Reasonableness*

Scott Baker
Washington University in St. Louis
bakerscott@wustl.edu

Giri Parameswaran
Haverford College
gparames@haverford.edu

May 28, 2023

Abstract

This paper investigates what makes behavior reasonable. Two actors exert effort towards a goal. The planner knows each actor's cost of effort. The actors know their own cost, but not their counter-party's. We find that the planner will not base incentives on the actors' cost of care (information that is free and accurate). Instead, the planner identifies a common standard of "reasonableness" for many agents to follow to foster coordination and avoid waste. Meanwhile, the planner forgives the least able and holds them to a lower standard customized to their costs, while never upping the standard for the most able.

Key Words: Reasonable Person, Asymmetric Information, Coordination, Task-specialization.

JEL Codes: D7, K1, K4, L2

*We thank Gary Biglaiser, Dan Epps, Claude Fluet, Simon Grant, Louis Kaplow, Pauline Kim, Lewis Kornhauser, Claudio Mezzetti, Rohan Pitchford, Steven Shavell, and Kathy Spier for comments. We also thank workshop participants at the following institutions: Australian National University; Columbia Law School; Harvard Law, Economics and Organization Workshop; Haverford College; Macquarie University; Melbourne University; Monash University; Queensland University; Sydney University; Toulouse School of Economics; Washington University in St. Louis; and participants at the 2022 Society for the Advancement of Economics Theory annual meeting, and the 2022 Law and Economic Theory Conference.

In many everyday interactions—driving, dating, working, selling—"reasonableness" defines acceptable behavior. Reasonable behavior is permissible; unreasonable behavior is not. The "reasonable" person standard is used to decide legal cases from torts, contracts, and property, to employment discrimination and ERISA, to required disclosures under securities laws. But what exactly does reasonable even mean and why do we require it? Our model explains why "being reasonable" is such a pervasive cultural and legal idea. It also explains why, sometimes, excusing or forgiving some individuals for "being unreasonable" leads to a superior allocation of resources.

In the model, two agents take efforts that contribute towards a common goal. Effort decisions are complements: the effectiveness of an agent's effort depends on her counter-party's. As a result, a mismatch of effort can generate inefficiencies; circumstances where an actor exerts effort that is wasted due to a lack of effort by her counter-party. Each party knows her own cost of effort, but, importantly, is ignorant of the counter-party's. A planner wishes to provide incentives, constrained by the fact that the actors very know little about each other.

Unlike the classic moral hazard problem, we assume that the planner observes each party's effort. Moreover, the planner knows both actors' costs. The planner determines 'appropriate' (or 'ideal') choices for each agent, mindful of what can be feasibly implemented given the information possessed by the agents. In so doing, the planner must decide whether to "personalize" the scheme—i.e., make it responsive to the individual's cost or talent or "objectify" the scheme—i.e., make it unresponsive to the individual's cost or talent.

Beside the law, the type of problem we study arises in many settings. First, take a manager overseeing a supply chain. The supply chain is non-integrated and has a number of links. Production depends on the effort of each supplier in the chain. The suppliers are located around the globe and know little about each other. Should the manager customize or objectify each supplier's performance standard? What happens to the performance standards if the supply chain is only as strong as its weakest link, as is sometimes suggested (Quade, 2023; Shih, 2023).

Second, take codes of conduct, like the ones often adopted by universities and businesses. In formulating standards of behavior, the question is whether different agents should be permitted to act differently (standards should be personalized and excuses generously given) or should everyone be held to the same standard (be reasonable) irrespective of their cost of compliance.

Although there are many settings where reasonableness governs behavior, our primary motivation roots in law. A recurring character in the common law is the "reasonable" person—a disembodied hypothetical construct intended to establish the proper measure of behavior. If a party's behavior is as good as, or better than, the objective "reasonable person's", then it bears no liability. But if a party fails to live up to what a reasonable person would do, then she is responsible—at least in part—for whatever harm was caused. The reasonable person, in turn, is the 'average person', the figurative "man on the Clapham Omnibus"[1] whose traits are those commonly found in the community. Indeed, as Holmes (1881) remarks: "the standards of law are standards of general application. ... [They require] a certain average of conduct, a sacrifice of individual peculiarities going beyond a certain point."[2]

Reliance on the reasonable person standard entails costs. Under the standard, less able individuals must take precautions whose cost to them is greater than the benefits of those precautions to others, and gifted persons are told the law does not expect them to use precautions that are personally cheap for them, but expensive for the average person. The cost-benefit calculation of the reasonable person is a fiction that rarely matches the cost-benefit calculation of the real person. By relying on objective standards, the law seems to ignore information that would be relevant to any cost-benefit analysis. This is a puzzle.

More puzzling is that, after ignoring costs for most agents, the law then finely tailors the standard of conduct for some individuals according to their cost — and it does so asymmetrically. It lowers the standard and forgives those with a high-cost of compliance (for example, children and individuals with physical disabilities) while seldom heightening the standard for those with a low-cost of compliance.

Well-known examples of the reasonable person standard in the law include:

- Torts: Liability attaches when a defendant fails to use reasonable care to prevent an accident (Dobbs, Hayden, & Bublick, 2015, p.213). Courts define reasonable care as the "care, attention or skill a reasonable person would use under similar circumstances" (MSBNA Standing Committee on Pattern Jury Instructions, 2009).

- Property: An owner of a piece of property abutting a river, stream, or lake is entitled to a "reasonable use [of the surface water], with due regard to the rights and necessities of others interested."[3]

[1] *McQuire v Western Morning News* [1903] 2 K.B. 100 at 109 per Collins MR

[2] Scholars debate whether courts equate reasonable with how the average person would behave or instead see reasonable as a prescriptive judgment about how the typical person should behave (Tobia, 2018). For our purposes, either view will suffice.

[3] *Martin v. Brit. Am. Oil Producing Co.*, 1940 OK 218.

- Contracts: An offer for sale arises when a reasonable person standing in the shoes of the offeree would conclude that his assent, and only his assent, is needed to create an enforceable agreement.

So why use objective standards? For one, the court might find it costly to measure an individual's aptitude. The law and economics literature has explored this reason in detail (Holmes, 1881; Posner, 2014; Shavell, 1987). We place this concern to one side by assuming, as noted above, the planner knows (1) each actor's cost of effort and (2) the effort they undertook. As such, the planner could, if they wanted to, perfectly personalize standards of conduct to an individual's costs/abilities, and customize compensation schemes accordingly.

Our first result shows that a planner will couple objective standards for the most-able agents with forgiveness for the least able agents. In so doing, the planner trades off the benefits of coordination and waste avoidance against the costly failure to customize standards of performance. On the one hand, by holding actors to the same conduct or benchmark, the planner ensures coordination of efforts, which due to complementarity mitigates waste. On the other hand, objective standards sacrifice the benefits of linking the effort required to how much that effort costs the agent.

The planner induces pooling among enough of the more able actors to ensure that gains from coordinated effort are maximized for the "average" ability of the actor in the pool — a definition of reasonableness under the law. The region of pooling always contains the most able actors, the ones with the lowest cost of effort. In this region, the standard is objective, meaning actors are held to the same standard despite heterogeneity in their costs of effort. On the other hand, the region of excuses contains the least able, the ones with the highest cost of effort, and is fully customized to the actor's cost.

To flesh out the intuition, consider the following example: To prevent a traffic accident, both the pedestrian and the motorist should exercise care. Since the motorist and the pedestrian are strangers, they do not know how costly the effort is for the other party. Suppose that, given the ways accidents happen, the party taking the least amount of care determines the probability of an accident. For example, a motorist driving recklessly significantly increases the likelihood of an accident and the severity of harm, even if the pedestrian exercises due care, and vice versa. Given this complementarity, whatever efforts at accident prevention the motorist takes above and beyond what the pedestrian does are wasted. The motorist pays for those extra efforts, but they do not reduce the probability of an accident. Likewise, it is unwise for the pedestrian to devote more care to accident prevention than the motorist. In this admittedly extreme example, the need to coordinate how much effort each party devotes to accident prevention is striking. The best thing is for the pedestrian and the motorist to take the same amount of care.

Asymmetric information makes coordination difficult. Suppose the motorist has a low cost of care. If the law induces him to take lots of effort—a decision consistent with his low cost—there is a fairly good chance that the effort will be wasted. The reason resonates. Requiring the motorist to use a high degree of care only makes sense if the pedestrian can be expected to mirror that choice. But the pedestrian will only take great care if she also has a low cost of care, which is far from certain. By applying the reasonable person standard, courts improves the chance of coordination by having the low-cost motorist chisel on his care. In other words, society wants the low-cost motorist to ignore his talent for harm prevention; to instead do what the "typical" motorist would do. Likewise, society is unforgiving to some (but not all) higher-cost pedestrians. It forces these pedestrians to bump up their care to some "average" level. Notably, through an objective standard, the law tosses away seemingly relevant information, namely each individual's cost of care.

Viewed from the low-cost motorist's perspective, the presence of pedestrians with higher costs of care infects the pool of counter-parties. The more pedestrians with high costs in the pool, the more (mis)-coordination becomes an issue and the greater the distortion in the care decision of the low-cost motorist. High-cost pedestrians are like having lemons in the used car market. Compression to the mean solves this problem: it provides sufficient certainty to low-cost agents that their efforts in taking care will not be wasted. With that said, compression to the mean necessarily entails a sacrifice in the fine-tuning of the law to the particular abilities of the actors. This results in a mismatch between what parties could do to prevent accidents and what the law demands. Stated differently, the law is both over and under-inclusive. It demands too little care from some actors and too much care from others.

In the used car market, it is not necessary to rid the market of all lemons to resolve the adverse selection problem — as long as the number of lemons remaining is not too large. The same insight applies in our setting. Coordinating sufficiently many agents on the same 'objective' care level may sufficiently mitigate adverse selection, as to enable the planner to tailor the standard of care for the remaining (highest cost) agents, deploying for them a standard more closely in align with their costs.

Second, we show that the breadth of the objective and excuse regions turns on the complementarity of the actors' efforts. As complementarity increases, the width of the pooling region increases and the width of the excuse region narrows. At the same time, the effort induced under the objective benchmark falls. In other words, with increased complementarity, more actors are subject to a less onerous objective standard. In fact, under certain conditions, we show that the planner will not customize the scheme at all. She will refuse to create incentives rooted in the actor's cost of effort, a parameter she learns for free. Instead, the planner holds all actors to the same, objective reasonable person standard, and that standard will correspond to the first-best standard of care for someone with the average cost of effort in the population.

Finally, we demonstrate that the decision of whether an agent should be held to a rigorous or forgiving standard (given her cost) can be decoupled from the decision what those standards should be. Specifically, the threshold that separates the types of agents who are pooled and those who are excused, depends only on the distribution of costs and is independent of the technology that translates effort into the common goal.

An important implication is that, even if the harm technology operates differently in different circumstances, the set of agents who can avail themselves of excuses will be unchanged. Thus, while what the law demands of adults will differ depending on the context (e.g., more care should be exercised when driving in raining conditions), the fact that all adults are held to the same standard will not. Likewise, what the law demands of children will always be more forgiving than what is demands of adults. But what exactly the more forgiving standard is will depend on the type of accident.

Similarly, take a senior manager overseeing a large number of different units. Each unit has a manager and a number of employees who work in teams. The model demonstrates that the senior manager can instruct the cadre of middle managers to only relax their employee's performance standards (whatever they are) for certain classes of individuals and adopt a non-discrimination or identical treatment policy for everyone else (i.e., treat all employees the same). The senior manager can then delegate to the units the task of constructing the standards based on local knowledge.

Likewise, a university can set a presumption of non-discrimination in tenure standards coupled with list of excuses for relaxation of the standard (parental obligations, COVID relief, etc.). After articulating these policies, the university can delegate to the departments decisions about the standard to be met, meaning these standards can vary dramatically across disciplines in line with local knowledge held by department chairs.[4]

[4]One might think of research as involving team production between the culture of the department established by senior faculty and the assistant professor.

We see this separation concretely in legal practice. The judge decides whether to grant an excuse from the reasonable person standard. The jury then decides whether the litigant met the appropriate standard (given the judge's ruling) based on their knowledge of local circumstances.

Before turning to the analysis, we pause to make a point about the relationship between objective standards and recent technological advances. Technological advances have made learning about individualized characteristics, such as the cost of accident prevention, cheaper and easier. Building off this, some scholars advocate that liability should become more and more personalized (Ben-Shahar & Porat, 2016). For instance, the skillful driver—revealed as such through data analysis by the state or insurer—should be held to a higher standard of care than the average driver.

Our analysis shows that, even if courts or regulators could use big data to learn everyone's personalized costs, the state could not harness this information to create better incentives. And thus there is no benefit to collecting the information in the first place. The issue is that individuals often must interact with one another before they know much about each other. And the promise of big-data is unlikely to plug this knowledge gap in the fleeting encounters between strangers.

After a brief discussion of related literature, the paper unfolds as follows. Section 1 develops a model of accident law. Section 2 articulates the first-best benchmarks, assuming the planner and the actors are fully informed. Section 3 studies efficiency in a second-best environment, where each agent's standard of conduct can only depend on their cost and not their counterparty's. In increasingly more generality, Section 3 demonstrates that the court prefers to hold actors to a objective reasonableness standard meshed with excuses for the least able. Further, it shows that the decision about which agents should be excused and which agents should be treated as identical can be disentangled from the decision about the level of care to require among these two sets of agents. Section 4 pivots to implementation, explaining how the legal rules found in practice can induce agents to make the second-best effort choices established in the prior sections. Section 5 considers some extensions and limitations of the model. We offer some concluding remarks and discussion about the predictions from the model in Section 6.

Related Literature

Our work relates and builds off different strands of literature. First, legal scholars and philosophers have explored the reasonable person standard. Philosophers argue that the reasonable person standard embodies positive virtues such as mutual respect, reciprocity, and fair terms of cooperation (Keating, 1995). Zipursky (2015) aptly summarizes this position, stating

> Reasonableness requires a sense of fitting one's demands alongside the multiple demands of others, which one accommodates to a certain extent (Zipursky, 2015, p. 1243)

Our model fleshes out what makes an interaction between two strangers 'fair'. It explains when (and why) the law should hold certain actors to the unwavering standard of conduct and when it should be more forgiving.

Other scholars view the reasonable person standard in a less positive light. For example, Bender (1988) argues that the reasonable person standard disguises what is, in fact, the raw exercise of judicial power, a power used to preserve existing social hierarchies. Bernstein (2001) suggests that through the reasonable person standard courts impose liability while pinning the reasons for it on some external community—the community of reasonable actors.

Not all legal scholars are so critical of objective standards. Landes and Posner (1987) define reasonable as actions that are consistent with cost-benefit analysis. Of course, as noted above, different people have different costs of accident prevention, suggesting the law should be finely tailored. Landes and Posner (1987) and Shavell (1987) show that the one-size-fits-all reasonable person standard arises when the courts cannot observe cost differences among individuals. Shavell (1987) offers another account, suggesting that the reasonable person standard operates as a tax, encouraging individuals with a high cost of compliance to shift away from the activity.

Unlike the prior literature, we assume the court knows each actor's cost of exercising care. Instead, frictions arise because the parties do not *themselves* know the cost of others with whom they interact, and thus cannot predict the actions of their counter-parties.[5]

We are unaware of any economic models that explore the coordination/customization trade-offs associated with the reasonable person standard. That said, our message is that the

[5]Garoupa and Dari-Mattiacci (2007) examine this same information problem in a model where care decisions are perfect substitutes. They show that a court will find it taxing to create appropriate incentives using a negligence standard, and advocate for fines instead. We characterize the optimal legal rule for any degree of complementarity between care decisions. Our focus is the reasonable person standard rather than the choice of the vehicle for controlling conduct.

planner does not want to use information that is perfectly accurate and free. A similar message appears elsewhere in the economics literature. Work on the principal-agent model reveals that, sometimes, the principal benefits from remaining ignorant about the characteristics of the agent (Cremer, 1995; Sappington, 1986). Ignorance enables the principal to avoid renegotiating in a way that damages the agent's ex-ante incentives. In a similar vein, Taylor and Yildirim (2011) demonstrate that by committing to blind review a principal can create better incentives for the agent to produce good projects. Likewise, ignorance about the past behavior of an opponent can create a strategic advantage, dulling the consequences of the first-mover advantage (Schelling, 1980, p.161).

