A Theory of the Reasonable Person Standard

Scott Baker
Washington University in St. Louis

Giri Parameswaran
Haverford College

PRELIMINARY & INCOMPLETE

November 6, 2022

Abstract

Why is what a reasonable person would do the standard for assessing whether an individual's actions should result in liability? To answer this question, we develop a model where two actors must decide how much to invest in accident prevention. Each actor knows their own cost, but is ignorant of their counter-party's. If coordination is critical, the court prefers to hold actors to a standard associated with some average cost, rather than a standard targeted to what is, in fact, the actor's true cost—a cost the court knows perfectly. Further, the court optimally meshes the objective standard with a path towards forgiveness for the least able. Throughout, we show how many otherwise puzzling legal rules and doctrine are consistent with the predictions of the model.

1 Introduction

Through the rules governing criminal law, torts, contracts, property, and intellectual property, judge-made law prevents externalities, facilitates trades and encourages investment in research and development. To do these tasks, courts consistently refer to the 'reasonable' person — a disembodied hypothetical construct. Here are a few examples:

- A person can be held liable in tort only if they fail to use reasonable care to prevent the accident (Dobbs et al., 2015, p.213). Reasonable care is defined as the "care, attention or skill a reasonable person would use under similar circumstances" (MSBNA Standing Committee on Pattern Jury Instructions, 2009).
- In a lawsuit claiming the employer created a hostile work environment, the jury must consider the" objective severity of harassment . . . from the perspective of a reasonable person in the plaintiff's position, considering 'all the circumstances."'[1]
- 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."[2].
- To obtain a patent, an inventor must show that the invention is non-obvious (35 U.S.C. §103). To do so, the court asks how difficult the invention would have been to a develop for a person skilled in the art, the typical or average inventor in the field.
- Privacy rights attach when the revealed conduct occurs in a location where a reasonable person would expect privacy.

Who is the reasonable person, and why is she so important in determining rights and obligations? Is the reasonable person standard a good idea? What problem is it solving? Further, how broad is the group that makes up the reasonable person reference? Who is in and who is out?

Let us expand on the last point by way of example. In the hostile work environment case does the reasonable person refer to both men and women, only women, or only a women with the same education and experience as the plaintiff? As the group expands or contracts

[1] *Oncale v. Sundowner Offshore Servs., Inc.*, 523 U.S. 75, 81 (1998).
[2] *Martin v. Brit. Am. Oil Producing Co.*, 1940 OK 218.

so does the required standard of conduct: what is reasonable depends on who is included in the reference group.

To defend the reasonable person standard, Oliver Holmes provided two justifications. The first involved the cost of adjudication. Because courts find it costly to observe differences in agents' capacities and circumstances, it is efficient to treat all litigants as if they were the 'average' person (Holmes, 1881; Posner, 2014; Shavell, 1987). This rationale has dominated the discussion in the scholarly debate. Yet courts routinely sort false claims from true claims. Why are claims about aptitude so different from claims about, say, intent to commit a crime?

As to the second justification, Holmes wrote:

> [W]hen men live in society, a certain average of conduct, a sacrifice of individual peculiarities going beyond a certain point, is necessary to general welfare. 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 than if they sprang from guilty neglect. His neighbors accordingly require him, at his peril, to come up to their standard, and the courts which they establish decline to take his personal equation into account (Holmes, 1881, p. 108).

We explore the second rationale. In so doing, we provide an answer to why an individual might wish that the law forced his neighbor to meet an objective standard, to act as a reasonable person. Our account makes sense of the role of legal rules in inducing coordination among strangers. Because of this fact, courts, we show, will rationally ignore information about an individual's cost, even if that data is freely available and perfectly accurate.

The reasonable person standard has attracted much attention from legal scholars, philosophers, and scholars in the law and economics tradition. Yet economists have neglected the topic, taking the development and existence of the background rules of contract, tort, and property as given. Our simple model fills that gap. Specifically, we show that the reasonable person standard arises endogenously where the law is trying to coordinate efforts at harm reduction in the presence of asymmetric information.

An example illustrates the argument. 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 effort is for the other party. Suppose that the technology is such that the party taking the least amount of care determines the probability of an accident. Like the case of perfect complements in input productions, the first best involves coordination: both the pedestrian and the motorist should devote the same effort towards accident avoidance.

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. Lots of care by the motorist only makes sense if the pedestrian mirrors that choice. The reflection only happens if the pedestrian also has a low cost of care, which is far from certain. Judge-made law improves the chance of coordination by having the low-cost motorist chisel on his care. In other words, society wants to have the low-cost motorist ignore his talent for harm prevention to do what the "average" 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 information that would seemingly be relevant to any cost-benefit analysis, namely each individual's cost of care. Remarkably, despite the perfect complementarity, we find that the law need not force all agents to the same level of care. Instead, the law often grants the least able agents forgiveness, and allows them to invest less in accident prevention than everybody else.

From the low-cost motorist's perspective, the presence of pedestrians with higher costs infects the pool of counter-parties. The more pedestrians with high costs in the pool, the more coordination becomes an issue and the greater the distortion in the care decision of the low-cost motorist. High cost pedestrians are akin to 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 from taking care will not be wasted.

As noted, the court does not force all actors to take the same care level. Rather than rid the market of lemons, the law accommodates the highest-cost actors as much as possible. The reasonable person standard ignores individual differences in costs among low and medium cost actors in a hunt for the successful coordination of efforts among strangers. Coordination demands that each actor subject to the objective care level be paired with sufficiently high probability to a counter-party subject to the same objective care. Once that pairing is achieved, the law grants excuses to everyone else. In other words, while the objective standard 'pools' types on a constant care level, it need not do so fully. The reasonable person standard with exceptions — what we observe in practice — arises endogenously.

Our primary purpose is the study of judge-made law; that said, the use of objective standards spreads beyond the courts. Employment contracts, for example, often require that the employee exert reasonable efforts.

The reasonable person standard has been subject to exploration by three different strands of literature. First, philosophers suggest that the reasonable person standard imports ideas of 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 provides one way to flesh out what makes the terms of an interaction between two strangers "fair."

A second strand comes from feminists and other critical scholars. For example, Bender (1988) suggests that the reasonable person standard is a vehicle through which judicial discretion is granted; and judges use this discretion as a mechanism to preserve social hierarchies. Alternatively, Bernstein (2001) argues that the reasonable person standard provides courts with cover. The court can impose liability while pinning the reasons for it on some external community—the community of reasonable actors.

The most closely related literature comes from the law and economics scholars. Landes and Posner (1987) define reasonable as actions that are consistent with cost-benefit analysis. Of course, different people have different costs of accident prevention, suggesting the law should be finely tailored. The one-size-fits-all reasonable person standard arises because the courts cannot observe cost differences among individuals (Shavell, 1987).

Unlike Shavell (1987), we assume the court knows each actor's cost of exercising care. Instead, a friction arises because parties do not *themselves* know the cost of others with whom they interact, and thus cannot predict the actions of their counterparties.

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 appropriate vehicle for controlling conduct.

Before turning to the model, we pause to make a point about the relationship between 'objective' standards and recent technological advances. Technological advances has made learning about individualized characteristics cheaper and easier. Building off this, some scholars advocate that the law account for more and more individual traits–that the law become more and more personalized (Ben-Shahar and Porat, 2016). Our model counters this suggestion. We show that what matters is what each individual knows about others with whom he interacts. Even if courts and regulators could use big data to learn everyone's personalized cost, the reasonable person standard is still necessary.

