove and cannot. His best credited alternative then yields $\chi(c^*)b(c^*)\sigma(c^*) = 0$, so the *migo* condition of Proposition 6 cannot hold for any modest claim, however high its σ. A defending possessor, whose unrefuted claims the rule credits, retains a credited c^* and a live *migo*. The bound acts on crediting χ, not on refutation σ; conflating the two would

make it and the Tosafot–Ramban dispute (Proposition 8) incoherent, since that dispute is precisely over whether *migo* can move χ. □

Proof of Proposition 8. Migo operates only on a claimant who forwent a *credited* alternative. By Proposition 6 the modest claim c_0 is credited when a lying claimant would have strictly preferred an available bolder claim c^*; that comparison has force only if c^* would itself have been credited, $\chi(c^*) = 1$, so that it was worth making. For a defending possessor the burden sets $\chi = 1$ on unrefuted claims, so such a c^* exists and the sorting of Proposition 6 can credit c_0. For a challenger the burden sets $\chi(c) = 0$ on every claim of his menu (Proposition 7(iii)): no unrefuted claim of his is credited, so there is no credited c^* to forgo, and the *migo* condition fails for want of a credited alternative – not for want of a strong one. Hence *migo* credits defensively but cannot by itself extract. This is *ein omrim migo l'hotzi*, derived from the structure of the crediting rule rather than assumed.

The exception is whether *migo* can confer crediting rather than merely relay it – whether the revealed-preference standing of a modest claim can set $\chi = 1$ where the burden leaves $\chi = 0$. On the Tosafot reading it cannot: χ is fixed by the burden and *migo* works only within it, so $\chi = 0$ on extraction and Proposition 7(iii) holds with full force. On the Ramban and Rashba reading a sufficiently compelling *migo* raises χ from 0 to 1, carrying a challenger past the burden. The model derives the default and houses both exceptions; the recorded ruling on which party may invoke *migo* selects between them. □

References

[1] Aumann, Robert J., and Michael Maschler. 1985. "Game Theoretic Analysis of a Bankruptcy Problem from the Talmud." *Journal of Economic Theory* 36(2): 195–213.

[2] Aumann, Robert J. 2003. "Risk Aversion in the Talmud." *Economic Theory* 21(2–3): 233–239.

[3] Brown, John P. 1973. "Toward an Economic Theory of Liability." *Journal of Legal Studies* 2(2): 323–349.

[4] Bull, Jesse, and Joel Watson. 2007. "Hard Evidence and Mechanism Design." *Games and Economic Behavior* 58(1): 75–93.

[5] Crawford, Vincent P., and Joel Sobel. 1982. "Strategic Information Transmission." *Econometrica* 50(6): 1431–1451.

[6] Daughety, Andrew F., and Jennifer F. Reinganum. 2000. "On the Economics of Trials: Adversarial Process, Evidence, and Equilibrium Bias." *Journal of Law, Economics, and Organization* 16(2): 365–394.

[7] Demougin, Dominique, and Claude Fluet. 2006. "Preponderance of Evidence." *European Economic Review* 50(4): 963–976.

[8] Dewan, Torun, and David P. Myatt. 2007. "Scandal, Protection, and Recovery in the Cabinet." *American Political Science Review* 101(1): 63–77.

[9] Holmström, Bengt. 1979. "Moral Hazard and Observability." *Bell Journal of Economics* 10(1): 74–91.

[10] Finkelstein, Jacob J. 1981. "The Ox That Gored." *Transactions of the American Philosophical Society* 71(2): 1–89.

[11] Glazer, Jacob, and Ariel Rubinstein. 2004. "On Optimal Rules of Persuasion." *Econometrica* 72(6): 1715–1736.

[12] Grossman, Sanford J. 1981. "The Informational Role of Warranties and Private Disclosure about Product Quality." *Journal of Law and Economics* 24(3): 461–483.

[13] Hart, Sergiu, Ilan Kremer, and Motty Perry. 2017. "Evidence Games: Truth and Commitment." *American Economic Review* 107(3): 690–713.

[14] Kaplow, Louis, and Steven Shavell. 1994. "Optimal Law Enforcement with Self-Reporting of Behavior." *Journal of Political Economy* 102(3): 583–606.

[15] Kamenica, Emir, and Matthew Gentzkow. 2011. "Bayesian Persuasion." *American Economic Review* 101(6): 2590–2615.

[16] Kartik, Navin. 2009. "Strategic Communication with Lying Costs." *Review of Economic Studies* 76(4): 1359–1395.

[17] Landes, William M., and Richard A. Posner. 1987. *The Economic Structure of Tort Law*. Cambridge, MA: Harvard University Press.

[18] Levmore, Saul. 1986. "Rethinking Comparative Law: Variety and Uniformity in Ancient and Modern Tort Law." *Tulane Law Review* 61(2): 235–287.

[19] Milgrom, Paul R. 1981. "Good News and Bad News: Representation Theorems and Applications." *Bell Journal of Economics* 12(2): 380–391.

[20] Milgrom, Paul, and John Roberts. 1986. "Relying on the Information of Interested Parties." *RAND Journal of Economics* 17(1): 18–32.

[21] Shavell, Steven. 1980. "Strict Liability versus Negligence." *Journal of Legal Studies* 9(1): 1–25.

[22] Shavell, Steven. 1986. "The Judgment Proof Problem." *International Review of Law and Economics* 6(1): 45–58.

[23] Shin, Hyun Song. 1994. "The Burden of Proof in a Game of Persuasion." *Journal of Economic Theory* 64(1): 253–264.

[24] Sinai, Yuval. 2008. "The Doctrine of Affirmative Defense in Civil Cases: Between Common Law and Jewish Law." *North Carolina Journal of International Law and Commercial Regulation* 34(1): 111–164.