Our insight differs from these works. Here, the planner perfectly observes effort. As a result, if the planner were able to condition each agent's incentives on their own cost of effort *and* their counter-parties, the planner could achieve the first-best. Nonetheless, because the agents are ignorant of each other, the planner prefers not to base incentives on information the agents do, in fact, know. In other words, if the planner cannot use all of the information it has to create incentives, it is better off using none of it.

This idea finds an analog in the literature on strategic ambiguity (Bernheim & Whinston, 1998). There, the parties refuse to explicitly condition behavior on verifiable information about, say, the manager if they cannot also explicitly condition the behavior of the worker. The refusal to condition enhances the freedom of the manager and, in so doing, facilitates more effective punishment for deviations by the worker from the implicit arrangement. This model has nothing to do with the credibility of punishment. Instead, the planner refuses to use free and accurate information because of a need to foster coordination and avoid waste among asymmetrically informed agents.

1 A Model of Accidents

The model consists of a large number of motorists (m) and pedestrians (p). The interactions between a motorist and a pedestrian can lead to an accident. Each motorist and each pedestrian must decide how much care to take in preventing the accident. The court or planner decides on a liability schedule. Facing that schedule, the motorist selects a care level of $x_m \geq 0$, and the pedestrian selects a care level of $x_p \geq 0$.

Motorists and pedestrians differ in the cost of exerting care. Let c_i be the unit cost of care for each player $i \in \{m, p\}$. The cost of care parameter, the player's type, is drawn from a continuous distribution $G_i(c)$, that admits a strictly positive density $g_i(c)$, with support on $[\underline{c}_i, \bar{c}_i]$, where $\underline{c}_i > 0$. For technical reasons, we require that $G_i(c)$ be 'sufficiently' log-concave.[6] Formally, let $r_i(c) = -c\frac{d^2 lnG_i(c)/dc^2}{dlnG_i(c)/dc}$ be the analog of the coefficient of relative risk aversion. We assume that $r_i(c) > 1$.[7]

The court observes each c_i perfectly. Thus, the court can, if it so chooses, successfully use a fully customized standard for both actors. While unrealistic, this assumption allows us to put aside the most common explanation for objective standards: namely, that courts find them cheap to apply. (Holmes, 1881; Landes & Posner, 1987; Shavell, 1987).

The efforts of the pedestrian and the motorist combine to determine the likelihood and size of harm. We wish to examine how the interdependence of effort choices influences what the court deems permissible behavior. Denote the expected harm from the accident as $\Pi(a)$, where a is a measure of the central tendency of the two agent's effort choices. As is standard in models of accidents, assume that the function Π is twice continuously differentiable, strictly decreasing and strictly convex ($\Pi'(a) < 0$ and $\Pi''(a) > 0$). Additionally, assume that Π satisfies the Inada conditions.

The measure of the central tendency of care between the agents, a,is constructed to reflect complementarity between care decisions. For tractability, we adopt the *ordered weighted average* (OWA) technology as the central tendency measure[8]:

$$a(x_p, x_m; \lambda) = \lambda \max\{x_m, x_p\} + (1 - \lambda) \min\{x_m, x_p\}$$

where $\lambda \in [0, 0.5]$. The OWA technology returns a weighted average of the two care levels, guaranteed to lie between the minimum and maximum care taken. Since $\lambda \leq 0.5$, the technology assigns more weight to the agent taking less care, reflecting the intuition that the more reckless actor drives the likelihood of harm.

[6] In settings where the agents draw costs from the same distribution, we can replace this assumption with the requirement that the distribution be unimodal.

[7] Many common distributions with positive support, including uniform, triangular, chi-square, gamma, log-normal, exponential, Pareto, and Beta (provided $\alpha, \beta > 1$) distributions, satisfy this assumption.

[8] Though it is not commonly used, the OWA technology has antecedents in the economics literature, most famously in the Hurwicz criterion (Hurwicz, 1951), which provides a method to balance optimism and pessimism when agents make decisions under uncertainty. Most commonly, it has been used to study decision-making under conditions of ambiguity (see Xiong & Liu, 2014; Yager, 2002, 2004). But the OWA technology has been applied in a wider range of contexts, including the modeling and measurement of inflation (León-Castro, Espinoza-Audelo, Merigó, Gil-Lafuente, & Yager, 2020), asset valuation (Doña, la Red, & Peláez, 2009), exchange rate forecasting (León-Castro, Avilés-Ochoa, & Gil Lafuente, 2016), risk analysis (Blanco-Mesa, León-Castro, & Merigó, 2018), and government accountability (Avilés-Ochoa, León-Castro, Perez-Arellano, & Merigó, 2018), amongst others.

This technology enables a parsimonious articulation of the degree of substitutability between the agents' care decisions. The parameter $\lambda \in [0, 0.5]$ captures the degree of substitutability of care between agents. When $\lambda = 0.5$, the technology simplifies to the arithmetic mean, $a(x_m, x_p) = \frac{1}{2}x_m + \frac{1}{2}x_p$, and care levels can be perfectly substituted for each other. When $\lambda = 0$, the technology reduces to the Leontief technology, $a(x_m, x_p) = \min\{x_m, x_p\}$, and substitution is not possible. Any degree of substitutability between these extremes corresponds to a value of $\lambda \in (0, 0.5)$.

We use the OWA technology to capture substitutability instead of the more familiar CES technology.[9] In our setting, the CES technology has an unappealing feature: when the motorist and pedestrian choose similar care levels, the CES function is locally approximated by a perfect substitutes technology regardless of the value of ρ.[10] We conjecture that complementarity in care motivates the planner's need and desire to coordinate behavior between heterogeneous agents. Yet the CES technology does not meaningfully capture complementarity when agents match their care decisions. And that makes it a poor fit for this problem.

2 Benchmarks

The planner's (or an efficiency-minded court's) objective is to minimize the sum of accident costs and prevention costs.[11] We begin with two benchmarks.

2.A. Unilateral Problem

Consider the optimal level of care in an analogous model with only a single actor. The average care, a, is the care taken by the actor. The optimal level of care x_u satisfies:

$$\min_{x_u} \Pi(x_u) + cx_u.$$

[9]Recall the CES technology is given by: $b(x_m, x_p; \rho) = \left(\frac{1}{2}x_m^\rho + \frac{1}{2}x_p^\rho\right)^{\frac{1}{\rho}}$, where $\frac{1}{1-\rho}$ is elasticity of substitution. Like the OWA technology, the CES technology includes perfect substitutes ($\rho = 1$) and perfect complements ($\rho \to -\infty$) as special cases. In section 5.A. we discuss the implications of assuming this alternative "smoother" technology. With some small qualifications, the results are robust to this alternative specification.

[10]To see this, note that for all ρ, a first-order Taylor approximation of the CES aggregator centered at (x_m^0, x_p^0) with $x_m^0 = x_p^0$ gives: $b(x_m, x_p; \rho) \approx \frac{1}{2}x_m + \frac{1}{2}x_p = b(x_m, x_p; 1)$.

[11]In *United States v. Carroll Towing Co.*, 159 F.2d 169 (2d. Cir. 1947), Judge Learned Hand set forth this understanding of the objectives of accident law. See also Brown (1973); Calabresi and Hirschoff (1971); Posner (2014); Shavell (1987).

The efficient care level is $x_u = [\Pi']^{-1}(-c) = z(c)$. The function z maps the cost of care into its optimal level for the unilateral actor and will play a starring role in what follows. It is easily shown that $z'(c) < 0$, so that as the cost of care rises, the efficient unilateral care level falls. The level of care, x_u, is efficient, and it is what the court strives to implement when designing the legal rules.

2.B. Bilateral Problem with Full Information

Now take two actors. In the full information benchmark, the actors' costs are observable to everyone — the court and the actors themselves — and so the planner (or court) can condition each agent's care level on *both* agents' costs. As in the unilateral benchmark, the first-best care decisions minimize the social loss associated with accidents and accident prevention:

$$\min_{x_m, x_p} \Pi(a(x_m, x_p; \lambda)) + c_m x_m + c_p x_p$$

The first-best care schedule is characterized as follows:

Proposition 1. *Let $\overline{\lambda}(c_m, c_p) = \frac{\min\{c_m, c_p\}}{c_m + c_p} \leq \frac{1}{2}$. If the care technology is characterized by:*

- *(Imperfect) Substitutes (i.e. if $\lambda > \overline{\lambda}(c_m, c_p)$), then:*

$$x_i^{1st}(c_m, c_p) = \frac{1}{\lambda} z\left(\frac{c_i}{\lambda}\right) \cdot \mathbf{1}[c_i < c_{-i}]$$

- *(Imperfect) Complements (i.e. if $\lambda \leq \overline{\lambda}(c_m, c_p)$), then:*

$$x_m^{1st}(c_m, c_p) = x_p^{1st}(c_m, c_p) = z(c_m + c_p)$$

where $\mathbf{1}[\cdot]$ denotes the indicator function.

The first-best care decisions exist in one of two regimes. If the degree of complementarity is high (i.e., λ is small), then the planner wants to coordinate both agents to take the same level of care. Moreover, since both agents must pay the cost of providing this care, the optimal care level coincides with the one that would be optimal for a unilateral agent facing unit cost $c_m + c_p$.

If, instead, the degree of complementarity is low (i.e., λ is relatively high), then the planner will assign the full burden of taking care to the 'least cost avoider', while allowing the higher cost agent to sit idle.[12] This result tracks the insight from the early literature in law and economics (Calabresi & Hirschoff, 1971; Demsetz, 1972) that courts should decide on liability by hunting for the least cost avoider, a call that holds sway among some the justices of the United States Supreme Court.[13]

Implementing the first-best is difficult, if not impossible. With substitutes, the court must assign the legal responsibility to take due care to the motorist when, and only when, her cost is less than the pedestrian's. To react appropriately to this rule, the motorist must know her own costs and the costs of every pedestrian with whom she interacts—a Herculean information requirement. Likewise, with complements, what counts as due care (and thus non-negligent behavior) is a function of the sum of the motorist's and the pedestrian's costs. Absent knowledge of the pedestrian's cost, the motorist cannot compute the sum and the associated marker. She therefore cannot understand what behavior the court would deem negligent.

Plainly stated, to induce first-best care decisions, the court must do more than uncover each actor's cost of care through litigation, as noted by Garoupa and Dari-Mattiacci (2007). At the time of the accident, the actors must know the cost of care for everyone else they might be involved in an accident with. This information is, in reality, unavailable.

With these benchmarks in mind, we next examine what the standards of conduct should be, given that the actors know a lot about themselves but little about others with whom they interact.

3 Second-Best Analysis

In the first-best, the planner (or court) was able to condition the care level of each agent on the costs of both agents. In the second-best, we constrain the planner to choose care levels

[12]The 'bang-bang' nature of this result, with the least cost abater being wholly responsible for care, is an artifact of the piece-wise linear average care technology. As we note in Appendix B, with a more convex technology, the first-best care levels in this region would be more 'continuous'.

[13]*Air & Liquid Sys. Corp. v. DeVries*, 139 S. Ct. 986, 997 (2019)(Gorsuch, J., dissenting)("The manufacturer of a product is in the best position to understand and warn users about its risks; in the language of law and economics, those who make products are generally the least-cost avoiders of their risks.").

for each agent only on the basis of that agent's cost and without regard to the counter-party's cost, a cost the agent does not know.

To motivate and illustrate the main ideas of the second-best, we start with a baseline case. Let us assume the care decisions are perfect complements and the actor's costs are drawn from the same distribution. The court observes each actor's costs, but the actors do not observe the cost realization of their counterparty.

With this information structure, the court is able to impose liability as a function of each actor's costs, but not as a function of the "pair" of costs. Should it do so? Seemingly yes.

The planner (or court) chooses care schedules $x_m(c)$ and $x_p(c)$ to solve:

$$W = \min_{x_p(c), x_m(c)} \iint \{\Pi(\min\{x_p(c_p), x_m(c_m)\}) + c_p x_p(c_p) + c_m x_m(c_m)\} g(c_p) g(c_m) dc_p dc_m$$

Straightforwardly, the second-best care schedules $x_i(c_i)$ are continuous (by Berge's Theorem of the Maximum) and weakly decreasing in c_i.

If the second-best care schedules are strictly decreasing, they *separate* each agent's types according to their cost of care; the care schedules perfectly tag the agent's effort to the cost of its provision. Yet tailoring care to how much it costs is only part of the court's calculus. The court also wants to facilitate coordination to avoid wasted effort.

We will now show that coordination concerns ensure that the court holds actors whose costs lie an interval $[\underline{c}, \hat{c}]$ to the same objective standard. In other words, the planner sets the standard without reference to these actors' costs. The argument proceeds by first positing a perfectly separating schedule, and then showing that this schedule cannot be strictly decreasing as is required in any separating solution.

In the separating schedule, the care level for each type must satisfy the first-order condition. Notice that a marginal increase in the pedestrian's care only reduces accidents when she takes less care than the motorist. Otherwise, the additional care is wasted. But, because the pedestrian does not know the motorist's cost of care (or equivalently, the care level the motorist with that unit cost has been encouraged to take by the court), the pedestrian's care can only be fixed according to the distribution of costs. The pedestrian therefore must treat the motorist's care as a random variable. The court, of course, learns the cost parameters *ex-post*. Yet the court cannot condition behavior on something the agents themselves do not know *ex-ante*.

Focusing on the pedestrian, we have the following first order condition:

$$\Pi'(x_p)\,\Pr[x_m(c_m) > x_p(c_p)] + c_p = 0$$

where the probability is taken with respect to the distribution over c_m.

Next we exploit the symmetry in our setup (and in particular the assumption of identical cost distributions) to assert that the motorist and pedestrian must have the same care schedules, i.e. $x_p(c) = x_m(c) = x(c)$ for all c.

The probability that the pedestrian's care actually reduces accidents is:

$$\Pr[x_m(c_m) > x_p(c_p)] = \Pr(c_m < c_p) = G(c_p)$$

Because the schedule must decrease in c, the pedestrian's care is lower when they have a higher cost than the motorist. And that happens with probability $G(c_p)$.

Accordingly, the first-order condition becomes:

$$\begin{aligned} c_p + G(c_p)\Pi'(x_p) &= 0 \\ x(c_p) &= [\Pi']^{-1}\left(-\frac{c_p}{G(c_p)}\right) = z\left(\frac{c_p}{G(c_p)}\right) \end{aligned} \tag{1}$$

The ratio $\frac{c_p}{G(c_p)}$ represents what we denote as the pedestrian's *effective* cost of care. It is the cost of increasing the average care level by 1 unit *in expectation*, given that marginal effort is sometimes wasted. The effective cost of care increases in the actor's unit cost and decreases as the agent's effort becomes more likely to be pivotal and not wasted. It is this cost, the court uses to determine the care level it wants the agent to take.

Consider now the lowest cost type. Suppose this type takes a hefty dose of care, care consistent with her meager cost of providing it. Further, suppose every other type takes care tailored to their cost. In that setting, the care of the lowest cost type is wasted for sure.

As we approach the lowest cost type, we have $\lim_{c\to\underline{c}} \frac{c}{G(c)} = \infty$ (since $\underline{c} > 0$). With an infinite effective cost of care, it follows from the Inada conditions that $x(\underline{c}) = 0$. Yet, in a separating decreasing schedule, agents with costs a little above $\underline{c}$ must do strictly less than zero, and this cannot be.