The paper unfolds as follows. Section 2 presents a model of accident prevention. In designing the law, the court's task breaks down into two parts. First, the legal system must figure out what it would like each party to do—how much effort to spend on accident prevention. Second, the court must design legal doctrine that induces parties to make those choices. Should the court use negligence, contributory negligence, comparative negligence, strict liability, or some other legal rule? In solving this problem, the court must be mindful of the information the parties possess about each other. Section 3 articulates the first best decision rule assuming all the actors and the court know everything about each other.

Section 4 relaxes the assumption that the parties know each other's costs. There, we demonstrate that objective standards arise endogenously when the parties' care decisions are complementary. Section 6 pivots to legal doctrine, showing how the various legal rules found in practice serve to implement the standards of care derived in Section 4. Section 7 offers some concluding remarks. Throughout the analysis, we draw connections between the predictions of the model and the common law and use those connections to make testable predictions.

2 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, imposing a loss of one on the pedestrian. Each motorist and each pedestrian must decide how much care to take in preventing the accident. 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 and unimodal distribution $G_i(c)$, that admits a strictly positive density $g_i(c)$, with a support on $[\underline{c}_i, \overline{c}_i]$, where $\underline{c}_i > 0$. Under a subjective or fully tailored standard, a court evaluates the actor's care decision in light of each actor's individual cost of care. Under an objective standard, the court does not. Under a mixed objective/tailored standard, the court evaluates the conduct of some actors with respect to their individualized personal cost, but holds others to a standard associated with some average or group.

The court observes each c_i perfectly. Thus, the court can, if it so chooses, successfully use a fully tailored standard for both the pedestrian and the motorist. While unrealistic, this assumption decouples the motivations for objective standards originally presented by Holmes (1881). One motivation, discussed above and subject to a large literature, roots the reason in adjudication costs for the court. Our frictionless setup nvestigates the second, and heretofore unexplored motive: the coordination of care decisions.

The efforts of the pedestrian and the motorist combine to determine the probability of harm. The literature commonly assumes a probability technology $P(x_m, x_p)$ that is decreasing in each agents' care level, and jointly convex. We specialize the probability technology as follows: the probability of harm is a function $\Pi(a)$ of some measure of the central tendency a of the two agents' care. As usual, we assume that Π is twice continuously differentiable, strictly decreasing and strictly convex, so that $\Pi' < 0$ and $\Pi'' > 0$. Strict convexity implies diminishing marginal returns to care. We additionally assume that Π satisfies the Inada conditions, which guarantees that the solution to the first order condition will be interior.

The measure of central tendency is determined by an *ordered weighted average* (OWA) technology[3]:

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

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.[4]

[3]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 and Liu, 2014; Yager, 2002,0). But the OWA technology has been applied in a wider range of contexts, including the modelling and measurement of inflation (León-Castro et al., 2020), asset valuation (Doña et al., 2009), exchange rate forecasting (León-Castro et al., 2016), risk analysis (Blanco-Mesa et al., 2018), and government accountability (Avilés-Ochoa et al., 2018), amongst others.

[4]The simplicity of the OWA technology — it is piecewise linear — recommends its use. In the appendix, we show that the main insights of our analysis will continue to hold if we instead chose a more generalized (e.g. convex) order weighted technology.

Specifying the probability technology in this way enables a clean and parsimonious articulation of the degree of substitutability or complementarity of agents' care decisions in reducing the likelihood of harm. Indeed, the parameter $\lambda \in [0, 0.5]$ captures the degree of substitutability between the agents' care decisions. Starting from a baseline of equal care levels $x_m = x_p$, a one unit increase in care by some agent enables a $\frac{\lambda}{1-\lambda} \leq 1$ unit decrease by the other agent to keep average care unchanged. When $\lambda = 0.5$, the technology simplifies to the arithmetic mean: $a(x_m, x_p) = \frac{1}{2}x_m + \frac{1}{2}x_p$. Care levels can be substituted one-for-one; there is a perfect substitutes technology. When $\lambda = 0$, the technology reduces to the Leontief (perfect complements) technology: $a(x_m, x_p) = \min\{x_m, x_p\}$; there is no substitutability between care levels.

We use the OWA technology to capture substitutability instead of the more familiar CES technology.[5] In this setting, the CES technology has an unappealing feature: if the motorist's and pedestrian's care decisions are close to equal in value, the CES function is locally approximated by a perfect substitutes technology for any value of ρ.[6] In other words, different values of ρ do not meaningful capture differences in complementarity when agent match their care.

3 The Social Planner's Problem: First Best

In *United States v. Carroll Towing Co.*,[7] Judge Learned Hand set forth an economic understanding of tort law. Judge Hand and many others in the law and economics tradition argue that accident law trades off the cost of precautions against the benefits those precautions have in reducing accidents (Landes and Posner, 1987; Posner, 1972; Shavell, 1987).[8] Following Judge Hand's lead, our court adopts legal rules with an eye towards minimizing the sum of accident costs and prevention costs.

[5]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. Many of our results, including all of the baseline analysis under perfect complements (in Section 4.1), would continue to hold if we replaced the OWA technology with the CES technology.

[6]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: $b(x_m, x_p; \rho) \approx \frac{1}{2}x_m + \frac{1}{2}x_p = b(x_m, x_p; 1)$.

[7]159 F.2d 169 (2d. Cir. 1947).

[8]Indeed, Judge Hand wrote: "Possibly it serves to bring this notion into relief to state it in algebraic terms: if the probability be called P; the injury L; and the burden B; liability depends on whether B is less than L multiplied by P: i.e. whether $B < PL$.

3.1 Unilateral Benchmark

To begin, consider the optimal level of care in a unilateral accidents problem, where the probability of harm depends on the level of care of a unitary actor. Since there is a single actor, there is no need to aggregate the care levels of multiple actors into an 'average' care level. The optimal level of care z satisfies: $\min_z \Pi(z) + cz$.

Straightforwardly, the efficient unitary care level is $z(c) = [\Pi']^{-1}(-c)$. The function $z(c)$ returns the optimal care level for every level of cost. Moreover, $z'(c) < 0$, so that the optimal unitary care level decreases in the cost of care.

3.2 Bilateral First Best

Return now to the bilateral problem. Consider the full information benchmark, in which the actors' cost are observable to everyone — the court and the actors themselves — and so each agent can condition their choice on *both* agents' costs. The first-best care decisions minimize the social loss associated with harms and harm mitigation:

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

The first best care levels are 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:*

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

The first best care decisions exist in one of two regimes. If the degree of complementarity is relatively high (i.e. λ is relatively low), then the planner will coordinate both agents on the same level of care. Moreover, since the both agents must pay the cost of providing this care, the optimal care level will coincide with the one that would be chosen by a unilateral agent facing unit cost $c_m + c_p$.

If, instead, the degree of complementarity is relatively 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. Moreover, the standard of care that the least cost avoider is held to depends on λ. For concreteness, suppose $c_m < c_p$, so that the motorist is responsible for taking care. The motorist's effective unit cost of care is $\frac{c_m}{\lambda}$, since the motorist must take $\frac{1}{\lambda}$ units of care to increases the average care level by 1. The first best standard of care for the motorist is the one that causes the average care level to coincide with the unilaterally optimal care level for an agent facing the same effective cost as the motorist.[9]

Early literature in law and economics referenced the least cost avoider to explain the common law's assignment of liability (Calabresi and Hirschoff, 1971, p.1060).[10]

To summarize if the agents' care decisions are complements, it is optimal for the agents to coordinate their care, even if one agent has higher costs of care than the other. Further, the level of coordinated care turns on the sum of the two costs. By contrast, when care decisions are substitutes, then the agents will take different levels of care, such that agents with lower costs will take more care.