To explore the logic another way, notice that since $z' < 0$, the second-best schedule will be decreasing whenever $\frac{c}{G(c)}$ is increasing. But, given that $\underline{c} > 0$, it must be that $\frac{c}{G(c)}$ is decreasing for c close to $\underline{c}$. In fact, combined with the assumption that G is sufficiently log-concave (or that g is unimodal), we can show that $\frac{c}{G(c)}$ is first decreasing and then increasing.

Thus, we have shown both that $x(c)$ cannot be strictly decreasing when c is low and that it may be strictly decreasing when c is large. This explains an asymmetry: excuses or relaxed standards exist for high-cost actors but low-cost actors are not subject to heightened standards. As we have explained, the reason is that adverse selection has its strongest bite when applied to agents taking higher care levels. Furthermore, tailoring is only optimal when c is large enough.

This baseline case assumes perfect complementarity of care, creating a big push to coordinate conduct. Yet, our result is not simply that complementarity demands perfect coordination. We still find that the court has a counter-veiling incentive to forgive high-cost agents and have them put forth less effort than others. Indeed, if the court can induce sufficient matching among low-cost agents, it becomes worthwhile to grant relief to high-cost agents.

Having established that the solution must involve pooling, the next step involves: (a) determining the breadth of the pool region and (b) the care level for the pool members.

Suppose the court pools agents with costs $[\underline{c}, \tilde{c}]$ for some arbitrary $\tilde{c} > \underline{c}$. The care level that minimizes the social loss within the pool is the solution to:

$$\min_x \Pi(x)G(\tilde{c})^2 + 2x \int_{\underline{c}}^{\tilde{c}} cg(c)dc$$

The first term is the probability the pedestrian and motorist both draw costs in the pooling region and thus are pivotal to determining harm. The second term is the cost born by agents in the pool (whether their care is pivotal or not). Taking the first order condition gives:

$$\tilde{x} = z\left(\frac{2E[c \mid c < \tilde{c}]}{G(\tilde{c})}\right) \tag{2}$$

Let us interpret this. As expressed in proposition 1, with perfect complements, the first-best care level turns on the sum of the cost realizations of the pedestrian and motorists. In the second-best, conduct is determined by the sum of the expected *effective* costs among those in the region adjusted for the probability of waste. (Since the agents are drawn from the same cost distribution, the sum of expected costs is simply twice the conditional expectation.) The second-best care level is simply the first-best standard in one focal circumstance: when the motorist realizes the average cost and she is paired with a pedestrian who realizes the average cost. Thus, the pooling standard is, in fact, a reasonable person standard — it treats all agents as if they were the average person.

We are left to characterize the boundary (which we denote by $\hat{c}$) between the pooling and separating regions of the second-best schedule. Intuitively, the threshold agent must be indifferent between the objective standard and lodging an excuse. Let $\hat{x}$ denote the pooling care level implied by pooling region $[\underline{c}, \hat{c}]$. The threshold type's care level must satisfy:

$$\hat{x} = z\left(\frac{\hat{c}}{G(\hat{c}}\right) = z\left(\frac{2E[c \mid c < \hat{c}]}{G(\hat{c})}\right)$$
$$\hat{c} = 2E[c \mid c < \hat{c}] \tag{3}$$

Figure 1 illustrates the mechanics behind the result. The horizontal axis reflects the cost parameter. The solid (red) curve represents the optimal separating level of care, where the standard of care is tailored to the agent's cost. This is given by equation (1) above. Notice that this function assigns strictly more care to higher cost agents for $c < c'$, which we know cannot be in any solution. As a result, The planner must at least pool agents in the region $[\underline{c}, c']$. However, an even broader pool may be optimal.

The dashed (blue) curve is the optimal care for the pooling types when the interval $[\underline{c}, \tilde{c}]$ forms the pool (where the horizontal axis now measures $\tilde{c}$). This is given by equation (2). This function must be increasing whenever it lies below the red curve, and decreasing when the opposite is true. Intuitively, if the agent at the threshold or margin of joining the pool would individually be willing to take more care than the average agent in the pool, then adding that agent to the pool will increase the optimal care level within the pool (adding a marginal type above the average increases the average).

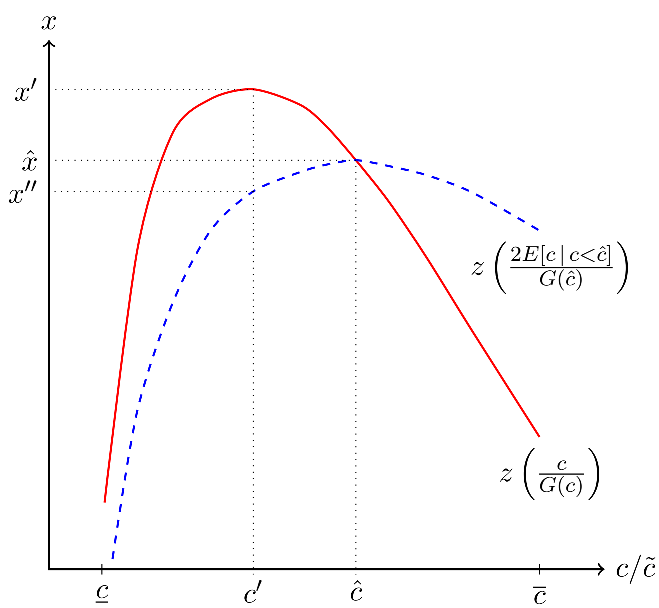
If the pooling region were limited to $[\underline{c}, c']$, then the optimal standard within the pool x'' would be below the optimal care for the threshold type x' (as well as for agents with costs slightly above the threshold). But this (1) violates the requirement that the second-best schedule be decreasing since many excused types exert more care than the pooled types and (2) breaks the needed indifference for the threshold type between pooling and separating.

Instead, the cost type where the solid and dashed curves intersect identifies the marker between the objective and tailored standards. At that point, the border type is indifferent.

Moreover, with this threshold type, the care level taken by the agents within the pool is maximized.

To understand why, notice that the breadth of the pooling region trades-off two competing forces. On the one hand, broadening the pool increases the average cost within the pool, which causes the optimal pooling care level to decrease, *ceteris paribus*. On the other hand, broadening the pool decreases the probability that agents within the pool will have their effort wasted by being matched to an agent outside the pool (who takes less care). Improved matching reduces the *effective* costs of care and thus increases the optimal pooling care level. The $\hat{c}$ defined by (3) makes this trade-off optimally, resulting in the highest possible care from agents within the pool.

Figure 1: Breadth of the Objective Test



Drawing these insights together, we have the following proposition.

Proposition 2. *Suppose c_m and c_p are independent draws from the same distribution. Then $x_m(c) = x_p(c) = x(c)$. There exists a unique threshold $\hat{c} > \underline{c}$ characterized by $\hat{c} = 2E[c|c < \hat{c}]$, s.t.*

$$x^{2nd}(c) = \begin{cases} z\left(\frac{2E[c|c<\hat{c}]}{G(\hat{c})}\right) = z\left(\frac{\hat{c}}{G(\hat{c})}\right) & \text{if } c < \hat{c} \\ z\left(\frac{c}{G(c)}\right) & \text{if } c \geq \hat{c} \end{cases}$$

Furthermore, $\frac{\partial x^{2nd}(c)}{\partial c} < 0$ whenever $c > \hat{c}$.

These results shed light on a number of phenomena. First, suppose that $\hat{c} > \overline{c}$. (This will occur if $\overline{c} > 2E[c]$.) In that case, the court holds all motorists and all pedestrian's to the same objective benchmark. Related to the discussion above, this standard coincides with the first-best decision rule when: (1) effort decisions are complements and (2) a motorist with average costs is paired with a pedestrian with average costs, i.e. $x^{2nd} = z(E(c)+E(c))$. Thus, 'reasonableness' reflects an appeal to the statistical average cost to manage coordination challenges and avoid waste. The planner is not constrained to use the cost of the average agent to determine standards of conduct. Instead, she finds it desirable to do so, given the agents' ignorance about each other.

If $\hat{c} < \overline{c}$, then the second-best care schedule is characterized by partial pooling; agents with low costs are held to an objective standard, while agents with high costs are excused and held to a tailored and lower standard, instead. The optimal pooling care level is a 'modified reasonable person standard." Moreover, this modified standard demands a higher level of care from agents in the pool than would be the case if the pool included all agents. Thus, the availability of excuses not only provides relief to high-cost agents. It also makes certain that low-cost agents are able to provide the highest level of care achievable, given the incentive problems that arise in the private information environment.

For excuses to arise, we need $\hat{c} < \overline{c}$, which in turn requires that $\overline{c} > 2E[c]$. Given the bounded support, this condition is satisfied when the distribution of costs has a sufficiently long tail relative to the mean cost - i.e. most agents have low costs, but there are a small tail with relatively high costs. As we discuss in a later section, forces that tend to truncate the distribution of costs or abilities, such as licensing requirements, will also make pure reasonable person standards without the possibility of excuses more likely.

The following example illustrates these ideas.

Example 1. Suppose the costs are distributed according to a triangle distribution with a mode at $\underline{c}$, so that $g(c) = 2\left(\frac{\overline{c}-c}{\overline{c}-\underline{c}}\right)$. Then, the conditional expectation is:

$$E[c \,|\, c < \hat{c}] = \begin{cases} \frac{\overline{c}(\hat{c}+\underline{c}) - \frac{2}{3}\left(\hat{c}^2+\underline{c}\hat{c}+\underline{c}^2\right)}{2\left(\overline{c}-\frac{\hat{c}+\underline{c}}{2}\right)} & \text{if } \hat{c} < \overline{c} \\ \frac{2}{3}\underline{c} + \frac{1}{3}\overline{c} & \text{if } \hat{c} > \overline{c} \end{cases}$$

.

The fixed point of expression (3) is:

$$\hat{c} = \begin{cases} \frac{\sqrt{24\overline{c}\underline{c} - 15\underline{c}^2} - \underline{c}}{2} & \text{if } \overline{c} > 4\underline{c} \\ \frac{4}{3}\underline{c} + \frac{2}{3}\overline{c} & \text{if } \overline{c} < 4\underline{c} \end{cases}$$

There is a pure reasonable person standard whenever $\overline{c} < 4\underline{c}$ — i.e. whenever the tail of the distribution is relatively short. Otherwise, there is a reasonable person standard coupled with excuses.

3.A. Doctrinal Implications and Discussion

Now we turn to the legal implications of the analysis. As noted in the introduction, the law most often assesses behavior against the "reasonable person" standard. Beyond this benchmark, the law lowers and customizes the standard for certain classes of individuals who experience a high-cost of compliance, while rarely holding individuals with a low-cost of compliance to an elevated standard.

Torts, for example, defines the duty of care as what a "reasonable person" would have done under like circumstances. But children and the individuals with physically disabilities are held to a standard of care consistent with a child of that age and experience or an individual with that specific disability. When it comes to excuses from the reasonable person standard, the conduct standard is finely tailored to the costs and abilities of the agent in question.

Some doctrinal examples showcase how the law reflects the insights in proposition 2. On tailoring, consider *Friedman v. State*.[14] The plaintiff was a sixteen-year-old girl who worked as a counselor at a summer camp. On her day off, she went with a male camp counselor to a ski resort. The chair lifts were open to take individuals to the top of the mountain for picnicking. The ski resort stopped the chair lift for the night while the plaintiff and her fellow counselor were still on it. Rather than spend the night alone with a male, the plaintiff jumped off the lift, sustaining injuries. The issue was whether her conduct rendered her contributorily negligent and therefore ineligible for relief.

In finding for the plaintiff, the court followed the customized standard for children, stating "In evaluating the issue of contributory negligence, as it related to this infant, the fact of freedom from negligence is even more evident when we consider her age, judgment, experience, and

[14] 282 N.Y.S.2d 858 (Ct. Cl. 1967).

education." The court further commented on the characteristics of that particular plaintiff, opining:

> [I]t does not require much imagination or experience to determine that a lightly dressed 16-year-old city girl might become hysterical at the prospect of spending a night on a mountainside, suspended in the air and with no apparent reason to hope for rescue until the next morning. Secondly, we must add to the fact of expectable hysteria, the moral compulsion *this* young lady believed she was under, not to spend a night alone with a man. *Id.* at 862 (emphasis added)

The ski resort likely had little knowledge of the plaintiff's enhanced cost of care. It is unclear whether the ski resort employees even saw the plaintiff. If they did, would they have presumed she was religious and with a male who was not her husband or brother? In other words, the *Friedman* court based its ruling on traits that the the defendant could not have—or would find it difficult to—observe.

The *Friedman* ruling should be contrasted with the case established in the reasonable person standard in torts: *Vaughn v. Menlove*.[15] In that case, the defendant stacked a hay rick on the edge of his land. His neighbor told him it might catch fire. The defendant responded that he would "chance it." The hay rick ignited and burned down his neighbor's house. At trial, the defendant stated the issue as "whether [the defendant] had acted bona fide to the best of his judgment; if he had, he ought not to be responsible for the misfortune of not possessing the highest order of intelligence." The court refused to relax the standard to account for the defendant's limited intelligence. Instead, the court held the defendant to the standard of ordinary prudence.

Presumably, the defendant in *Vaughn* had a higher cost of care than the average person. But, unlike the plaintiff in *Friedman*, he didn't fall into one of the classes eligible for tailoring and the court forced him to the pooling standard.

Commenting on the reasonable person standard, Holmes (1881) honed in on expectations as a reason for objective standards, concerns formalized in the model. He wrote:

> If, for instance, a man is born hasty and awkward, is always having accidents and hurting himself or his neighbors, no doubt his congenital defects will be allowed for in the courts of Heaven, but his slips are no less troublesome to his neighbors

[15] 132 Eng. Rep. 490 (1837).

> than if they sprang from guilty neglect. His neighbors accordingly require him, at his proper peril, to come up to their standard, and the courts which they establish decline to take his personal equation into account.

Finally, while the law is on occasion forgiving of agents with high costs, it is often not demanding of agents with low costs. In *Fredericks v. Castora*,[16], for example, the plaintiff was a passenger in a truck driven by an experienced trucker. The driver did a U-turn across four lanes of traffic, and, in so doing, was struck by another truck. The plaintiff argued that the truck drivers should be held to a higher standard of care than the typical driver of a motor vehicle because they were professional truck drivers. The court refused. The court did not allow evidence of a lower cost of care into the proceeding, a holding consistent with the predictions of the model.[17]

Other properties of the solution generate additional understanding. As a direct consequence of expression (3), we have the following corollary:

Corollary 1. *The threshold $\hat{c}$ can be determined without reference to the accident reduction technology Π and the conduct required under the standard.*

The standard of care $x(c)$ that the court assigns to an agent will depend on the accident reduction technology Π (through the function $z(c) = [\Pi']^{-1}(-c)$). This dependence on Π is reflected in the instruction to juries to assess the agents' conduct "under the circumstances". Conduct that might have been reasonable on a sunny day might not remain so during a snow storm.[18] The Corollary shows that neither this mapping, nor the circumstances prevailing in the case at hand, bear on whether to hold a particular agent to the pooling standard or to offer them an excuse. The intuition again stems from the nature of the adverse selection problem; if the purpose of pooling is to mitigate the problem of there being too many lemons in pool, then the extent of pooling will depend on the distribution of types, and not the external technology that maps types into choices.

This dis-aggregation has implications for institutional design. An institution can separate **policy** decisions—who gets leniency?—from **application**—what should the standard be and

[16] 360 A.2d 696 (1976).