Consider now the difficulty of implementing this first best care decisions. With substitutes, the court would only assign liability to the motorist when her cost was less than the pedestrian's. To know what level of care to take, the motorist would need to know not only her own cost, but the cost of every pedestrian she encounters. Likewise, with complements, the court would find, say, the motorist negligent if her care fell short of a marker that is a function of the sum of the motorist's cost and the pedestrian's cost. Absent knowledge of the pedestrian's cost, the motorist could not compute the sum, and thus she could not understand how much care to take.

To induce efficient care decisions, Garoupa and Dari-Mattiacci (2007) make plain that the court must do more than uncover each individual's cost of care through litigation. At the time of the accident, the actors must themselves know the cost of care for everyone else they might be involved in an accident with. This information is, in reality, unavailable, rendering it impossible for the court to implement the first best.

[9] 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 Section 4.2.3, with a more convex technology, the first best care levels in this region would be more 'continuous'.

[10] Building on this, Shavell (2004) argued that the applicability of the least cost avoider principle was limited to cases where care was highly substitutable between the parties, a result in accord with this model.

4 Second Best Analysis: Objective Standards and Excuses

Having studied the benchmark, the analysis now examines what happens when (1) the actors can observe their own costs; (2) the court can observe the actors' cost; yet (3) the actors cannot observe each other's costs. In this setting, the court can make each actor's care decision a function of their costs only. Should it do so? Seemingly yes, given that this information is relevant to creating proper incentives. Motorists, for instance, who have a low cost of care should be held to a higher standard, forced to put in more effort than motorists with a high cost of care.

In fact, we show that it may benefit social welfare for the court (or planner) to ignore each agent's costs, and instead hold them to a common objective standard. In Section 4.1, we focus on the special case where the care technology is characterized by perfect complements ($\lambda = 0$), so that $a(x_m, x_p) = \min\{x_m, x_p\}$. In section 4.1, we will additionally assume that the agent's costs are independent draws from the same distribution. Both of these assumptions are purely to simplify the exposition of results. In section 4.2, we generalize our results on a variety of dimensions, including by allowing the care technology to be characterized by imperfect complements, and allowing the agents' costs to be drawn from different distributions.

4.1 Perfect Complements and Identical Distributions

Given identical distributions of cost, the second best care levels solve:

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

It is straightforward to show that the second best care schedules $x_i(c_i)$ must be continuous (by Berge's Theorem of the Maximum) and weakly decreasing in c_i. Suppose they are strictly decreasing, $x_p'(c_p) < 0$ and $x_m'(c_m) < 0$. If so, the second best schedules *separate* each agent's types according to their cost of care; they are subjective or purely tailored standards.

The care schedule for each type satisfies the first order condition. Focusing on the pedestrian, we have:

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

where the probability is taken with respect to the distribution over c_m. Because the pedestrian does not know the motorist's cost of care, her care can only be fixed according to the distribution of costs. The pedestrian therefore must treat the motorist's care as a random variable (a function of the unknown cost). Likewise, the pedestrian must treat the pedestrian'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 parties do not know *ex ante*.

The first term is the marginal benefit of care. Additional care by the pedestrian reduces the probability of harm, but only when the pedestrian takes less care than the motorist. The second term is the marginal cost of care.

We can use the fact that care schedules are invertible and exploit the symmetry in the setup to make more sense of the marginal benefit of care. Symmetry means that the motorist and pedestrian have the same schedules, that is, the care induced for each level of cost ($x_p(c) = x_m(c) = x(c)$) must the same.

The probability that the pedestrian's care level is not wasted reduces to:

$$\Pr[x_m(c_m) > x_p(c_p)] = \Pr[c_m < x^{-1}(x(c_p))] = \Pr(c_m < c_p) = G(c_p)$$

Because the schedule is decreasing, the pedestrian's care is pivotal (i.e., lower) when they have a higher cost then the motorist. And that happens with probability $G(c_p)$.

Accordingly, the first order condition becomes:

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

An agent's willingness to take care depends on the cost of doing so, adjusted by the likelihood that their care will be wasted. Wasted effort occurs whenever the counter-party takes less care than the agent. The agent's care effort determines harm reduction when they are matched against an opponent whose cost is at least as low as their own. Indeed, we can interpret the expression $\frac{c}{G(c)}$ as the agent's effective cost of care — it is the cost of increasing the average care level by one unit, in expectation. The optimal care level is the unilaterally optimal care level appropriate to the agent's effective cost.

When the agents' care levels are complementary, each agent's willingness to take care depends on the likelihood that their counter-party will take at least as much care. An agent with relatively low costs may be dissuaded from taking high care, if she thinks the likelihood of being matched with an agent who takes less care is sufficiently high. Analogous to the classic adverse selection problem, the presence of the high-cost pedestrians distorts the investment decisions of low-cost motorists. Holding low cost actors to high levels of care — care more in tune with their costs — is inefficient. Such care is likely to be wasted because the low-cost actor is unlikely to be the party pivotal in determining the probability of harm.

Indeed, the incentive for the very lowest cost agents to take care collapses entirely.

The problem can be brought into stark relief by investigating the shape of $x(c) = z\left(\frac{c}{G(c)}\right)$. Since z is a decreasing function, the slope of $x(c)$ has the same sign as the slope of the average CDF function, $\frac{G(c)}{c}$. For every continuous distribution with $\underline{c} > 0$, the average CDF must be increasing when c is near $\underline{c}$. Similarly, the average CDF must be decreasing whenever $c > \overline{c}$. If G is unimodal so that g is single-peaked, then $\frac{G(c)}{c}$ will be single-peaked as well. Hence, for c low enough, the assumption that $x(c)$ is strictly decreasing generates the conclusion that, it is, in fact, strictly increasing. Obviously this cannot be. $x(c)$ cannot be strictly decreasing everywhere; it must be flat when c is sufficiently low. There must be some pooling.

Note well that our analysis has only showed that $x(c)$ cannot be strictly decreasing when c is low. It may still be consistent with our analysis that $x(c)$ is decreasing for higher values of c, provided that $\frac{G(c)}{c}$ is decreasing in that region. This explains the asymmetry in the availability of excuses. The court may well want to tailor care levels when c is high, but not when c is low. As we have explained, the reason is that the adverse selection has its strongest bite when applied to agents taking higher care levels. These insights are summarized in the following result:

Proposition 2. *There exists a threshold cost $\hat{c} > \underline{c}$, such that $x(c)$ is constant (i.e. an objective standard) for all agents with $c < \hat{c}$, and $x(c)$ is strictly decreasing for all $c > \hat{c}$.*

The next question is the breadth of the pooling region: how many actors should be able to avail themselves to excuses and how many actors should be forced to comply with the objective standard of care? Denote by $\hat{c}$ the type on the boundary between the pooling and separating regions of the second-best schedule. Such an agent must be indifferent between the objective standard and lodging an excuse. The indifferent type's care is implicitly defined by:

$$\Pi'(\hat{x}) = \frac{\hat{c}}{G(\hat{c})}. \tag{2}$$

The court chooses the pooling (objective) care level to minimize the social loss to agents in the pool from complying with this standard, subject to the constraint in expression (2), that the threshold type $\hat{c}$ must be indifferent joining the pool and lodging an excuse.