[17] The same result arose in *LaVine v. Clear Creek Skiing Corp.* , 557 F.2d 730 (10th Cir. 1977). There, a ski instructor crashed into the plaintiff while skiing. The plaintiff sought a jury instruction that a ski instructor should be held to a higher level of care, given his expertise. The court rejected the invitation. Others argue that courts should take these facts into consideration. Restatement (Third) of Torts: Gen. Principles § 10 (1999). This analysis suggests a reason to avoid doing so.

[18] See *Staunton v. City of Detroit*, 46 N.W.2d 569, 573 (1951)('[T]he proofs in the case indicate that at the time of the accident weather conditions were such as to require special care and caution on the part of the driver of defendant's bus.").

did the agent meet it? In so doing, the institution can then assign different decision rights to different actors.

For example, as discussed before, an institution might consist of a high-level policymaker and a cadre of low-level "managers." The policy-maker decides who is excused from hitting objective benchmarks and who is not. The "managers" can then use local knowledge to compute the needed benchmark against which behavior (for both pooled and excused workers) is judged.

As noted briefly in the introduction, tort law reflects this design choice. To prevail on a tort claim, the plaintiff must prove that the defendant owned the plaintiff a duty and that the defendant breached that duty. The judge decides the "duty" question: whether the defendant should be held to the reasonable person standard or granted a reprieve and held to a lower standard. The jury decides the breach question: whether the defendant acted as a reasonable person (possibly subject to an excuse) would "under like circumstances."

Prior models of accidents specify the standard of care as the result of minimizing social costs. These models do not specify which actors, judge or jury, should decide liability. They do not separate breach from duty. The reason is that, unlike us, prior work does not confront the difficulty of coordinating behavior in the face of asymmetric information between strangers.

Notably, Corollary 1 also implies that, while the standard required to be deemed non-negligent may vary across time and context, the frequency with which the court grants excuses from that standard should not. This insight contributes to a longstanding debate in tort law: why are courts resistant to extending excuses to new classes of individuals? For example, scholars and advocates have lobbied courts to apply a more lenient standard of care for the mentally disabled (Dark, 2004; Eggen, 2015; Lindquist, 2020). Courts have refused, reasoning that mental disability is easier to fake than physical disability.[19] But because of advances in mental health diagnosis, this argument is less persuasive today.

At the same time, mental disability and insanity are considered relevant for determining criminal culpability.[20] This analysis identifies a key difference between criminal law and tort

[19]See *Creasy v. Rusk*, 730 N.E.2d 659, 66 (Ind. 2000)("The public policy reasons most often cited for holding individuals with mental disabilities to a standard of reasonable care in negligence claims include . . . [It] removes inducements for alleged tortfeasors to fake a mental disability in order to escape liability."). See also Restatement (Second) of Torts Â§ 283B cmt. b(2).

[20]See State v. Searcy, 798 P.2d 914, 917 (1990); Model Penal Code, Â§ 4.01 ("(A) person is not responsible for criminal conduct if at the time of such conduct as a result of mental disease or defect he lacks substantial capacity to appreciate the criminality (wrongfulness) of his conduct or to conform to the requirements of the law."

law. Crimes are much less likely to involve coordinating behavior, and thus the need to fix the expectation of others—our key driver—becomes less important.

As indicated, prior work justifies the reasonable person standard as a response to the court's costs of measuring differences in aptitude. Under that view, changes in the measurement technology should result in changes in the standard. Yet, as noted, the law has not reflected this theme. Our model, by contrast, predicts stability in the classes of excused actors.

3.B. Comparative Statics

We now turn to comparative statics as to the breadth of the pooling region and the breadth of the excuse region. Both of these regions depend on the location of $\hat{c}$, which in turn depends on the distribution of costs $G(c)$.

To do comparative statics on the properties of that distribution, let $c \sim G$ and $c \sim H$ denote two different distributions from which both the pedestrian's and motorist's cost are drawn. Let $\hat{c}_G$ and $\hat{c}_H$ be the corresponding thresholds for both the motorist and the pedestrian. Recall that these thresholds are defined by $2E_i[c|c < \hat{c}_i] = \hat{c}_i$ for each $i \in \{G, H\}$, so that the comparative static results largely depend on the behavior of the conditional expectation as the distribution changes. We have the following results:

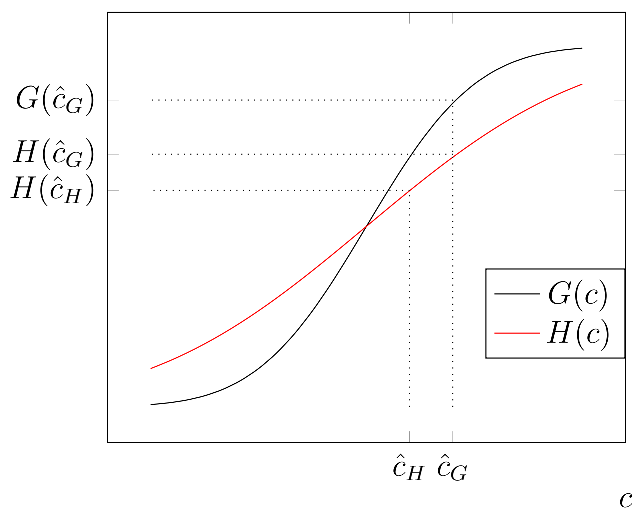
Lemma 1. *The threshold $\hat{c}$ is responsive to the distribution of costs in the following way:*

1. ***Scaling**: Suppose $c_H = \kappa c_G$ with $\kappa > 0$. Then $\hat{c}_H = \kappa \hat{c}_G$, and $H(\hat{c}) = G(\hat{c})$. The threshold scales and the probability of excuses is unchanged.*
2. ***Mean Preserving Spread**: Suppose that (a) H is a MPS of G, and (b) $H(\hat{c}_G) < G(\hat{c}_G)$ (i.e. the c_G standard applied under the spread distribution produces more excuses). Then $\hat{c}_H < \hat{c}_G$, and $H(\hat{c}_H) < H(\hat{c}_G) < G(\hat{c}_G)$. There will be more excuses.*

The scaling comparative static teaches us, unsurprisingly, that the units of cost measurement do not impact the probability of an excuse. The mean preserving spread result is of more interest. The planner cares about (a) the average cost of those agents in the pool and (b) the possibility of waste. By contracting the pool size, the planner decreases the average cost of pool members while increasing the probability of waste.

The reshuffling of agent types under a mean preserving spread makes shrinking the pool attractive. The reason is that the pool has many more low-cost agents in it (and thus a lower average cost). Thus, the planner gladly tolerates an increased chance of mismatched effort to achieve a higher standard of care among those subject to the objective standard. In obtaining this result, we restrict attention to the most natural setting, cases where excuses are rare. Formally, we assume that the old standard applied to the new distribution results in more excuses, as represented in the Figure 2, where the distribution H is a mean-preserving spread over G. Since $H(\hat{c}_H) < G(\hat{c}_G)$, the higher spread environment pools fewer agents and offers excuses to more frequently.

Figure 2: Mean Preserving Spread and Excuses



Observe the analog to licensing. A licensing requirement prohibits high-cost actors from engaging in the activity. Our model provides a new justification for this prohibition. With fewer individuals at the high-cost end, coordination becomes less of a challenge. The planner then can restrict the pool size. With a smaller pool, the planner optimally imposes a higher standard of conduct on anyone subject to the objective standard. This idea finds support in the law: licensed professionals are held to a higher standard of care. Further, a pure reasonable person standard will be more likely applied in situations where licensing occurs, and the availability of excuses will likely be curtailed.

To summarize, the analysis thus far gestures toward an efficiency-based account of the reasonable person standard that rationalizes the law turning a blind eye to the individual particularities of an agent's conduct. The court's self-imposed ignorance solves an adverse selection problem: namely the distortion in care that would arise among low-cost actors because they fear involvement in an accident with a high-cost actor. This problem is endemic to interactions between strangers; it will arise whenever the agents do not know salient details about their counterparty at the time of their interaction and arises even when the court can observe these details *ex post* and condition its rulings accordingly. To fix this distortion, the law coalesces around a unified standard of care, occasionally coupling this standard with a limited number of opt-outs for the least able.

3.C. Generalizations

Having explored the baseline case, we turn to more general settings. We first extend the results to the case where care decisions are imperfect complements, retaining the assumption that costs are drawn from identical distributions. Next, we derive results assuming non-identical distributions of costs, but retaining the assumption of perfect complements.

3.C.1. Imperfect Complements

In the baseline case whenever an agent took more care than her counterparty that excess care was wasted. With imperfect complements, excess care is no longer completely wasted; it reduces the probability of harm, though the size of the harm reduction may be small.

The planner/court might nonetheless wish to coordinate the efforts of agents with different costs because the gain of having the lower-cost type take more care, though it is positive, is not worth the additional cost. Formally, we have the following result:

Proposition 3. *For every $\lambda \in [0, 0.5]$, there exists $\hat{c}(\lambda) \geq \underline{c}$, such that the second-best schedule applies an objective standard $\hat{x}(\lambda)$ to all agents with costs $c < \hat{c}(\lambda)$. Additionally, $\hat{c}(\lambda)$ is strictly decreasing in λ. Finally, if $\lambda = 0$ (perfect complements) then $\hat{c}(0) = 2E[c\,|c < \hat{c}]$, and if $\lambda = .5$ then $\hat{c}(0.5) = \underline{c}$.*

Proposition 3 demonstrates that our results do not turn on the strong assumption of perfect complements. Generically, for any degree of substitutability/complementarity, the optimal second-best schedule will apply an objective standard to low-cost agent and excuse the conduct of the highest cost agents. The breadth of the pooling interval, however, narrows as the agents' care becomes increasingly substitutable. In the perfect substitutes limit, the pooling region disappears entirely.

The comparative static is testable. As care can more easily be interchanged (e.g., more care by the doctor in evaluating injury can be replaced to much the same effect by more care by the patient in explaining the circumstances surrounding the injury), the legal regime should entail more and more tailoring.

Now consider the implication of proposition 3 for the emergence of a pure reasonable person standard, the case where the court ignores all cost information in setting the standard. Recall, in the case of the perfect complements, the court disregards all personalized cost information, and holds *all* agents to the same standard when $\hat{c}(0) > \overline{c}$ (which occurs when $\overline{c} < 2E[c]$). Further, this standard aligned with the first-best standard when matching the average pedestrian and the average motorist. The same remains true as the care technology becomes somewhat substitutable. We formalize this result in Corollary 2 below.

Corollary 2. *Suppose $\overline{c} < 2E[c]$, so that under perfect complements, the second-best schedule is a pure objective rule. Then, there exists $\overline{\lambda} = 1 - \frac{\overline{c}}{2E[c]} > 0$, such that the second-best schedule remains a pure objective rule for all $\lambda < \overline{\lambda}$ (equivalently, if $\overline{c} < 2(1-\lambda)E[c]$). Moreover, this objective standard is the reasonable person standard $\hat{x}(\lambda) = z(2E[c])$.*

Finally, consider the separation of policy and application. The baseline case showed that decisions about the breadth of the pooling region could be made independently from decisions about the conduct demanded under the standard. That result extends:

Corollary 3. *With imperfect complements, the location of $\hat{c}(\lambda)$ is independent of the accident reduction technology Π.*

3.C.2. Non-Identical Distributions

Let us return to the perfect complements environment.Suppose that the motorist and pedestrian draw costs from the different distributions. Let $G_m(c)$ and $G_p(c)$ be two different continuous distributions satisfying the assumption previously described.

Because the distributions are different, the second-best care schedules, $x_m(c_m)$ and $x_p(c_p)$, will have different shapes. More, the motorist and the pedestrian will have different pooling thresholds.

Given these complications, some new notation helps explicate the results. With perfect complements, the motorist with cost c_m must estimate the probability that the pedestrian he encounters will take less care. Define the pedestrian type $c_p(c_m)$ as the pedestrian who the court induces through legal rules to take the same care as the motorist with cost c_m;

that is, $x_m(c_m) = x_p(c_p(c_m))$. In the separating region, if it exists, the first-order conditions imply

$$x_m(c_m) = z\left(\frac{c_m}{Pr(x_p(c_p) > x_c(c_m))}\right) = z\left(\frac{c_m}{G_p(c_p(c_m))}\right)$$
$$x_p(c_p(c_m)) = z\left(\frac{c_p(c_m)}{Pr(x_m(c_m) > x_p(c_p(c_m)))}\right) = z\left(\frac{c_p(c_m)}{G_m(c_m)}\right),$$

where we use the fact that schedules are locally invertible. By construction $x_m(c_m) = x_p(c_p(c_m))$ and thus we can define the function $c_p(c_m)$ implicitly by:

$$c_m G_m(c_m) = c_p G_p(c_p).$$

Analogously we can define $c_p(c_m)$, and note that $c_p(c_m)$ and $c_m(c_p)$ are inverse functions.

We have the next result, a generalization of Proposition 2:

Proposition 4. *There exist threshold $\hat{c}_m > \underline{c}_m$ and $\hat{c}_p > \underline{c}_p$ uniquely defined by:*

1. *$\hat{c}_m = c_m(\hat{c}_p)$ (or equivalently, $\hat{c}_p = c_p(\hat{c}_m)$), and*
2. $\frac{E[c_m \,|\, c_m < \hat{c}_m]}{\hat{c}_m} + \frac{E[c_p \,|\, c_p < \hat{c}_p]}{\hat{c}_p} = 1$

such that:

$$x_i^{2nd}(c_i) = \begin{cases} z\left(\frac{E[c_m|c_m<\hat{c}_m]}{G_p(\hat{c}_p)} + \frac{E[c_p|c_p<\hat{c}_p]}{G_m(\hat{c}_m)}\right) = z\left(\frac{\hat{c}_i}{G_{-i}(\hat{c}_{-i})}\right) & \text{if } c_i < \hat{c}_i \\ z\left(\frac{c_i}{G_{-i}(c_{-i}(c_i))}\right) & \text{if } c_i \geq \hat{c}_i \end{cases}$$

Proposition 4 is a natural generalization of Proposition 2. In the region of excuses, the agent chooses the unilaterally best care level given their *effective* cost. Modified costs simply inflate the agents true cost by the inverse of the probability that the opponent takes more care. Since the opponent with cost $c_{-i}(c_i)$ takes the same care level of agent i with cost c_i, the modifier term is simply $G_{-i}(c_{-i}(c_i))$. Similarly, in the pooling region, the agents coordinate upon the optimal first-best care level given the expected *effective* costs of the agents in the pool. Yet again, when there is complete pooling (which will occur whenever $\max\{\overline{c}_m, \overline{c}_p\} < E[c_m] + E[c_p]$), all agents will be held to a common reasonable person standard, which is simply the first-best standard when applied to the average motorist and average pedestrian, having costs $E[c_m]$ and $E[c_p]$, respectively.

4 Implementation: Designing the Legal Rules

The analysis so far has focused exclusively on the planner's problem identifying the optimal care levels for the motorist and pedestrian, constrained by the information those agents possess at the time they act. But are these optimal choices implementable, and, in particular, are they implementable under the standard liability rules used by courts? The answer is yes. The standard logic from Shavell (1987) applies.

Take two common liability rules utilized by courts: a pure negligence rule, and a strict liability rule with a defense of contributory negligence. The pure negligence rule holds the defendant liable for harm suffered by the plaintiff only if the defendant took less than due care. A strict liability rule with a defense of contributory negligence holds the defendant liable for the harm unless the *plaintiff* took less than due care. The former does not specify a care level for the plaintiff, and the latter does not specify a care level for the defendant, though as we shall see, in equilibrium, both agents take efficient care. The two rules are broadly similar, and differ only in which agent is held to be the residual bearer of harm: under negligence, it is the plaintiff; under strict liability with contributory negligence, it is the defendant.