Suppose the planner 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 objective or pooling region, and thus are pivotal to determining harm. (Since the care schedule is decreasing, agents outside the pool will take less care, and so the care of agents within the pool will be wasted whenever this happens.) The second term is the cost born by agents in the pool. Taking first order conditions gives:

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

Let us interpret this. Recall that the 'effective cost' to an agent in the pool with cost c is $\frac{c}{G(\hat{c})}$; this is the cost of increasing the average care level by one unit in expectation, recognizing that care is sometimes wasted. Then, $\frac{E[c\,|\,c<\tilde{c}]}{G(\tilde{c})}$ is the expected effective cost of an agent in the pool. So, the optimal pooling care level is the optimal care level that would be taken by a unilateral agent whose cost of care was the sum of the expected effective costs of the motorist and pedestrian, conditional upon being in the pool. (Since the agents are drawn from the same cost distribution, the sum of expected costs is simply twice the conditional expectation.) But, by Proposition 1, this is the first best care level under perfect complements if all agents in the pool are treated as if they were the average agent in the pool, with costs adjusted for the likelihood of mismatch. Thus, the objective pooling standard is a reasonable person standard — it treats all agents as if they were the average agent.

As we noted previously, the optimal pooling standard must not only minimize the social loss amongst agents in the pool, but must also keep the threshold agent indifferent between joining the pool and separating. Thus, at the optimal threshold $\hat{c}$, we must have:

$$z\left(\frac{\hat{c}}{G(\hat{c}}\right) = z\left(\frac{2E[c\,|\,c < \hat{c}]}{G(\hat{c})}\right)$$

$$\hat{c} = 2E[c\,|\,c < \hat{c}] \tag{4}$$

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 is increasing for $c < c'$, which we know cannot be. 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 (3). 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 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.

The cost type where the solid and dashed curves intersect identifies the marker between the objective and subjective 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 optimal pooling care level 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). This reduces the *effective* costs of care, and thus increases the optimal pooling care level. The $\hat{c}$ defined by (4) makes this trade-off optimally, resulting in the highest possible care from agents within the pool.

Drawing these insights together, we have the following proposition.

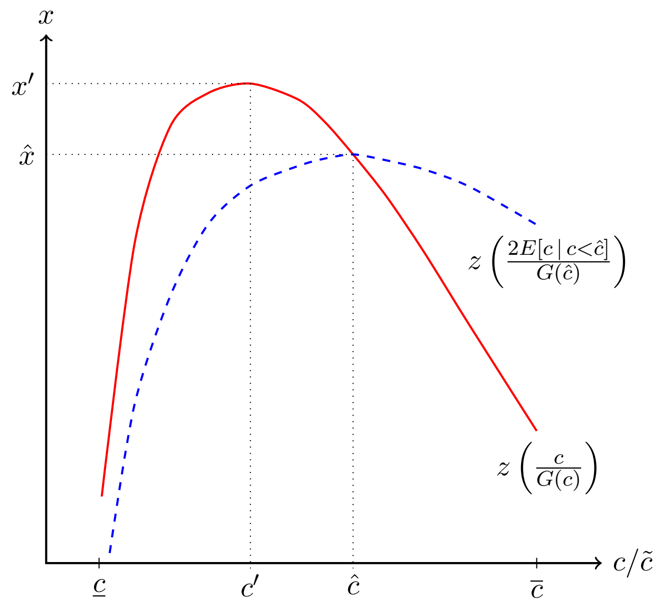
Proposition 3. *Suppose c_m and c_p are independent draws from the same distribution. Then $x_m(c) = x_p(c) = x(c)$. There exists a unique threshold $\hat{c} > \underline{c}$ characterized by $\hat{c} = 2E[c|c < \hat{c}]$, s.t.*

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

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

Proposition 3 guarantees that $\hat{c} > \underline{c}$, so that there will always be a region of types who are pooled. It does not guarantee that $\hat{c} < \overline{c}$. If $\hat{c} \geq \overline{c}$, then the second best care schedule is characterized by complete pooling; all agents are held to a pure objective standard. Moreover,

Figure 1: Breadth of the Objective Test



the pooling standard simplifies to $z(2E[c])$. But this is simply the first best care level under perfect complements when matching the (unconditional) average motorist and the average pedestrian. Thus, the objective standard precisely coincides with a reasonable person standard. It holds all agents to the level of care that would be efficient for the average person.

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 may meet a lower standard, instead. The optimal pooling care level is a 'modified reasonable person standard'; it is the first-best care level applied to average person in the pool, given their *effective* cost of care. 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, but ensures 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.

The breadth of the pooling region, and the possibility and likelihood of excuses depends on the size of $\hat{c}$, which in turn depends on the distribution of costs $G(c)$. (Notice, by Proposition 3, that the condition that defines $\hat{c}$ is independent of the accident technology Π.) In Lemma 1, below, we determine whether the law will be characterized by a pure or partial objective standard. Furthermore, when excuses are available, we provide some insights into how the breadth of the pool and the likelihood of agents availing themselves of excuses vary with the distribution of costs.

Let $c_1 \sim G_1$ and $c_2 \sim G_2$ denote two different distributions of costs, and let $\hat{c}_1$ and $\hat{c}_2$ be the corresponding thresholds. Recall that these thresholds are defined by $\frac{E[c_i \,|\, c_i < \hat{c}_i]}{\hat{c}_i} = \frac{1}{2}$, 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 second best schedule is characterized by a pure objective standard whenever the costs are not distributed too broadly above the mean. (Formally, $\hat{c} > \overline{c}$ whenever $\overline{c} < 2E[c]$.) Additionally, the threshold $\hat{c}$ is responsive to the distribution of costs in the following way:*

1. ***Scaling**: Suppose $c_2 = \kappa c_1$ with $\kappa > 0$. Then $\hat{c}_2 = \kappa \hat{c}_1$, and $G_2(\hat{c}_2) = G_1(\hat{c}_1)$. The threshold scales, and the probability of excuses is unchanged.*
2. ***Translation**: Suppose $c_2 = c_1 + \kappa$ with $\kappa > 0$. Then $\hat{c}_2 > \hat{c}_1 + \kappa$, and $G_2(\hat{c}_2) > G_1(\hat{c}_1)$. There is a lower probability of excuses.*
3. ***Mean Preserving Spread**: Suppose G_2 is a MPS of G_1, and $G_2(\hat{c}_1) < G_1(\hat{c}_1)$ (i.e. the c_1 standard applied in the c_2 setting produces more excuses). Then $\hat{c}_2 < \hat{c}_1$, and $G_2(\hat{c}_2) < G_2(\hat{c}_1) < G_1(\hat{c}_1)$. There will be more excuses.*

In summary, we have provided 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 among low-cost actors because they fear involvement in an accident with a high-cost actor. This problem is inherent to interactions between strangers; it will arise whenever the agents do not know salient details about their counter-party 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.

4.2 Generalizing the Results

In presenting our baseline results, we made several simplifying assumptions. Many of these assumptions were purely for ease of exposition, and our key insights continue to hold when we relax them. In this subsection, we consider a series of generalizations of the baseline assumptions to demonstrate the robustness of our results.