Proposition 5. *A pure negligence rule and a strict liability rule with contributory negligence will both implement the second-best schedules $(x_m^{2nd}(c_m), x_p^{2nd}(c_p))$. Formally, consider*

- *A pure negligence rule that establishes a standard of care for the defendant (i.e. the motorist) $x_m^{2nd}(c_m)$.*
- *A strict liability rule with a defence of contributory negligence that establishes a standard of care for the plaintiff (i.e. the pedestrian) $x_p^{2nd}(c_p)$.*

Under either rule, it is Nash equilibrium for both agents to take their second-best care level.

We have assumed the court can observe the actors' costs perfectly. This is not strictly necessary. Emmons and Sobel (1991) show that the second-best schedules can be implemented as a Nash equilibrium of a game between the motorist and pedestrian so long as society allows for sufficiently rich liability rules (including the possibility of punitive damages).

5 Extensions and Limitations

5.A. Central Tendency Technology

The main analysis relied on an Order Weight Average technology to express complements and substitutes in care. The technology is attractive for its simplicity. Also, as noted, unlike the more familiar CES technology, the OWA technology allows for a parameter that meaningfully captures complementarity even when the actors take the same level of care. The isoquant for this technology is piecewise linear and has a kink when the actors take the same level of care. Appendix B shows that the linearity assumption can be relaxed without altering the results. Non-differentiability is a more important feature. Yet the insights obtained with this simple technology extend without much loss of generality. Let us explain.

Take, for instance, the smoother CES technology. When the CES function represents perfect complements ($\rho \to \infty$), we have the same result as derived in section 3; there is a region of pooling possibly coupled with tailoring for the highest cost types. Moving away from perfect complements, with the CES technology, the second-best care schedule becomes strictly decreasing, and the pooling result formally disappears. However, it is easily shown that the second-best schedule $x(c; \rho)$ is continuous in the CES parameter ρ, and so with imperfect complements, the second-best schedule will be relatively flat for low levels of cost and relatively steep for high levels of cost.

Now suppose the planner paid a small price for imposing differential standards on actors with costs that were very close together. The efficiency benefits of making those distinctions for actors with costs below $\hat{c}$ are tiny because the optimal schedule has a fairly flat slope. In the excuse region, by contrast, the efficiency benefits from making distinctions among types are much higher because the care schedule dictates that the law induces much different care decisions for each type (the schedule's slope is quite negative). Thus, the non-differentiable schedules associated with our model mirror the schedules in second-best with the CES function under the assumption that the planner pays a small cost to make fine distinctions between agent types.

5.B. Other Applications

We motivated the model through a discussion of accident law. By straightforward re-characterizations of the problem, the insights can be applied in other contexts.

Take a contract between a buyer and seller. The buyer has a valuation the seller does not know. The seller has a cost the buyer does not know. The returns on the seller's investment turn on the buyer's valuation and vice versa. Both the buyer and seller make investment choices that increase the gains from trade. The parties write a contract to maximize the gains from trade. We formalize this model, and provide results, in an Online Appendix.

Suppose the court can observe the seller's cost and the buyer's valuation ex-post. Should the contract dictate tailored or objective standards for the investment? That is to say, should the contract hold the buyer with a high valuation or the seller with a low cost to higher investment markers?

The same trade-off arises. The seller with low costs effectively wastes her hefty investment if the buyer realizes a low valuation. Likewise, a buyer with a high valuation wastes her significant effort if she is matched with a seller with high cost of effort. And so, the contract should mandate objective rather than tailored standards for low-cost sellers and high valuations buyers. Further, the contract shouldexcuse skimpy investment choices by a buyer with a low valuation or a seller with a high cost.

Contract law reflects these ideas. Consider two examples of default rules, rules that are designed to maximize the gains from trade when the parties leave gaps in a contract.

5.B.1. Good Faith and Formation

Under the Uniform Commercial Code, every contract includes an obligation of good faith in its enforcement and performance.[21] Good faith means "honesty in fact" and "observations of reasonable commercial standards of fair dealing."[22] The good faith standard takes into account the particularities of the transaction and what happened between the parties. In this way, good faith inquiries are finely tailored to the transaction at issue. Translated into the model, a court, in filling out what good faith means, would likely consider the seller's actual cost and the buyer's actual valuation.

Notably, the good faith obligation does not apply at formation. At formation, the court asks whether a "reasonable" party in the position of, say, the offeree would conclude an offer has been made.[23] Painting with a broad brush, we might say contract law "pools" seller cost and

[21] U.C.C. 1-304.

[22] U.C.C. 1-201.

[23] In re JGB Indus., Inc., 223 B.R. 901, 907 (Bankr. E.D. Va. 1997)("The content of the offer will be construed from the standpoint of a reasonable person in accordance with the objective theory of contract formation.").

buyer valuation types during formation through the reasonableness inquiry but does not do so in assessing performance or enforcement obligations. Why the discord?

Our model explains the distinction. During formation, the parties do not know much about each other. In this context, reasonableness induces coordination of efforts and the avoidance of waste. It assures that both parties adequately invest—but do not over-invest—in communicating their needs and desires. After the parties have formed a contract, the parties spend time together. In the course of doing so, the buyer learns something about the seller's costs and the seller learns something about the buyer's valuation. As the information asymmetry vanishes, the parties can root the seller's obligations in the seller's actual cost and the buyer's actual valuation (which the seller now knows). Doing so induces an efficient (first-best) level of investment. The good faith standard effectuates that goal.

5.B.2. Conditions Involving Taste, Judgment and Fancy

Contracting parties often condition performance on the triggering of uncertain events. A buyer, say, might condition his obligation to buy a piece of property on his ability to obtain financing. The buyer then might have to make some investment in securing financing. But how much is required under the contract? How hard must the buyer try to find a lender? Alternatively, the owner of a restaurant might condition its obligation to have an ongoing relationship with a live music band on its satisfaction with the band's performance. How much effort must the owner spend working and promoting the band? Can the owner fire the band and then immediately hire back three of its members, leaving the lead singer without a gig?[24]

On the one hand, when the condition involves "taste, fancy, or judgment," the law applies a more personalized good faith standard. The question is whether the specific restaurant owner was satisfied with the band or the specific buyer found obtaining financing difficult. On the other hand, when the condition involves satisfaction as to mechanical fitness, utility, or marketability of the goods, the law applies the reasonable person standard. For example, suppose the contract called for delivery of an automatic rug cleaning machine.[25] In deciding whether the sale should be consummated, the court inquires whether a reasonable person would conclude the rug cleaning machine did not work as promised.

The cases display a pattern: In "taste, judgment, and fancy" cases, the courts apply a more tailored standard than in cases involving satisfaction as to mechanical performance. Like

[24]See Ferris v. Polansky, 191 Md. 79, 82, 59 A.2d 749, 750 (1948).
[25]See McKendrie v. Noel, 146 Colo. 440, 441, 362 P.2d 880, 881 (1961).

with formation versus enforcement, there is a discord in the law tailoring obligations. Why? The model provides one possible answer.

According to Lemma 1 above, a mean preserving spread of the buyer's valuation should result in a law that is more tailored. To the extent that transactions rooted in "taste, fancy, or judgment" involve a larger spread, the common law's reliance on reasonableness for conditions involving mechanical utility and shying away from reasonableness in taste, judgment, and fancy cases makes some sense.

6 Concluding Remarks and Discussion

In conclusion, the debate about the law's reliance on the reasonable person standard has been simmering for years. On one side sits the philosophers. They claim that the reasonable person standard embeds into the law notions of reciprocity, fairness, and the proper expectations of the behavior of others. On the other side sits the law and economics scholars. They argue that the reasonable person standard arises because courts find it expensive to measure cost on an individual basis. The court, then, avoids these costs by treating everyone the same: as a person with some average cost of avoiding accidents.

The two camps largely talk past each other. This model provides economic content to the philosopher's position. It explains (a) when expectations about the behavior of others matters and (b) how a legal rule consisting of an objective standard with releases for the least able in the population leads to a proper construction of those expectations and a superior allocation of resources.

Notably, the expectations concern strikes hardest where care decisions are strong complements. In those cases, the adverse selection problem is most acute, and so is the need for the antidote of the objective standard. The urge to be "reasonable" arises as a way to facilitate coordination and avoid waste. The cost is a failure to fully utilize the talents of high-ability actors.

From a modeling standpoint, the pooling result resembles the "ironing" technique from mechanism design (Baron & Myerson, 1982). Yet the need to iron in this model arises from a different economic challenge for the planner. In the mechanism design problem, incentive compatibility demands that say, the quantity offered (weakly) increases with the consumer's taste for the good. If the hazard rate for the distribution of consumer types is increasing, the schedule separating the types satisfies the monotonicity requirement. If the hazard rate does not have this familiar property, the seller must iron and offer some interval of types the same quantity of the good. Ironing arises when the distribution of consumer types is "unusual" and can happen anywhere along the support.

In our model, the planner 'irons' to coordinate two agents who are uninformed about each other. Because the motivation differs, we find that ironing becomes the rule, not the exception (i.e., it happens always, not just when the distribution lacks an increasing hazard rate). More, the model produces a clean prediction about the location in the distribution where the ironing occurs—always with respect to the lowest cost types.

Finally, the model made a number of assumptions. Most importantly, we assumed that the parties interacted only one time. With repeated interactions (like with the buyer and seller in a relational contract), the actor would learn about each other's costs over time. The law would then need to be responsive to this learning dynamic. We hope to consider this extension in future work.

References

Avilés-Ochoa, E., León-Castro, E., Perez-Arellano, L. A., & Merigó, J. M. (2018). Government transparency measurement through prioritized distance operators. *Journal of Intelligent & Fuzzy Systems*, *34*(4), 2783–2794.

Baron, D. P., & Myerson, R. B. (1982). Regulating a monopolist with unknown costs. *Econometrica*, *50*(4), 911–930.

Bender, L. (1988). Lawyer's primer on feminist theory and tort, a women in legal education–pedagogy, law, theory, and practice. *Journal of Legal Education*, *38*(1), 3–38.

Ben-Shahar, O., & Porat, A. (2016). Personalizing negligence law. *New York University Law Review*, *91*(3), 627–688.

Bernheim, B. D., & Whinston, M. D. (1998). Incomplete contracts and strategic ambiguity. *American Economic Review*, *88*(4), 902-932.

Bernstein, A. (2001). The communities that make standards of care possible symposium on negligence in the courts: The actual practice. *Chicago-Kent Law Review*, *77*(2), 735–772.

Blanco-Mesa, F., León-Castro, E., & Merigó, J. M. (2018). Bonferroni induced heavy operators in erm decision-making: A case on large companies in colombia. *Applied Soft Computing*, *72*, 371–391.

Brown, J. P. (1973). Toward an economic theory of liability. *Journal of Legal Studies*, *2*, 323–349.

Calabresi, G., & Hirschoff, J. T. (1971). Toward a test for strict liability in torts. *Yale Law Journal*, *81*(6), 1055–1085.

Cremer, J. (1995). Arm's length relationships. *Quarterly Journal of Economics*, *110*(2), 275-295.

Dark, O. C. (2004). Tort liability and the unquiet mind: A proposal to incorporate mental disabilities into the standard of care. *Thurgood Marshall Law Review*, *30*(1), 169–214.

Demsetz, H. (1972). When does the rule of liability matter? *Journal of Legal Studies*, *1*, 13-28.

Dobbs, D., Hayden, P., & Bublick, E. (2015). *Hornbook on torts* (2nd ed.). West Academic Publishing.

Doña, J. M., la Red, D. L., & Peláez, J. I. (2009). A mix model of discounted cash-flow and owa operators for strategic valuation. *International Journal of Interactive Multimedia and Artificial Intelligence*, *1*(2).

Eggen, J. M. (2015). Mental disabilities and duty in negligence law: Will neuroscience reform tort doctrine medical myths: Exploring effectiveness, misinformation and scientific rigor. *Indiana Health Law Review*, *12*(2), 591–650.

Emmons, W., & Sobel, J. (1991). On the effectiveness of liability rules when agents are not identical. *Review of Economic Studies*, *58*(2), 375-390.

Garoupa, N., & Dari-Mattiacci, G. (2007). Least cost avoidance: The tragedy of common safety. *Journal of Law, Economics, & Organization*, *25*(1), 235-261.

Holmes, O., Jr. (1881). *The common law*. Dover Publications.

Hurwicz, L. (1951). The generalized bayes minimax principle: a criterion for decision making under uncertainty. *Cowles Comm. Discuss. Paper Stat*, *335*, 1950.

Keating, G. C. (1995). Reasonableness and rationality in negligence theory. *Stanford Law Review*, *48*(2), 311–384.

Landes, W. M., & Posner, R. A. (1987). *The economic structure of tort law* (1st ed.). Harvard University Press.

León-Castro, E., Avilés-Ochoa, E., & Gil Lafuente, A. M. (2016). Exchange rate usd/mxn forecast through econometric models, time series and howma operators. *Economic Computation and Economic Cybernetics Studies and Research*, *50*.

León-Castro, E., Espinoza-Audelo, L. F., Merigó, J. M., Gil-Lafuente, A. M., & Yager, R. R. (2020). The ordered weighted average inflation. *Journal of Intelligent & Fuzzy Systems*, *38*(2), 1901–1913.

Lindquist, G. (2020). Science or status quo? disregard for a defendant's mental illness in tort suits. *Washington Law Review Online*, *95*, 115–167.

MSBNA Standing Committee on Pattern Jury Instructions. (2009). *Maryland civil pattern jury instructions* (4th ed. ed.). Md Institute for the Continuing Prof Ed of Lawyers, Inc.

Posner, R. A. (2014). *Economic analysis of law* (9th ed.). Aspen Publishing.

Quade, P. (2023). Council post: Securing your supply chain: Where is the weakest link? *Forbes*. (https://www.forbes.com/sites/forbestechcouncil/2019/01/30/securing-your-supply-chain-where-is-the-weakest-link/)

Rothschild, M., & Stiglitz, J. E. (1970). Increasing risk: I. a definition. *Journal of Economic theory*, *2*(3), 225–243.

Sappington, D. E. M. (1986). Commitment to regulatory bureaucracy. *Information Economics and Policy*, *2*(4), 243–258.

Schelling, T. C. (1980). *The strategy of conflict* (Rev. ed. ed.). Harvard University Press.

Shavell, S. (1987). *Economic analysis of accident law.* Harvard University Press.

Shih, W. (2023). Global supply chains are only as strong as their weakest link, and that matters during the coronavirus crisis. *Forbes*.

Taylor, C. R., & Yildirim, H. (2011). Subjective performance and the value of blind evaluation. *Review of Economic Studies*, *78*(2), 762-794. Retrieved from `https://www.jstor.org/stable/23015873`

Tobia, K. P. (2018). How people judge what is reasonable. *Alabama Law Review*, *70*(2), 293-360.

Xiong, W., & Liu, H. (2014). On solving some paradoxes using the ordered weighted averaging operator based decision model. *International journal of intelligent systems*, *29*(1), 1–25.

Yager, R. R. (2002). On the valuation of alternatives for decision-making under uncertainty. *International Journal of Intelligent Systems*, *17*(7), 687–707.

Yager, R. R. (2004). Uncertainty modeling and decision support. *Reliability Engineering & System Safety*, *85*(1-3), 341–354.

Zipursky, B. C. (2015). Reasonableness in and our of negligence law. *University of Pennsylvania Law Review*, *163*(7), 2131–2170.

Appendices

A Proofs

Proof of Proposition 1. Without loss of generality, suppose $c_m < c_p$. Then, it must be that $x_m \geq x_p$. (To see this, note that if $x_m < x_p$, we can achieve the same average cost of care — and hence the same probability of harm — by reversing the care levels, but at lower total cost.)

Suppose $x_m > x_p$. Then, the planner's problem is:

$$W = \min_{x_m, x_p} \Pi(\lambda x_m + (1-\lambda)x_p) + c_m x_m + c_p x_p$$

The first order conditions (FOCs) are:

$$\begin{aligned} \frac{\partial W}{\partial x_m} &= \lambda \Pi'(\lambda x_m + (1-\lambda)x_p) + c_m \geq 0 \\ \frac{\partial W}{\partial x_p} &= (1-\lambda)\Pi'(\lambda x_m + (1-\lambda)x_p) + c_p \geq 0 \end{aligned}$$

Moreover, if $x_i > 0$, then $\frac{\partial W}{\partial x_i} = 0$, and if $\frac{\partial W}{\partial x_i} > 0$, then $x_i = 0$.

Since $x_m > x_p$ by assumption, then $x_m > 0$, and so $\frac{\partial W}{\partial x_m} = 0$. Notice that the Π' term is common to both FOCs. This means that, except for knife-edge cases, it cannot be that both FOCs hold to zero simultaneously. Hence, we must have $\frac{\partial W}{\partial x_m} = 0 < \frac{\partial W}{\partial x_p}$ and so $x_p = 0$. Since the first equation holds with equality, we have:

$$\begin{aligned} \lambda \Pi'(\lambda x_m) &= -c_m \\ x_m &= \frac{1}{\lambda}[\Pi']^{-1}\left(-\frac{c_m}{\lambda}\right) = \frac{1}{\lambda} z\left(\frac{c_m}{\lambda}\right) \end{aligned}$$

Next, substituting out for the common term, the second equation will be positive provided that:

$$\begin{aligned} c_p + (1-\lambda)\left(-\frac{c_m}{\lambda}\right) &> 0 \\ \lambda &> \frac{c_m}{c_m + c_p} \end{aligned}$$

If this condition is not met (i.e. if $\lambda \leq \frac{c_m}{c_m+c_p}$) then asserting $x_m > x_p$ gives a contradiction. Hence, it must be that $x_m = x_p$. The constrained problem becomes:

$$\min_x (c_m + c_p)x + \Pi(x)$$

Straightforwardly, we have $x_m = x_p = [\Pi']^{-1}(-(c_m + c_p)) = z(c_m + c_p)$. $\square$

Proofs of Propositions 2 and 4, and Corollary 1. Proposition 2 is simply a special case of Proposition 4. Thus, in this section, we prove the latter proposition.

The court's problem is to choose functions $x_m(c_m)$ and $x_p(c_p)$ to minimize the expected social loss:

$$W = \iint_{c_m,c_p} [\Pi(\min\{x_m(c_m), x_p(c_p)\}) + c_m x_m(c_m) + c_p x_p(c_p)]\, g_m(c_m) g_p(c_p) dc_m dc_p$$

Taking the derivative w.r.t. $x_i(c_i)$ gives:

$$\begin{aligned} \frac{\partial W}{\partial x_i(c_i)} &= c_i + \Pi'(x_i(c_i)) \int_{c_{-i}} \mathbf{1}[x_i(c_i) < x_{-i}(c_{-i})] g_{-i}(c_{-i}) dc_{-i} \\ &= c_i + \Pi'(x_i(c_i)) \Pr[x_{-i}(c_{-i}) > x_i(c_i)] \end{aligned}$$

First, we make a technical point about first order conditions (FOCs). The FOCs characterize the optimum wherever the first derivative is continuous (in x_i). Notice that this will be true whenever $\Pr(x_{-i}(c_i) > x_i)$ is also continuous. Since the distribution of c's is itself continuous, the probability function will be continuous except at values of x at which there is (partial)-pooling. Moreover, at these points of discontinuity, there may be a range of c_i's for which the first order condition cannot be satisfied (because $\frac{\partial W}{\partial x_i}(x) < 0$ but $\lim_{x' \uparrow x} \frac{\partial W}{\partial x_i}(x') > 0$). Naturally, for these c_i's, the planner does best to pool on x as well. But, since the first order condition is not met, changes in x will have first order effects on social welfare. Hence, we must additionally consider a joint deviation where both m and p types switch from x to some other pooling level x'. (When the first order condition holds exactly, this isn't necessary, since the benefits of any such deviation will be second order.)

We now begin the proof proper. The proof is in many steps. We proceed in the following order: (1) We show that the second-best schedules must be continuous and weakly decreasing; (2) we characterize the schedule whenever it is strictly decreasing; (3) we show that the schedule must be constant when c_i lies below some threshold $\hat{c}_i$; (4) we characterize that threshold as well as the pooling care level; (5) we show that the thresholds are unique; and (6) we show that the schedule must be decreasing beyond this threshold. Taken together, these steps prove the claims.

First, since the objective function is strictly concave, the optimizers must be singleton-valued. Furthermore, since the objective function is continuous, and the optimization is over a compact set, the optimizers $x_i(c_i)$ must themselves be continuous, by Berge's Theorem of the maximum. Moreover, the second-best functions $x_i(c_i)$ must be weakly decreasing in c_i. To see this, note that for any schedule $x_i(c_i)$ that is strictly increasing over some interval, we can construct an alternative schedule $y_i(c_i)$ that is strictly decreasing, and generates the same marginal distribution over care levels. The alternative schedule y_i produces the amount of care and the same likelihood of harms as x_i, but assigns the higher care levels to agents with lower costs. This clearly reduces the social loss.

Second, we show that whenever the second-best schedule is strictly decreasing, it is characterized by $x_i(c_i) = z\left(\frac{c_i}{G_{-i}(c_{-i}(c_i))}\right)$. To see this, note by the above logic that the first order conditions must be satisfied in this case. Thus, we have:

$$\Pi'(x_i(c_i))\,\Pr[x_{-i}(c_{-i}) > x_i(c_i)] = -c_i$$

Moreover, if $x_i(c_i)$ is decreasing in a neighborhood where care level x is chosen, then $x_{-i}(c_i)$ must also be decreasing in a corresponding neighborhood where that same care level is taken. (If not, the pooling by one side would cause pooling by the other side.) Hence, $x_i(c_i)$ and $x_{-i}(c_{-i})$ must both be locally invertible in this region. Let $c_{-i}(c_i) = x_{-i}^{-1}(x_i(c_i))$. Then, the first order condition becomes:

$$\Pi'(x_i(c_i))\,\Pr[c_{-i} < c_{-i}(c_i)] = -c_i$$
$$x_i(c_i) = [\Pi']^{-1}\left(-\frac{c_i}{G_{-i}(c_{-i}(c_i))}\right) = z\left(\frac{c_i}{G_{-i}(c_{-i}(c_i))}\right)$$

Now, suppose $c_m = c_m(c_p)$. It follows that $c_p = c_p(c_m)$. Then, since z is strictly decreasing, and since $x_m(c_m) = x_p(c_p)$ (by construction):

$$
\begin{aligned}
z\left(\frac{c_m}{G_p(c_p(c_m))}\right) &= z\left(\frac{c_p}{G_m(c_m(c_p))}\right) \\
\frac{c_m}{G_p(c_p(c_m))} &= \frac{c_p}{G_m(c_m(c_p))} \\
c_m G_m(c_m) &= c_p G_p(c_p)
\end{aligned}
$$

This expression implicitly defines the functions $c_{-i}(c_i)$. Moreover, by the implicit function theorem:

$$
\frac{\partial c_{-i}}{\partial c_i} = \frac{G_i(c_i) + c_i g_i(c_i)}{G_{-i}(c_{-i}(c_i)) + c_{-i}(c_i) g_{-i}(c_{-i}(c_i))}
$$

We have an explicit characterization of the second-best schedule whenever it is strictly decreasing. We must confirm that this schedule is indeed strictly decreasing in costs. Since $z'(\cdot) < 0$, it suffices to show that $\frac{c_i}{G_{-i}(c_{-i}(c_i)}$ is a strictly increasing. Differentiating gives:

$$
\frac{\partial}{\partial c_i}\left(\frac{c_i}{G_{-i}(c_{-i}(c_i))}\right) = \frac{1}{G_{-i}(c_{-i}) + c_{-i} g_{-i}(c_{-i})}\left[1 - \frac{c_i g_i(c_i)}{G_i(c_i)} \cdot \frac{c_{-i} g_{-i}(c_{-i})}{G_{-i}(c_{-i})}\right]
$$

where we occasionally suppress the dependence of c_{-i} on c_i, and repeatedly use the fact that $c_i G_i(c_i) = c_{-i} G_{-i}(c_{-i})$. Hence, to have a decreasing second-best schedule, it suffices that $\frac{c_i g_i(c_i)}{G_i(c_i)} \cdot \frac{c_{-i} g_{-i}(c_{-i})}{G_{-i}(c_{-i})} < 1$ — a property that we establish, below.

Third, we show that there must exist $\hat{c}_m > \underline{c}_m$ and $\hat{c}_p > \underline{c}_p$ s.t. $x_m(c_m) = \hat{x} = x_p(c_p)$ for all $c_i < \hat{c}_i$. Suppose not. I.e. suppose there exists $\varepsilon > 0$ s.t. $x_p(c_p)$ is strictly decreasing on the interval $[\underline{c}_p, \underline{c}_p + \varepsilon]$. We know that neither agent-type will take a care level that they know (for sure) will be larger than their opponent's. Hence, since the x's are weakly decreasing, it must be that $x_m(\underline{c}_m) = \overline{x} = x_p(\underline{c}_p)$. Consider now types close to the lower bound. Since $x_p(c_p)$ is strictly decreasing on $[\underline{c}_p, \underline{c}_p + \varepsilon]$, it must be that $\Pr[x_p(c_p) > x_m]$ is continuous for $x_m \in [x_p(\underline{c}_p + \varepsilon), \overline{x})$. Hence, there exists $\delta(\varepsilon)$ s.t. $x_m(c_m)$ is characterized by the FOCs for $c_m \in [\underline{c}_m, \underline{c}_m + \delta]$. Hence, over this interval, $x_m(c_m) = z\left(\frac{c_m}{\Pr[x_p(c_p) > x_m(c_m)]}\right)$. But, $\lim_{x_m \uparrow \overline{x}} \Pr[x_p(c_p) \geq x_m] = 0$, and so $x_m(\underline{c}_m) = 0$ (by the Inada conditions). Hence $\overline{x} = 0$, and since $0 \leq x_i(c_i) \leq \overline{x}$ for each i, it must be that $x_i(c_i) = 0$ for all i. But this contradicts the assumption that x_p was strictly decreasing on the interval $[\underline{c}_p, \underline{c}_p + \varepsilon]$. Hence, it must be that both x_m and x_p are constant for $c_i \leq \hat{c}_i$.

Fourth, we characterize the pooling care level and the threshold defining the pool. Let $x_i(c_i) = \hat{x}$ for $c_i \leq \hat{c}_i$. By construction, there exists some $\varepsilon > 0$ s.t. $x_i(c_i)$ is strictly decreasing on the interval $(\hat{c}_i, \hat{c}_i + \varepsilon)$. Hence, on this interval, the care levels are characterized by the FOCs. Moreover, since $x_i(c_i)$ is continuous, it must be that the FOC is satisfied at $\hat{c}_i$ (for each i). It follows that:

$$z\left(\frac{\hat{c}_m}{G_p(\hat{c}_p)}\right) = \hat{x} = z\left(\frac{\hat{c}_p}{G_m(\hat{c}_m)}\right)$$

which implies that $\hat{c}_m G_m(\hat{c}_m) = \hat{c}_p G_p(\hat{c}_p)$. Hence, the thresholds satisfy $\hat{c}_m = c_m(\hat{c}_p)$.

Now, noting that $\hat{x}$ implicitly pins down $\hat{c}_m$ and $\hat{c}_p$, $\hat{x}$ is chosen to minimize the social loss:

$$\begin{aligned} W(\hat{x}) = &\int_{\underline{c}_m}^{\hat{c}_m(\hat{x})} \hat{x} c_m g_m(c_p) dc_m + \int_{\underline{c}_p}^{\hat{c}_p(\hat{x})} \hat{x} c_p g_p(c_p) dc_p + G_m(\hat{c}_m) G_p(\hat{c}_p)\Pi(\hat{x}) + \\ &+ \int_{\hat{c}_m(\hat{x})}^{\bar{c}_m} x_m(c_m) c_m g_m(c_P) dc_m + \int_{\hat{c}_p(\hat{x})}^{\bar{c}_p} x_p(c_p) c_p g_p(c_p) dc_p + \\ &\quad + \int_{\hat{c}_m(\hat{x})}^{\bar{c}_m} G_p(c_p(c_m))\Pi(x_m(c_m)) g_m(c_m) dc_m + \int_{\hat{c}_p(\hat{x})}^{\bar{c}_p} G_m(c_m(c_p))\Pi(x_p(c_p)) g_p(c_p) dc_p \end{aligned}$$

Taking the first order condition, and noting that all indirect effects through $\hat{c}_m$ and $\hat{c}_p$ cancel, we have:

$$\begin{gathered} G_m(\hat{c}_m) E[c_m \,|\, c_m < \hat{c}_m] + G_p(\hat{c}_p) E[c_p \,|\, c_p < \hat{c}_p] = -G_m(\hat{c}_m) G_p(\hat{c}_p) \Pi'(\hat{x}) \\ \frac{E[c_m \,|\, c_m < \hat{c}_m]}{\hat{c}_m} + \frac{E[c_p \,|\, c_p < \hat{c}_p]}{\hat{c}_p} = 1 \end{gathered}$$

where we use the fact that $\hat{c}_m G_m(\hat{c}_m) = \hat{c}_p G_p(\hat{c}_p)$ and that $\Pi'(\hat{x}) = -\frac{\hat{c}_p}{G_m(\hat{c}_m)}$. Thus, we have the conditions that characterize the thresholds $\hat{c}_m$ and $\hat{c}_p$, and the pooling level $\hat{x}$.

Fifth, we must show that the thresholds are unique. As a preliminary step, note that:

$$\frac{\partial}{\partial c_i}\left(\frac{E[c_i \,|\, c_i < \hat{c}_i]}{\hat{c}_i}\right) = \frac{1}{\hat{c}_i}\left[\frac{\hat{c}_i g_i(\hat{c}_i)}{G_i(\hat{c}_i)} - \left(1 + \frac{\hat{c}_i g_i(\hat{c}_i)}{G_i(\hat{c}_i)}\right)\frac{E[c_i \,|\, c_i < \hat{c}_i]}{\hat{c}_i}\right]$$

Recall also that:

$$\frac{\partial c_p(c_m)}{\partial c_m} = \frac{G_m(c_m)}{G_p(c_p(c_m))} \cdot \frac{1 + \frac{c_m g_m(c_m)}{G_m(c_m)}}{1 + \frac{c_p(c_m) g_p(c_p(c_m))}{G_p(c_p(c_m))}}$$

Now, define:

$$\phi(\tilde{c}_m) = \frac{E[c_m \,|\, c_m < \tilde{c}_m]}{\tilde{c}_m} + \frac{E[c_p \,|\, c_p < c_p(\tilde{c}_m)]}{c_p(\tilde{c}_m)} - 1$$

We know that $\phi(\hat{c}_m) = 0$. To prove uniqueness, it suffices to show that $\phi(\cdot)$ has a unique root. Notice that $\phi(\underline{c}_m) = \frac{\underline{c}_m}{\underline{c}_m} + \frac{\underline{c}_p}{\underline{c}_p} - 1 = 1 > 0$, which makes use of the fact that $c_p(\underline{c}_m) = \underline{c}_p$. Also, $\lim_{\tilde{c}_m \to \infty} \phi(\tilde{c}_m) = -1$. Then, since ϕ is a continuous function, there must be at least one $\tilde{c}_m \in (\underline{c}_m, \infty)$ s.t. $\phi(\tilde{c}_m) = 0$. Let $\hat{c}_m$ be the first such instance. Since $\phi(\tilde{c}_m) > 0$ for $\tilde{c}_m < \hat{c}_m$, we must have that $\phi'(\hat{c}_m) < 0$.