4.2.1 Imperfect Complements

The baseline analysis assumed that the agents' care choices were perfect complements. This meant that whenever an agent took more care than her counter-party, 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. As we saw in Proposition 1, when λ is positive but small (i.e. with imperfect complements), agents with different costs may nevertheless be coordinated on the same level of care because the gain to having the low-cost type take more care, though it is positive, is not worth the additional cost. This basic insight motivates the following result:

Proposition 4. *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 $c < \hat{c}(\lambda)$. The threshold $\hat{c}(\lambda)$ is defined implicitly by:*

$$\hat{c}(\lambda) = E[c \,|\, c < \hat{c}(\lambda)] - \frac{1-2\lambda}{2} G(\hat{c}(\lambda))\Pi'(\hat{x}(\lambda))$$

Additionally, $\hat{c}(\lambda)$ is strictly decreasing in λ, and $\hat{c}(0) = 2E[c \,|\, c < \hat{c}(0)]$ and $\hat{c}(0.5) = \underline{c}$.

Proposition 4 tells us that the optimal second best schedule will apply an objective standard to low cost agents. However, the breadth of the pooling interval narrows as the agents' care becomes increasingly substitutable, such that in the perfect substitutes limit, the pooling region disappears entirely. An implication of the proposition is that, if under perfect complements the second best schedule is a pure reasonable person standard (i.e. $\hat{c}(0) > \overline{c}$) , then this will remain true even if the care technology becomes somewhat substitutable. We formalize this result in Corollary 1 below.

Corrolary 1. *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 $\hat{x}(\lambda) = z(2E[c])$.*

4.2.2 Non-Identical Distributions

In the baseline analysis, we assumed that the motorist and pedestrian's costs were drawn from the same distribution. This assumption simplified the exposition of results, though it was not in any way crucial to the analysis. In this subsection, we generalize the results to the case when the distributions of costs differ amongst the agents.

Let $G_m(c)$ and $G_p(c)$ be continuous and unimodal distributions. For $c_p \geq \underline{c}_p$, let $c_m(c_p)$ be a function implicitly defined by $c_m G_m(c_m) = c_p G_p(c_p)$. Analogously define $c_p(c_m)$, and note that $c_p(c_m)$ and $c_m(c_p)$ are inverse functions. It will turn out that the second best care schedules assign the same care level to a motorist with cost c_m as to a pedestrian with cost $c_p(c_m)$; equivalently, they will assign the same care level to a pedestrian with cost c_p as to a motorist with cost $c_m(c_p)$.

We have the following result that generalizes Proposition 3:

Proposition 5. *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. $\frac{E[c_m \mid c_m < \hat{c}_m]}{\hat{c}_m} + \frac{E[c_p \mid c_p < \hat{c}_p]}{\hat{c}_p} = 1$

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 5 is a natural generalization of Proposition 3. 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.

4.2.3 Generalized Order Weighted Averages

In our baseline analysis, we used the ordered weighted average technology to aggregate the agents' individual care choices into an average care level. As noted in footnote **??**, 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 4.2.1, 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 2, 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:

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

4.2.4 Convex Costs of Care

Our baseline analysis assumed that the agents faced a constant marginal cost of care. More reasonably, one might think that the agents' cost technology is convex in their care level, so that taking care becomes increasingly more onerous the higher is the care level.

We can generalize the cost technology so that the cost of care is $\chi(x, c)$, where χ is increasing and convex in its first argument, and super-modular (so that $\frac{\partial \chi(x,c)}{\partial x} > \frac{\partial \chi(x,c')}{\partial x}$ whenever $c > c'$). Our results continue to hold qualitatively in this setting. Intuitively, pooling in our model solves the problem of low-cost agents otherwise being disincentivized to take higher levels of care. Adding increasing marginal costs only exacerbates this problem, and reinforces the value of pooling.

Additionally, if the cost function takes the form $\chi(x, c) = c\phi(x)$, where $\phi' > 0$ and $\phi'' > 0$, then our results continue to hold exactly, modifying only the definition of z, so that it is now defined by $z(c) = \left[\frac{\Pi'}{\phi'}\right]^{-1}(-c)$.

4.2.5 Distributional Assumptions

Our analysis relied upon two assumptions about the distribution of costs; first that $\underline{c} > 0$ and second that the distribution was unimodal. The first assumption ensured that the average CDF was increasing for c close to $\underline{c}$, which in turn guaranteed a region of pooling amongst low-cost types. The second assumption guaranteed that there would be a single pooling region, and that agents with higher costs would be held to an individually tailored standard.

We can weaken both assumptions. Allowing for $\underline{c} = 0$, the second best schedule will still involve pooling by the lowest types, provided that $g'(0) > 0$ (i.e. the density is increasing near the lower bound of the support). This guarantees that the average CDF is increasing in that region as well.

Allowing for multi-modal distributions simply introduces the possibility that the average CDF function increases, then decreases, then increases again (perhaps repeatedly). Necessarily, this introduces the possibility that there may be additional (distinct) regions of pooling, besides pooling at the bottom.

5 Discussion

The model illustrates several aspects of the common law of torts. To prevail on a negligence cause of action, the plaintiff must show that (1) the defendant owed the plaintiff a duty; (2) defendant breached that duty and (3) the defendant's breach caused the plaintiff's injury.

The "duty" element defines the class of potential plaintiffs. So, for example, a landowner owes a duty of care to visitors and a driver owes a duty of care to pedestrians. The duty of care is an obligation to act reasonably with respect to some counter-party. A landowner, for example, must exercise the care to avoid creating unreasonable risks of danger on their property. The motorist must exercise the care of an ordinary driver. "Breach" then is a failure to meet the legal standard. In the making the allegation, the plaintiff argues that the defendant:

1. took an action that a reasonable person would not have taken; or
2. failed to take an action that a reasonable person would have taken.

Like defendants, plaintiffs in tort cases must take actions to avoid injury. Historically, this obligation manifested itself via the defense of contributory negligence. Accused of negligence, the defendant counters that the plaintiff's own conduct was unreasonable and contributed to the accident. As a result, the argument goes, the defendant should not be held liable.

The negligence and contributory negligence inquiries measure the actor's conduct against a reasonable person standard. There are two notable exceptions where the court employs a more subjective standard of care: (1) children and (2) physical disabilities. We discuss each in turn.

5.1 Children

Take the plaintiff in *Daun v. Truax.*[11]. After school, the plaintiff, a five year old child, ran across the street via an unmarked crosswalk and was struck by the defendant's car. The defendant claimed that the child failed to yield before crossing the street and, as a result, was contributorily negligent. While that argument would be true for an adult pedestrian, the court made plain that the plaintiff should not be held to the adult standard. Instead,

[11] 56 Cal. 2d 647 (1961)

the court applied a subjective-standard based on the age of the child: a kindergartener. According to the court, a child must "exercise that degree of care expected of children of like age, experience and intelligence."[12] Such a child might not yield to a motorist before crossing the street.

The outcome exemplified in *Daun* fits the model. Irrespective of their age, adult pedestrians are held to one standard. By contrast, the law gives child pedestrians a break. The size of the break depends on the age and experience of the child, which is fair proxy for the child's cost of exercising care. The border type, $\hat{c}$ is the age of maturity: 18.

Notably, the rule for child motorists differs from the rule for child pedestrians. Because child motorists engage in a so-called adult activity, the law imposes the adult standard of care. Consider the case of *Goodfellow v. Coggburn*.[13]. There, the 13 year old plaintiff was driving a tractor on a public highway. The plaintiff failed to signal before turning left. As a result, the defendant crashed into the tractor, killing the plaintiff. Unlike in *Daun*, the *Goodfellow* court refused to customize the plaintiff's standard of care to a person of that youthful age and experience. The court opined:

> [We] take judicial notice of the hazards of automobile traffic, the frequency of accidents, the often catastrophic results of accidents, and the fact that immature individuals are no less prone to accidents than adults. While minors are entitled to be judged by standards commensurate with age, experience, and wisdom when engaged in activities appropriate to their age, experience, and wisdom, it would be unfair to the public to permit a minor in the operation of a motor vehicle to observe any other standards of care and conduct than those expected of all others. A person observing children at play with toys, throwing balls, operating tricycles or velocipedes, or engaged in other childhood activities may anticipate conduct that does not reach an adult standard of care or prudence. However, one cannot know whether the operator of an approaching automobile, airplane, or powerboat is a minor or an adult, and usually cannot protect himself against youthful imprudence even if warned.[14]

The model explains the asymmetry in tort law's approach to children in adult activities. The support of the distribution reflects the cost variation among actors who should engage the

[12] See *Daun v. Truax*, 56 Cal. 2d 647, 654 (1961).
[13] 98 Idaho 202 (1977)
[14] Id. at 203-204

activity. The unit cost for children engaging in adult activities is high, and, as a result, doing the activity at all is unlikely to be cost-justified. By prohibiting excuses for child in these activities, the law drives the child away from the activity. Accordingly, an adult motorist need not fear his high level of care will be wasted because he happens to be in accident with a child driver, where the child's low level of care is excused and pivotal in determining harm.

5.2 Physical Disability

Like children, the law holds the physically disabled to a subjective standard of care. The relevant standard is whether the actions of the disabled party conformed to the actions of a reasonable person with that disability. Consider *Roberts v. State, Through Louisiana Health Hum. Res. Admin.*[15] The defendant was a blind operator of a concession stand. While walking to the bathroom, he ran into the plaintiff, causing injuries. The court held the defendant to the standard of a reasonable person who is blind. The objective standard of care would have had the defendant look where before crossing the lobby. The blind defendant obtained a respite from this objective standard. He only had to demonstrate familiarity with the territory and a reasonable plan a blind man might adopt for getting back and forth to the restroom without harming others.

The model's predictions are consistent with these "jostling" cases, cases where care decisions are clearly complements. The law doesn't allow claims by defendants that they are clumsier than the average person and therefore should not be held liable. Likewise, the law doesn't hold the agile to a higher standard of care. Instead, the law imposes a constant standard of care with some leeway for the physically disabled.

5.3 Sticky Excuses

As noted above, the second-best scheme relies on a border type. For types above this border, the court customizes the care level to their individualized cost. The following expression defines the border type:

$$\hat{c} = 2E[c|c < \hat{c}].$$

[15] 396 So. 2d 566 (La. Ct. App.), writ granted sub nom. Roberts v. State Louisiana Health & Hum. Res. Admin., 400 So. 2d 667 (La. 1981), and aff'd, 404 So. 2d 1221 (La. 1981)

This expression is independent of the technology linking the amount of care with the reduction in harm ($[\Pi']^{-1}(c) = z(c)$). In other words, the demarcation between objective and subjective tests is independent of the technology of harm prevention.

By contrast, the optimal level of care is given by

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

Intuitively, the level of care within the objective and subjective regions goes up or down as care becomes more or less effective at preventing accidents.

These two expressions lead to a prediction. On the one hand, the frequency with which the court grants excuses from the reasonable person standard shouldn't change much across time or across types of accidents. On the other hand, the level of care the law demands should continuously vary within these two groups.

This insight sheds new light on a longstanding debate: why are courts resistant to extending subjective standards 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.[16]. But because of advances in mental health diagnosis, this argument is less persuasive today.

Notably, if limiting the reach of subjective standards ensures fewer distortions in the care decisions of others, then a focus on judicial verification costs misses the mark. Indeed, mental disability is considered for criminal culpability and can play a role in determining the intent needed for intentional torts; e.g., battery.[17] Our model explains the disparate treatment between these areas and accident cases. Intentional torts and crimes are much less likely to involve coordinating care decisions.

[16]See *Creasy v. Rusk*, 730 N.E.2d 659, 661â62 (Ind. 2000)("Adults with mental disabilities are held to the same standard of care as that of a reasonable person under the same circumstances without regard to the alleged tortfeasor's capacity to control or understand the consequences of his or her actions.")(American Law Institute, 2010)

[17]Insert cites

Table 1: Legal Standards as a Function of the Care Technology

		Legal Standard	
		Objective	Subjective
Care Technology	Complements	X	
	Substitutes		X

5.4 Substitutes versus Complements

Although our baseline analysis assumed a perfect complements care technology, we showed in Section 4.2.1 that the qualitative features of the main result — that all but the highest cost agents would be held to an objective standard —would continue to hold allowing for some substitutability in care. Moreover, in Proposition 4 and Corollary 1, we showed that the likelihood of a pure objective rule and the breadth of the pooling region when excuses were allowed, were both decreasing in the degree of substitutability in care. These results yield a testable prediction, reflected in table 1, namely that the law is more likely to be characterized by objective rules when the care technology is more complementary.

The common law reflects this division. Traffic accidents are a classic case of strong complements in care. And the law deploys an objective standard in those situations. Consider, by way of contrast, medical malpractice. Historically, physicians were held to a customary practice standard, not a reasonable person standard. For many doctors, the customary practice standard, then, was further tailored to local conditions. As such, the standard for medical malpractice has a more subjective flavor.

Medical mistakes take a variety of forms. Many, however, do not involve complementarity in the care decisions of the patient and the doctor. On this score, consider *Duvall v. Laidlow.*[18] After surgery, the physician failed to properly diagnose a post-operative infection.

In resolving the case, the court applied the "similar locality" rule, which "requires physicians to possess and apply the knowledge, skill, and care which a reasonably well-qualified physician in the 'same or similar community' would bring to a similar case. *Id.* at 722.

Under this rule, the law holds physicians from different communities to different standards. In the case at bar, the doctor's efforts at locating the infection did not hinge terribly much on the patient's efforts at, say, proper cleaning of the wound. Accordingly, our model predicts a more subjective standard for these kinds of cases, rationalizing the similar locality rule.

[18] 141 Ill. App. 3d 717, 722, 490 N.E.2d 1004, 1007 (1986)

6 Implementation: Designing the Legal Rules

The analysis so far has focused exclusively on the social planner's problem — in identifying the optimal care levels for the motorist and pedestrian. 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.

We consider two common liability rules utilized by courts: a pure negligence rule, and a strict liability rule with a defence of contributory negligence. The pure negligence rule holds the defendant liable for a harm suffered by the plaintiff only if the defendant took less than due care. A strict liability rule with a defence 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 harms: under negligence it is the plaintiff; under strict liability with contributory negligence, it is the defendant.

Proposition 6. *A pure negligence rule and a strict liability rule with contributory negligence will both implement the second best schedules* $(x_m(c_m), x_p(c_p))$*. Formally, consider*

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

Under either rule, it is Nash equilibrium for both agents to take their 2nd best care level.

Sobel (1989) show that the second best care levels can be implemented even if the court does not observe the agents' costs,

7 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.