In what follows, we write $\tilde{c}_p = c_p(\tilde{c}_m)$ and we suppress this dependence in the notation for convenience. Now:

$$\begin{aligned}
\phi'(\tilde{c}_m) &= \frac{1}{\tilde{c}_m}\left[\frac{\tilde{c}_m g_m(\tilde{c}_m)}{G_m(\tilde{c}_m)} - \left(1 + \frac{\tilde{c}_m g_m(\tilde{c}_m)}{G_m(\tilde{c}_m)}\right)\frac{E[c_m \,|\, c_m < \tilde{c}_m]}{\hat{c}_m}\right] \\
&\quad + \frac{1}{\tilde{c}_p}\left[\frac{\tilde{c}_p g_p(\tilde{c}_p)}{G_p(\tilde{c}_p)} - \left(1 + \frac{\tilde{c}_p g_p(\tilde{c}_p)}{G_p(\tilde{c}_p)}\right)\frac{E[c_p \,|\, c_p < \tilde{c}_p]}{\hat{c}_p}\right]\left(\frac{G_m(\tilde{c}_m)}{G_p(\tilde{c}_p)} \cdot \frac{1 + \frac{\tilde{c}_m g_m(\tilde{c}_m)}{G_m(\tilde{c}_m)}}{1 + \frac{\tilde{c}_p g_p(\tilde{c}_p)}{G_p(\tilde{c}_p)}}\right) \\
&= \frac{1}{\tilde{c}_m}\left[\frac{\tilde{c}_m g_m(\tilde{c}_m)}{G_m(\tilde{c}_m)} + \left(1 + \frac{\tilde{c}_m g_m(\tilde{c}_m)}{G_m(\tilde{c}_m)}\right)\left\{\frac{\frac{\tilde{c}_p g_p(\tilde{c}_p)}{G_p(\tilde{c}_p)}}{1 + \frac{\tilde{c}_p g_p(\tilde{c}_p)}{G_p(\tilde{c}_p)}} - \left(\frac{E[c_m \,|\, c_m < \tilde{c}_m]}{\hat{c}_m} + \frac{E[c_p \,|\, c_p < \tilde{c}_p]}{\hat{c}_p}\right)\right\}\right]
\end{aligned}$$

where we use the fact that $\tilde{c}_m = \tilde{c}_p \cdot \frac{G_p(\tilde{c}_p)}{G_m(\tilde{c}_m)}$. Evaluating this at $\tilde{c}_m = \hat{c}_m$ gives:

$$\begin{aligned}
\phi'(\hat{c}_m) &= \frac{1}{\hat{c}_m}\left[\frac{\hat{c}_m g_m(\hat{c}_m)}{G_m(\hat{c}_m)} + \left(1 + \frac{\hat{c}_m g_m(\hat{c}_m)}{G_m(\hat{c}_m)}\right)\left\{\frac{\frac{\hat{c}_p g_p(\hat{c}_p)}{G_p(\hat{c}_p)}}{1 + \frac{\hat{c}_p g_p(\hat{c}_p)}{G_p(\hat{c}_p)}} - 1\right\}\right] \\
&= \frac{1}{\hat{c}_m}\left[\frac{\hat{c}_p g_p(\hat{c}_p)}{G_p(\hat{c}_p)} \frac{1 + \frac{\hat{c}_m g_m(\hat{c}_m)}{G_m(\hat{c}_m)}}{1 + \frac{\hat{c}_p g_p(\hat{c}_p)}{G_p(\hat{c}_p)}} - 1\right] \\
&= \frac{1}{\hat{c}_m} \cdot \frac{1}{1 + \frac{\hat{c}_p g_p(\hat{c}_p)}{G_p(\hat{c}_p)}}\left[\frac{\hat{c}_m g_m(\hat{c}_m)}{G_m(\hat{c}_m)} \cdot \frac{\hat{c}_p g_p(\hat{c}_p)}{G_p(\hat{c}_p)} - 1\right]
\end{aligned}$$

Since, $\phi'(\hat{c}_m) < 0$, it follows that $\frac{\hat{c}_m g_m(\hat{c}_m)}{G_m(\hat{c}_m)} \cdot \frac{\hat{c}_p g_p(\hat{c}_p)}{G_p(\hat{c}_p)} < 1$. By the argument above, this ensures that $x_i(c_i)$ is strictly decreasing for c_i slightly above $\hat{c}_i$.

We need to show that $\hat{c}_m$ is the unique root of ϕ. Suppose not. Let $c_m^\dagger > \hat{c}_m$ denote the next smallest root. Then $\phi(c_m^\dagger)$. Moreover, since ϕ is continuous and $\phi(c_m) < 0$ for $c_m \in (\hat{c}_m, c_m^\dagger)$, it must be that $\phi'(c_m^\dagger) > 0$. Hence, by the previous argument: $\frac{c_m^\dagger g_m(c_m^\dagger)}{G_m(c_m^\dagger)} \cdot \frac{c_p^\dagger g_p(c_p^\dagger)}{G_p(c_p^\dagger)} > 1$, where $c_p^\dagger = c_p(c_m^\dagger)$. We show that this cannot be. It suffices to show that $\frac{c_i g_i(c_i)}{G_i(c_i)}$ is weakly decreasing in c_i. If so, then since $\frac{\hat{c}_m g_m(\hat{c}_m)}{G_m(\hat{c}_m)} \cdot \frac{\hat{c}_p g_p(\hat{c}_p)}{G_p(\hat{c}_p)} < 1$, it must be that $\frac{c_m g_m(c_m)}{G_m(c_m)} \cdot \frac{c_p g_p(c_p)}{G_p(c_p)} < 1$ for all pairs (c_m, c_p) s.t. $c_m > \hat{c}_m$ and $c_p > \hat{c}$.

The fact that $\frac{c_i g_i(c_i)}{G_i(c_i)}$ is decreasing in c_i follows from the assumption that $G_i(c_i)$ was sufficiently log-concave. Notice that:

$$\frac{\partial}{\partial c_i}\left(\frac{c_i g_i(c_i)}{G_i(c_i)}\right) = \frac{\partial}{\partial c_i}\left(c_i \frac{\partial \ln G_i(c_i)}{\partial c_i}\right) = c_i \frac{\partial^2 \ln G_i(c_i)}{\partial c_i^2} + \frac{\partial \ln G_i(c_i)}{\partial c_i} = \frac{\partial \ln G_i(c_i)}{\partial c_i} \cdot [1 - r_i(c_i)]$$

where $r_i(c_i) = -c_i \cdot \frac{\partial^2 \ln G_i(c_i) / \partial c_i^2}{\partial \ln G_i(c_i) / \partial c_i}$ is the analogue of the coefficient of relative risk aversion. Since G_i is weakly increasing on its support, and since $r_i(c_i) > 1$ for all c_i, it follows that $\frac{\partial}{\partial c_i}\left(\frac{c_i g_i(c_i)}{G_i(c_i)}\right) \leq 0$.

Sixth, we must show that $x_i(c_i)$ is strictly decreasing for all $c_i > \hat{c}_i$. One way to see this is to note that since $\frac{c_m g_m(c_m)}{G_m(c_m)} \cdot \frac{c_p(c_m) g_p(c_p(c_m))}{G_p(c_p(c_m))} < 1$ for all $c_m > \hat{c}_m$, that $x_m(c_m)$ is strictly decreasing $c_m > \hat{c}_m$, and likewise for c_p. (This follows from part two of the proof.)

But more strongly, assume the opposite and suppose there is an interval $[c_i', c_i'']$ with $c_i' > \hat{c}_i$, s.t. $x_i(c_i)$ is constant on $(c_i', c_i'']$. Then, it must be that the x_i chosen on this interval is optimal for the average cost type in the pool (similar to how $\hat{x}$ was computed above). Given that $\frac{c_m g_m(c_m)}{G_m(c_m)} \cdot \frac{c_p(c_m) g_p(c_p(c_m))}{G_p(c_p(c_m))} < 1$, every type in the interval, if separating themselves, would want to produce less than $x(c_i')$, and so on average, the pool must produce strictly less than $x(c_i')$. But then, necessarily, there will be a discontinuity in x_i at c_i', which cannot be. Hence, $x_i(c_i)$ is strictly decreasing for $c_i > \hat{c}_i$. This completes the proof. □

Proof of Lemma 1. First, we show that $\hat{c} \geq \overline{c}$ if $\overline{c} \leq 2E[c]$. To see this, recall that that $\hat{c} = 2E[c \,|\, c < \hat{c}]$. Suppose $\hat{c} \geq \overline{c}$. Then $E[c \,|\, c < \hat{c}] = E[c]$, and so $\hat{c} = 2E[c]$. Consistency requires that $2E[c] \geq \overline{c}$, as required.

Next, we verify the comparative statics. Consider two distributions of costs, $G(c)$ and $H(c)$, and let $\hat{c}_i$ satisfy $\hat{c}_i = 2E_i[c \,|\, c < \hat{c}_i]$ for $i \in \{G, H\}$. Begin with scaling — i.e. suppose $c_H = \kappa c_G$. Then $\kappa \hat{c}_G = 2E[\kappa c_G \,|\, \kappa_1 c_G < \kappa \hat{c}_G] = 2E[c_H \,|\, c_H < \kappa \hat{c}_G]$, and so $\hat{c}_H = \kappa \hat{c}_G$. Moreover, since $H(\kappa c) = G(c)$ for all c, we have: $H(\hat{c}_H) = H(\kappa \hat{c}_G) = G(\hat{c}_G)$.

Finally, suppose H is a mean preserving spread of G. Then, by the Rothschild and Stiglitz

(1970) condition, $\int^c H(t)dt \geq \int^c G(t)dt$ for all t. This implies that:

$$\begin{aligned}
\frac{\int^{\hat{c}_G} H(c)dc}{G(\hat{c}_G)} &\geq \frac{\int^{\hat{c}_G} G(c)dc}{G(\hat{c}_G)} \\
\frac{\int^{\hat{c}_G} H(c)dc}{H(\hat{c}_G)} &> \frac{\int^{\hat{c}_G} G(c)dc}{G(\hat{c}_G)} \\
\hat{c}_G - \frac{\int^{\hat{c}_G} H(c)dc}{H(\hat{c}_G)} &< \hat{c}_G - \frac{\int^{\hat{c}_G} G(c)dc}{G(\hat{c}_G)} \\
E_H[c \,|\, c < \hat{c}_G] &< E_G[c \,|\, c < \hat{c}_G] \\
\frac{E_H[c \,|\, c < \hat{c}_G]}{\hat{c}_G} &< \frac{E_G[c \,|\, c < \hat{c}_G]}{\hat{c}_G}
\end{aligned}$$

where the second inequality uses the fact that $H(\hat{c}_G) < G(\hat{c}_G)$. The fourth line uses the property that, for any function f with $f(a) = 0$, $\int_a^c xf(x)dx = c - \int_a^c F(x)dx$, where $F'(x) = f(x)$. Then, by expression (3) observe that $\frac{E_G[c \,|\, c<\hat{c}_G]}{\hat{c}_G} = \frac{1}{2} > \frac{E_H[c \,|\, c<\hat{c}_G]}{\hat{c}_G}$.

Also, by expression (3), $\frac{E_H[c \,|\, c<\hat{c}_H]}{\hat{c}_H} = \frac{1}{2}$, and since $\frac{E_H[c \,|\, c<\alpha]}{\alpha}$ is a decreasing function of α (which follows from the fact that the CDFs are sufficiently log-concave), it must be that $\hat{c}_H < \hat{c}_G$. Moreover, this implies that $H(\hat{c}_H) < H(\hat{c}_G) < G(\hat{c}_G)$. □

Proof of Proposition 3 and Corollary 3. Let $x(c; \lambda)$ denote the second-best schedule. (We will often omit the 2nd argument, for notational convenience.) We first show that there must be pooling for $\lambda > 0$ sufficiently small. Suppose not, i.e. suppose the second-best schedule $x(c)$ is purely separating, so that $x'(c) < 0$. Take agent i, and note by symmetry that $x_i(c_i) > x_{-i}(c_{-i})$ whenever $c_i < c_{-i}$. Since $x(c)$ satisfies the first order conditions, we have for agent i:

$$\frac{\partial W}{\partial x(c_i)} = c_i + (1-\lambda)\int_{\underline{c}}^{c_i} \Pi'(\lambda x(c_{-i}) + (1-\lambda)x(c_i))g(c_{-i})dc_{-i} + \lambda \int_{c_i}^{\overline{c}} \Pi'(\lambda x(c_i) + (1-\lambda)x(c_{-i}))g(c_{-i})dc_{-i} \tag{4}$$

Then, for $c_i = \underline{c}$, this reduces to:

$$\frac{\partial W}{\partial x(\underline{c})} = \underline{c} + \lambda \int_{\underline{c}}^{\overline{c}} \Pi'(\lambda x(\underline{c}) + (1-\lambda)x(c_{-i}))g(c_{-i})dc_{-i}$$

Since $\Pi'' > 0$, then $\Pi'(\lambda x(c_i) + (1-\lambda)x(c_{-i})) > \Pi((1-\lambda)x(c_{-i}))$ whenever $x(\underline{c}) > 0$. Hence:

$$\frac{\partial W}{\partial x(\underline{c})} > \underline{c} + \lambda \int_{\underline{c}}^{\overline{c}} \Pi'((1-\lambda)x(c_{-i}))g(c_{-i})dc_{-i}$$

Then, since $\underline{c} > 0$ and $\int_{\underline{c}}^{\overline{c}} \Pi'((1-\lambda)x(c_{-i}))g(c_{-i})dc_{-i} < 0$ is finite, there exists $\lambda' > 0$ s.t. $\frac{\partial W}{\partial x(\underline{c})} > 0$ for all choices of $x(\underline{c})$ whenever $\lambda < \lambda'$. It follows that $x(\underline{c}) = 0$. But, then since $x(c)$ is weakly decreasing, this implies that $x(c) = 0$ for all c, which cannot be. Hence, there must be some pooling of low-cost types.

Define $\chi(c_i, c_{-i}; \lambda) = z^{-1}\left(\lambda x(\max\{c_i, c_{-i}\}) + (1-\lambda)x(\min\{c_i, c_{-i}\})\right)$. When a type c_i and type c_{-i} agent interact, the resulting 2nd best average care level coincides with the unilateral optimal care level for an agent with cost $\chi(c_i, c_{-i}; \lambda)$. So, $z^{-}1(\cdot)$ maps care into costs. By construction, $x(c; \lambda) = z(\chi(c, c; \lambda))$, so to characterize the second-best schedule, it suffices to characterize the function χ.