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 model made a number of assumptions. Most important, we assumed that the parties interacted only one time. With repeated interactions (like with the buyer and seller to 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

American Law Institute (2010). Restatement of the law third, torts, liability for physical and emotional harm.

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

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

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.

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., and 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.

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

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.

Dobbs, D., Hayden, P., and Bublick, E. (2015). *Hornbook on Torts*. West Academic Publishing, 2nd edition.

Doña, J. M., la Red, D. L., and Peláez, J. I. (2009). A mix model of discounted cash-flow and owa operators for strategic valuation.

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.

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

Holmes, Jr., O. (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. and Posner, R. A. (1987). *The Economic Structure of Tort Law*. Harvard University Press, 1st edition.

León-Castro, E., Avilés-Ochoa, E., and 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, 2016, vol. 50, num. 4, p. 135-150.*

León-Castro, E., Espinoza-Audelo, L. F., Merigó, J. M., Gil-Lafuente, A. M., and 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, Fourth Edition with 2009 Supplement*. Md Institute for the Continuing Prof Ed of Lawyers, Inc., 4th edition edition.

Posner, R. A. (1972). A theory of negligence. 1(1):29–96. Publisher: [University of Chicago

Press, University of Chicago Law School].

Posner, R. A. (2014). *Economic Analysis of Law*. Aspen Publishing, 9th edition.

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

Shavell, S. (1987). *Economic Analysis of Accident Law*. Harvard University Press.

Shavell, S. (2004). *Foundations of Economic Analysis of Law*. Belknap Press: An Imprint of Harvard University Press.

Sobel, J. (1989). An analysis of discovery rules. *Law & Contemp. Probs.*, 52:133.

Xiong, W. and 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. Publisher: The University of Pennsylvania Law Review.

Appendices

A Proofs

Proof of Proposition 1. WLOG, 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 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:

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

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

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

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 $x_m = x_p = [\Pi']^{-1}(-(c_m + c_p)) = z(c_m + c_p)$. □

Proofs of Propositions 2,3 and 5. Propositions 2 and 3 are simply special cases of Proposition 5. Thus, in this section, we prove the latter proposition.

The planner'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 — a technical point about first order conditions. The first order conditions 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:

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

Now, suppose $c_m = c_m(c_p)$. It follows that $c_p = c_p(c_m)$. Then 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:

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

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 get a decreasing second best schedule, it suffices that $\frac{c_i g_i(c_i)}{G_i(c_i)} \cdot \frac{c_{-i} g_{-i}(c_{-i})}{G_{-i}(c_{-i})} < 1$.

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 $c_i < \hat{c}_i$ and that $x_i(c_i) < \hat{x}$ 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 $x_m(\underline{c}_m) = \overline{x} = x_p(\underline{c}_p)$. Now, since $x_p(c_p)$ is strictly decreasing on $[\underline{c}_p, \underline{c}_p + \varepsilon]$, it must be that $\Pr[x_p(c_p) \geq x_m]$ 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 $x_m(\underline{c}_m) = 0$. 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:

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

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 $c_p(\underline{c}_m) = \underline{c}_p$. Furthermore, $\lim_{\tilde{c}_m \to \infty} \phi(\tilde{c}_m) = -1$, and so 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:

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

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 $\hat{c}_i$.

Uniqueness follows if we can show that $\phi'(\tilde{c}_m) < 0$ for all $\tilde{c}_m > \hat{c}_m$. (If so, then $\phi(\tilde{c}_m) < 0$ for all $\tilde{c}_m > \hat{c}_m$, and so $\hat{c}_m$ must be the unique root.) Now, 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 $\frac{\hat{c}_i g_i(\hat{c}_i)}{G_i(\hat{c}_i)} < 1$ for at least one $i \in \{m, p\}$.

Suppose this condition is satisfied for both m and p.($*$) Then, by the unimodality of each G_i, it must be that for each i, $\frac{\tilde{c}_i g_i(\tilde{c}_i)}{G_i(\tilde{c}_i)} < 1$ for all $\tilde{c}_i > \hat{c}_i$. [We state this as a claim, and show that it is true, below.] But this ensures that $\phi'(\tilde{c}_m) < 0$ for all $\tilde{c}_m > \hat{c}$, which establishes uniqueness.

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 $c_i > \hat{c}_i$.

All that remains is to formalize and prove the claim used in step 5.

Claim: Suppose G has a unimodal distribution. If $\frac{cg(c)}{G(c)} < 1$, then $\frac{c' g(c')}{G(c')} < 1$ for all $c' > c$.

The proof is as follows: First, note that if $g(c_0) > \frac{G(c_0)}{c_0)}$ and g is increasing on $[c_0, c_1]$, then $g(c_1) > \frac{G(c_1)}{c_1)}$. (Similarly, if $g(c_0) < \frac{G(c_0)}{c_0)}$ and g is decreasing on $[c_0, c_1]$, then $g(c_1) < \frac{G(c_1)}{c_1)}$.) To see this, note that:

$$G(c_1) = G(c_0) + \int_{c_0}^{c_1} g(c)dc \leq G(c_0) + g(c_1)(c_1 - c_0)$$

and so:

$$\begin{aligned} \frac{G(c_1)}{c_1} &\leq \frac{G(c_0)}{c_1} + g(c_1)\left(1 - \frac{c_0}{c_1}\right) \\ &\leq g(c_1) - \frac{c_0}{c_1}\left(g(c_1) - \frac{G(c_0)}{c_0}\right) \\ &< g(c_1) \end{aligned}$$

which makes use of the fact that $g(c_1) - \frac{G(c_0)}{c_0} > 0$ since $g(c_1) \geq g(c_0) > \frac{G(c_0)}{c_0}$ by assumption.

Now, since $G(\underline{c}) = 0$, there exists $\varepsilon > 0$ s.t. $g(c) > \frac{G(c)}{c}$ for $c \in (\underline{c}, \underline{c} + \varepsilon)$. Moreover, this inequality will remain true as long as $g(c)$ is increasing. Hence, if $g(c') < \frac{G(c')}{c'}$ at some c' (which implies that $\frac{c'g(c')}{G(c')} < 1$), it must be that g is zero or decreasing at c'. But by unimodality, once g is decreasing, it must continue to decrease, and if so, the inequality $g(c) < \frac{G(c)}{c}$ is preserved for all $c > c'$. This completes the proof. $\square$

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

Next, we verify the comparative statics. Consider two distributions of costs, c_1 and c_2, with and let $\hat{c}_i$ satisfy $\hat{c}_i = 2E[c_i \,|\, c_i < \hat{c}_i]$ for $i \in \{1, 2\}$. Begin with scaling, $c_2 = \kappa c_1$. It follows that $\kappa\hat{c}_1 = 2E[\kappa c_1 \,|\, \kappa_1 c_1 < \kappa\hat{c}_1] = 2E[c_2 \,|\, c_2 < \kappa\hat{c}_1]$, and so $\hat{c}_2 = \kappa\hat{c}_1$. Moreover, $G_2(\hat{c}_2) = \Pr[c_2 < \hat{c}_2] = \Pr[\kappa c_1 < \kappa\hat{c}_1] = \Pr[c_1 < \hat{c}_1] = G_1(\hat{c}_1)$.

Next, consider a translation, $c_2 = c_1 + \kappa$. Notice that $2E[c_2 \,|\, c_2 < \hat{c}_1 + \kappa] = 2E[c_1 + \kappa \,|\, c_1 + \kappa < \hat{c}_1 + \kappa] = 2E[c_1 \,|\, c_1 < \hat{c}_1] + 2\kappa = \hat{c}_1 + 2\kappa > \hat{c}_1 + \kappa$. Hence $\frac{2E[c_2 \,|\, c_2 < \hat{c}_1 + \kappa]}{\hat{c}_1 + \kappa} > 1 = \frac{2E[c_2 \,|\, c_2 < \hat{c}_2]}{\hat{c}_2}$. Then, since $\frac{2E[c_2 \,|\, c_2 < \alpha]}{\alpha}$ is a decreasing function of α, it must be that $\hat{c}_2 > \hat{c}_1 + \kappa$. Moreover, this implies that $G_2(\hat{c}_2) \geq G_2(\hat{c}_1 + \kappa) = G_1(\hat{c}_1)$, where the first inequality is strict whenever $G_1(\hat{c}_1) < 1$. Finally, suppose c_2 is a mean preserving spread of c_1. Then, by the Rothschild and Stiglitz (1970) condition, $\int^c G_2(\chi)d\chi \geq \int^c G_1(\chi)d\chi$ for all χ. This implies that:

$$\frac{\int^{\hat{c}_1} G_2(c)dc}{G_1(\hat{c}_1)} \geq \frac{\int^{\hat{c}_1} G_1(c)dc}{G_1(\hat{c}_1)}$$

$$\frac{\int^{\hat{c}_1} G_2(c)dc}{G_2(\hat{c}_1)} > \frac{\int^{\hat{c}_1} G_1(c)dc}{G_1(\hat{c}_1)}$$

$$\hat{c}_1 - \frac{\int^{\hat{c}_1} G_2(c)dc}{G_2(\hat{c}_1)} < \hat{c}_1 - \frac{\int^{\hat{c}_1} G_1(c)dc}{G_1(\hat{c}_1)}$$

$$E[c_2 \,|\, c_2 < \hat{c}_1] < E[c_1 \,|\, c_1 < \hat{c}_1]$$

$$\frac{E[c_2 \,|\, c_2 < \hat{c}_1]}{\hat{c}_1} < \frac{E[c_1 \,|\, c_1 < \hat{c}_1]}{\hat{c}_1}$$

where the second inequality uses the fact that $G_2(\hat{c}_1) < G_1(\hat{c}_1)$. 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 $F'(x) = f(x)$. (This can be verified using integration by parts.) Then, since $\frac{E[c_1 \,|\, c_1 < \hat{c}_1]}{\hat{c}_1} = \frac{1}{2} > \frac{E[c_2 \,|\, c_2 < \hat{c}_1]}{\hat{c}_1}$, since $\frac{E[c_2 \,|\, c_2 < \hat{c}_2]}{\hat{c}_2} = \frac{1}{2}$, and since $\frac{E[c_2 \,|\, c_2 < \alpha]}{\alpha}$ is a decreasing function of α, it must be that $\hat{c}_2 < \hat{c}_1$. Moreover, this implies that $G_2(\hat{c}_2) < G_2(\hat{c}_1) < G_1(\hat{c}_1)$. □

Proof of Proposition 4. First, we 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:

$$\frac{\partial W}{\partial x(c_i)} = c_i + (1-\lambda)\int_{\underline{c}}^{c_i} \Pi'(\lambda x(c_{-i}) + (1-\lambda)x(c_i))g(c_{-i})dc_{-i} + \lambda\int_{c_i}^{\overline{c}} \Pi'(\lambda x(c_i) + (1-\lambda)x(c_{-i}))g(c_{-i})dc_{-i} \tag{5}$$

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 $\Pi'' > 0$, $\Pi'(\lambda x(c_i) + (1-\lambda)x(c_{-i}))$ becomes less negative as $x(\underline{c})$ increases. 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 $\int_{\underline{c}}^{\overline{c}} \Pi'((1-\lambda)x(c_{-i}))g(c_{-i})dc_{-i}$ is finite, there exists $\lambda' > 0$ s.t. $\frac{\partial W}{\partial x(\underline{c})} > 0$ for all values of $x(\underline{C})$ whenever $\lambda < \lambda'$. It follows that $x(\underline{c}) = 0$. But this implies that $x(c) = 0$ for all c, which cannot be. Hence, there must be some pooling of low-cost types.

Next, we characterize the breadth of the optimal pool. Using the same logic as in Proposition 3, we know that the threshold type $\hat{c}$ must be indifferent between pooling and separating. Hence, her care level of characterized by the first order condition (5):

$$\hat{c} + (1-\lambda)G(\hat{c})\Pi'(\hat{x}) + \lambda \int_{\hat{c}}^{\infty} \Pi'(\lambda\hat{x} + (1-\lambda)x(c))g(c)dc = 0 \tag{6}$$

where we make use of the fact that $x(c) = \hat{x}$ for all $c < \hat{c}$.

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:

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

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:

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

$$2E[c \,|\, c < \hat{c}] + G(\hat{c})\Pi'(\hat{x}) - 2[\hat{c} + (1-\lambda)G(\hat{c})\Pi'(\hat{x})] = 0$$

$$E[c \,|\, c < \hat{c}] - \frac{1-2\lambda}{2}G(\hat{c})\Pi'(\hat{x}) = \hat{c} \tag{8}$$

where the second line uses (6). This defines the location of the threshold $\hat{c}(\lambda)$.

If $\lambda = 0$, then using the fact that $\Pi'(\hat{x}) = -\frac{\hat{c}}{G(\hat{c})}$ when $\lambda = 0$, the condition reduces to $\hat{c}(0) = E[c \,|\, c < \hat{c}(0)] + \frac{1}{2}\hat{c}(0)$, which implies that $\hat{c}(0) = 2E[c \,|\, c < \hat{c}(0)]$. If $\lambda = 0.5$, then the condition reduces to $\hat{c}(0.5) = E[c \,|\, c < \hat{c}(0.5)]$, which can only be satisfied at $\hat{c} = \underline{c}$. Finally, as λ increases, the wedge between $\hat{c}$ and $E[c \,|\, c < \hat{c}]$ decreases, and this implies that $\hat{c}$ must be decreasing (since $\frac{\partial E[c \,|\, c<\hat{c}]}{\partial \hat{c}} < 1$. □

Proof of Corollary 1. Suppose $\hat{c} \geq \bar{c}$. This implies that $G(\hat{c}) = 1$, and $E[c \mid c < \hat{c}] = E[c]$. Now, By equation (7) in the Proof of Proposition 4, the optimal pooling standard satisfies $\Pi'(\hat{x}) = -2E[c]$, which implies $\hat{x} = z(2E[c])$. Substituting this into equation (8) gives:

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

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

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:

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

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:

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

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

$$\begin{aligned}
\frac{1-\lambda}{\lambda} \cdot \frac{c_m}{c_p} &< 1 \\
\lambda &> \frac{c_m}{c_m + c_p}
\end{aligned}$$

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 $x_m = x_p = z(c_m + c_p)$. $\square$

Proof of Proposition 6. See Shavell (1987). $\square$

B Online Appendix: A Model of Contracts

In this section, we briefly show how the model can be adapted to captured 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 and unimodal distributions $G_B(v_B)$ and $G_S(c_S)$. 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 $\overline{v}_B < \infty$.

B.1 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.

B.2 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 4.2.2. 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 7. *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:

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

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

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

Proposition 7 is analogous to Proposition 5. (Since the proof strategy is identical to the proof of Proposition 5, 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.