Recall, from the unilateral problem, that $z^{-1}(x) = -\Pi'(x)$. Now, for any agent in the separating region (i.e. $c_i \geq \hat{c}(\lambda)$), we know that $x(c_i)$ is characterized by the FOC:

$$\begin{aligned}\frac{\partial W}{\partial x(c_i)} &= c_i + (1-\lambda)\int_{\underline{c}}^{c_i} \Pi'(\lambda x(c_{-i}) + (1-\lambda)x(c_i))g(c_{-i})dc_{-i} + \lambda \int_{c_i}^{\overline{c}} \Pi'(\lambda x(c_i) + (1-\lambda)x(c_{-i}))g(c_{-i})dc_{-i} = 0 \\ &= c_i - (1-\lambda)\int_{\underline{c}}^{c_i} \chi(c_i, c_{-i}; \lambda)g(c_{-i})dc_{-i} - \lambda \int_{c_i}^{\overline{c}} \chi(c_i, c_{-i}; \lambda)g(c_{-i})dc_{-i} = 0 \qquad (5)\end{aligned}$$

Notice that (5) defines a system of equations (one for each c_i) that characterizes the function χ directly in terms of c's and λ. Moreover, in the case of $\hat{c}(\lambda)$ we have:

$$(1-\lambda)G(\hat{c}(\lambda))\underbrace{\chi(\hat{c}(\lambda), c(\hat{\lambda}); \lambda)}_{=\chi(\hat{c}(\lambda))} + \lambda \int_{\hat{c}(\lambda)}^{\overline{c}} \chi(\hat{c}(\lambda), c; \lambda)g(c)dc = \hat{c} \qquad (6)$$

Now, consider the optimal pooling standard. This standard must minimize the social loss amongst members of the pool, subject to the constraint that the threshold type is kept indifferent between pooling and separating. The pooling loss is:

$$2\hat{x}\int_{\underline{c}}^{\hat{c}} cg(c)dc + G(\hat{c})^2\Pi(\hat{x}) + 2G(\hat{c})\int_{\hat{c}}^{\infty} \Pi(\lambda\hat{x} + (1-\lambda)x(c))g(c)dc$$

where the final term reflects the probability that an agent in the pool is matched with an agent without. Taking the first order condition w.r.t $\hat{x}$ gives:

$$2E[c\,|\,c < \hat{c}] + G(\hat{c})\Pi'(\hat{x}) + 2\lambda\int_{\hat{c}}^{\infty} \Pi'(\lambda\hat{x} + (1-\lambda)x(c))g(c)dc = 0 \qquad (7)$$

$$2E[c\,|\,c < \hat{c}] - G(\hat{c})\chi(\hat{c}) - 2\lambda\int_{\hat{c}}^{\overline{c}} \chi(\hat{c}(\lambda), c; \lambda)g(c)dc = 0 \qquad (8)$$

Combining (6) and (8) gives:

$$\left(\frac{1}{2}-\lambda\right) G(\hat{c})\chi(\hat{c}(\lambda)) + E[c\,|\,c<\hat{c}(\lambda)] = \hat{c}(\lambda) \tag{9}$$

which implictly characterizes $\hat{c}(\lambda)$. Notice that this characterization is independent of Π, which proves the claim in Corollary 3.

The characterization is particularly straightforward in the cases of $\lambda = 0$ and $\lambda = \frac{1}{2}$. When $\lambda = 0$, we know that $\chi(\hat{c}(0)) = \frac{\hat{c}}{G(\hat{c})}$, and so (9) reduces to $\hat{c}(0) = 2E[c\,|\,c<\hat{c}(0)]$. When $\lambda = \frac{1}{2}$, (9) simplifies to $\hat{c}\left(\frac{1}{2}\right) = E[c\,|\,c<\hat{c}\left(\frac{1}{2}\right)]$, which can only be true if $\hat{c}\left(\frac{1}{2}\right) = \underline{c}$.

Finally, note that $\frac{\hat{c}-E[c\,|\,c<\hat{c}]}{G(\hat{c})}$ is increasing in $\hat{c}$. Then, since (9) can be written: $1-2\lambda = \frac{\hat{c}-E[c\,|\,c<\hat{c}]}{G(\hat{c})\chi(\hat{c})}$, and since the left-hand side term is decreasing in λ, $\hat{c}(\lambda)$ must be decreasing in λ. $\square$

Proof of Corollary 2. Suppose $\hat{c} \geq \overline{c}$. This implies that $G(\hat{c}) = 1$, and $E[c\,|\,c<\hat{c}] = E[c]$. Now, By equation (7) in the Proof of Proposition 3, the optimal pooling standard satisfies $\Pi'(\hat{x}) = -2E[c]$, which implies $\hat{x} = z(2E[c])$. Substituting this into equation (9) gives:

$$\hat{c} = E[c] - \frac{1-2\lambda}{2}(-2E[c]) = 2(1-\lambda)E[c]$$

Then, since $\hat{c} \geq \overline{c}$, it must be that $2(1-\lambda)E[c] \geq \overline{c}$, which implies that $\lambda \leq 1 - \frac{\overline{c}}{2E[c]}$. $\square$

Proof of Proposition 5. See Shavell (1987). $\square$

Proof Of Lemma 2. The logic mirrors the proof of Proposition 1. For concreteness, suppose $\phi' > 0$ and $\phi'' < 0$. Suppose $c_m < c_p$. Then it must be that $x_m \geq x_p$.

Suppose $x_m > x_p$. The planner's problem is:

$$W = \min_{x_m, x_p} c_m x_m + c_p x_p + \Pi(\phi^{-1}(\lambda\phi(x_m) + (1-\lambda)\phi(x_p)))$$

The first order conditions are:

$$\frac{\partial W}{\partial x_m} = c_m + \frac{\Pi'(a(x_m, x_p; \lambda))}{\phi'(a(x_m, x_p; \lambda))}\lambda\phi'(x_m) = 0$$
$$\frac{\partial W}{\partial x_p} = c_p + \frac{\Pi'(a(x_m, x_p; \lambda))}{\phi'(a(x_m, x_p; \lambda))}(1-\lambda)\phi'(x_p) = 0$$

where $a(x_m, x_p; \lambda) = \phi^{-1}(\lambda\phi(x_m) + (1-\lambda)\phi(x_p))$. Since $x_m > 0$, we know that $\frac{c_m}{\lambda\phi'(x_m)} = -\frac{\Pi'(a(x_m,x_p;\lambda))}{\phi'(a(x_m,x_p;\lambda))}$. Moreover, since $\frac{\partial W}{\partial x_p} \geq 0$, we know that: $\frac{c_p}{(1-\lambda)\phi'(x_m)} \geq -\frac{\Pi'(a(x_m,x_p;\lambda))}{\phi'(a(x_m,x_p;\lambda))}$. Hence:

$$\frac{c_m}{\lambda\phi'(x_m)} \leq \frac{c_p}{(1-\lambda)\phi'(x_p)}$$
$$\frac{1-\lambda}{\lambda}\cdot\frac{c_m}{c_p} \leq \frac{\phi'(x_p)}{\phi'(x_m)}$$

Then, since $\phi'' < 0$ and $x_m > x_p$, it must be that $\frac{\phi'(x_p)}{\phi'(x_m)} < 1$, and so:

$$\frac{1-\lambda}{\lambda}\cdot\frac{c_m}{c_p} < 1$$
$$\lambda > \frac{c_m}{c_m + c_p}$$

Notice that this condition is independent of ϕ. Then, taking the contra-positive, whenever $\lambda \leq \frac{\min\{c_m,c_p\}}{c_m+c_p}$, it must be that $x_m = x_p$. If so, the planner's problem becomes:

$$\min_x (c_m + c_p)x + \Pi(x)$$

whose solution is very clearly $x_m = x_p = z(c_m + c_p)$. □

B Generalized OWA Technology

In our main analysis, we used the ordered weighted average technology to aggregate the agents' individual care choices into an average care level. We chose this technology for its simplicity — it is piece-wise linear in x_m and x_p. In this sub-section, we briefly demonstrate that the results can easily accommodate a more general order weighted technology.

Let $\phi(x)$ be a continuous function satisfying either $\phi' > 0$ and $\phi'' < 0$, or $\phi' < 0$ and $\phi'' > 0$.

Define the average function:

$$\alpha(x_m, x_p; \lambda) = \phi^{-1}\left[\lambda\phi(\max\{x_m, x_p\}) + (1-\lambda)\phi(\min\{x_m, x_p\})\right]$$

where $\lambda \in [0, 0.5]$. The function α so defined returns an order weighted generalized average of x_m and x_p. Notice that, regardless of the choice of ϕ, $\alpha = \min\{x_m, x_p\}$ whenever $\lambda = 0$. Hence, all of the insights of our baseline analysis (under perfect complements) will continue to hold in this generalized setting.

Moreover, the insights will continue to hold even when the care technology is characterized by moderate complements. To see this, given the discussion in subsection 2.B., it suffices to show that when $\lambda > 0$ is small, the first best schedule continues to coordinate both agents on the same care level, even if their costs differ. Indeed, the following Lemma shows that the 'coordination regime' of the first best schedule remains unchanged if we replace the simple order weighted average with a generalized order weighted technology.

Lemma 2. *For any generalized order weighted average technology satisfying the conditions above, the first best schedule satisfies:*

$$x_m^{1st}(c_m, c_p) = x_p^{1st}(c_m, c_p) = z(c_m + c_p)$$

whenever $\lambda < \frac{\min\{c_m, c_p\}}{c_m + c_p}$.

The first best schedule will behave somewhat differently in the 'tailoring regime', where it may no longer be optimal to assign the entirety of care to the least cost avoider. Indeed, by convexifying the first best schedules in this regime will be more continuous, and have less of a 'bang-bang' flavor.

In section 1, we contrasted the OWA technology with the more familiar CES technology, both of which facilitate a parameterization of the degree of substitutability between the agents' care decisions. By setting $\phi(x) = x^\rho$, we can combine these approaches by defining the order weighted CES aggregator:

$$\alpha(x_m, x_p; \lambda, \rho) = [\lambda(\max\{x_m, x_p\})^\rho + (1-\lambda)(\min\{x_m, x_p\})^\rho]^{\frac{1}{\rho}}$$

C Online Appendix: A Model of Contracts

In this section, we briefly show how the model can be adapted to capture interactions in a contracts setting. As we will show, but for some (intuitive) modifications, the results from our main analysis carry over exactly.

Consider an interaction between a buyer B and a seller S. The buyer and seller may each invest effort $x_i \geq 0$ to facilitate the creation of a surplus. Effort is costly, and the unit cost of effort is c_i for agent $i \in \{B, S\}$. The value of the surplus depends on a measure of the 'average' effort exerted, and the intrinsic value of the item being transacted to the buyer. Formally, the surplus is $v_B \Pi(a(x_B, x_S))$, where $\Pi' > 0$ and $\Pi'' < 0$, so that effort increases the size of the surplus, but with diminishing returns. As usual, we construct the average effort using the order weighted average technology: $a(x_B, x_S) = \lambda \max\{x_B, x_S\} + (1 - \lambda) \min\{x_B, x_S\}$, where $\lambda \in \left[0, \frac{1}{2}\right]$. We will focus on the case of perfect complements ($\lambda = 0$).

We assume that the buyer's valuation v_B and the seller's cost c_S are private information. For simplicity, we assume that the buyer's cost c_B is commonly known. v_B and c_S are each (independent) draws from continuous distributions with CDFs $G_B(v_B)$ and $G_S(c_S)$ that are sufficiently log-concave. The supports of the distributions are $[\underline{v}_B, \overline{v}_B]$ and $[\underline{c}_S, \overline{c}_S]$ (respectively), with $\underline{c}_S > 0$ and $\overline{v}_B < \infty$.

C.A. The first-best

First, let us characterize the first-best effort levels $x_B(v_B, c_S)$ and $x_S(v_B, c_S)$, which are the solutions to:

$$\max_{x_B, x_S} v_B \Pi(\min\{x_B, x_S\}) - c_B x_B - c_S x_S$$

Straightforwardly applying first order conditions, we find that the efficient investments are:

$$x_i(v_B, c_S)^{1st} = [\Pi']^{-1} \left(\frac{c_B + c_S}{v_B} \right) = z \left(\frac{c_B + c_S}{v_B} \right)$$

where $z(a) = [\Pi']^{-1}(a)$. This is analogous to the first-best expression in our baseline model, except that each agent's costs are normalized by the marginal benefit parameter v_B. We can easily verify that $z'(a) < 0$, so the first-best schedule is decreasing in each agent's costs, and increasing in the buyer's valuation.

C.B. The second-best

Now, consider the second-best setting, where the planner can condition each agent's investment decision on their own type (which is private information), but not on the opponent's type. The second-best effort levels $x_B(v_B)$ and $x_S(c_S)$ are the solutions to:

$$\max_{x_B(v_B),x_S(c_S)} \iint \{v_B \Pi(\min\{x_B, x_S\}) - c_B x_B - c_S x_S\} \, g_B(v_B) g_S(c_S) dv_B dc_S$$

Let $v_B(c_S)$ and $c_S(v_B)$ be functions that are implicitly defined by:

$$\frac{c_B}{v_B} E[v \,|\, v > v_B](1 - G_B(v_B)) = c_S G_S(c_S)$$

This expression is the analogue of the expression $c_m G_m(c_m) = c_p G_p(c_p)$ that we defined in Section **??**. It will turn out that a seller with cost c_S will make the same investment as a buyer with valuation $v_B(c_S)$. Similarly, a buyer with valuation v_B will make the same investment as a seller with cost $c_S(v_B)$. Naturally, $v_B(c_S)$ and $c_S(v_B)$ are inverse functions of one another. We can verify, via the implicit function theorem, that $v_B'(c_S) < 0$, and similarly that $c_S'(v_B) < 0$. A seller with a higher cost will make the same investment as a buyer with a lower valuation, and vice versa.

The second-best effort schedules are characterized as follows:

Proposition 6. *There exist threshold $\hat{v}_B < \overline{v}_B$ and $\hat{c}_S > \underline{c}_S$ uniquely defined by:*

1. *$\hat{v}_B = v_B(\hat{c}_S)$ (or equivalently, $\hat{c}_S = c_S(\hat{v}_B)$), and*
2. *$\frac{E[c_S \,|\, c_S < \hat{c}_S]}{\hat{c}_S} + \frac{\hat{v}_B}{E[v_B \,|\, v_B < \hat{v}_B]} = 1$*

such that the second-best investment schedules are:

$$x_B^{2nd}(v_B) = \begin{cases} \hat{x} & \text{if } v_B > \hat{v}_B \\ z\left(\frac{c_B}{v_B G_S(c_S(v_B))}\right) & \text{if } v_B \leq \hat{v}_B \end{cases}$$

$$x_S^{2nd}(c_S) = \begin{cases} \hat{x} & \text{if } c_S < \hat{c}_S \\ z\left(\frac{c_S}{E[v_B \,|\, v_B > v_B(c_S)](1 - G_B(v_B(c_S)))}\right) & \text{if } c_S \geq \hat{c}_S \end{cases}$$

where $\hat{x} = z\left(\frac{\frac{c_B}{G_S(\hat{c}_S)} + \frac{E[c_S \,|\, c_S < \hat{c}_S]}{1 - G_B(\hat{v}_B)}}{E[v_B \,|\, v_B > \hat{v}_B]}\right)$.

Proposition 6 is analogous to Proposition 4. (Since the proof strategy is identical to the proof of Proposition 4, we do not replicate it here.) Similar to our main model, there is pooling amongst the agents who would ideally make larger investments; i.e. high valuation buyers and low cost sellers. This pooling mitigates the adverse selection problem that would otherwise arise, due to such agents understanding that their likely interactions would be with 'worse-type' opponents who made lower investments. Sellers with sufficiently high costs, and buyers with sufficiently low valuations are excused from meeting this 'reasonable' standard, and may instead take effort commensurate to their costs/valuations.

A few key points are worth noting. Whenever there is separation, both the buyer and seller condition their investment on their cost of effort *relative* to the buyers valuation. In the buyer's case, this valuation is known, and so the buyer's investment level depends purely on $\frac{c_B}{v_B}$. By contrast, the seller does not know v_B, and so can only user her expectation of the buyer's valuation (conditional upon the buyer taking more care than her). As in our baseline model, these relative costs of effort are converted into effective relative costs, reflecting the probability that each agent's effort is wasted.

Second, the pooling investment level is precisely the first-best investment for the average agent within the pool, given their effective (relative) costs of effort. This exactly matches the result from the baseline model. Furthermore, if $\hat{c}_S \geq \overline{c}_S$ and $\hat{v}_B \leq \underline{v_B}$ (which will occur if both $E[c_S] + c_B > \overline{c}_S$ and $\frac{1}{E[c_S]} + \frac{1}{c_B} > \frac{1}{\underline{v}_B}$), then there will be complete pooling. If so, then the pooling level will be:

$$\hat{x} = z\left(\frac{c_B + E[c_S]}{E[v_B]}\right)$$

which is precisely the first-best effort level when matching the average buyer with the average seller.