Torts, Expertise, and Authority: Liability of Physicians and Managed Care Organizations

Jennifer Arlen*

W. Bentley MacLeod**

The purpose of this paper is to examine optimal individual and entity-level liability for negligence when expected accident costs depend both on the agent's level of expertise and the principal's level of authority. We consider these issues in the context of physician and managed care organization (MCO) liability for medical malpractice. Under current law, physicians generally are considered independent contractors and hence MCOs are not liable for negligent acts by physicians. We find that the practice of reviewing the medical decisions of physicians affects their incentives to take care, which, in turn, implies that it is efficient for MCOs to be held liable for the torts committed by their physicians.

1. Introduction

The purpose of this paper is to explore the proper scope of physician and managed care organization (MCO) negligence liability for medical malpractice. Specifically, we determine how the exertion of control by the MCO over physician decision making affects the optimal damage award when a tort is the result of either physician or MCO negligence. It is shown that when the MCO is liable for all torts, including those committed by its physicians, then the standard rule for damages results in an efficient outcome. However, under current law MCOs generally are not liable for their own negligent treatment decisions. MCOs also are not liable for the negligence of their affiliated physicians because physicians affiliated with managed care plans generally

are deemed to be independent contractors and thus only they are personally liable for their torts. Our results imply that under current law patients are receiving suboptimal levels of care.

In contrast with traditional fee-for-service insurers, MCOs are not simply passive insurers whose role is limited to paying for any treatment provided. MCOs seek to affect care directly through the use of a "utilization review." Utilization review grants an MCO authority to review proposed treatments and to deny coverage for any treatment that the MCO deems to be either "experimental" or not "medically necessary." An MCO denying coverage for one treatment often does so with the understanding that it will pay for another. This decision as to what treatment to cover often determines the treatment the patient receives – particularly in the case of serious illness with expensive treatments.

Although MCOs influence medical treatment, they are generally insulated by federal law from liability for their own negligence in denying treatment coverage. Generally, they are insulated by state law governing entity-level liability from liability for negligence by their affiliated physicians. The doctrine of vicarious liability holds Principals liable for torts by their Agents but generally restricts liability to torts caused by Agents who are employees. Principals generally are not liable for torts caused by independent contractors – even if the independent contractor has insufficient assets to pay the tort judgment. Most MCOs hire physicians as independent contractors and thus they usually avoid liability for the negligence of their affiliated physicians. However, the MCOs' capacity to influence medical treatment – both directly and through their effect on physicians – presents the question of whether MCOs should be liable for either their own negligent treatment decisions or for physician negligence, and, if so, what is the optimal damage award?

Beginning with Kornhauser (1982) and Sykes (1984), the traditional analysis of entity-level liability is based upon the standard Principal-Agent model in which a risk-neutral principal offers an incentive contract to a risk averse Agent. They have shown that when the Agent has limited wealth it may be efficient to shift liability to the Principal, who then offers a contract that efficiently trades off risk against incentives to take care. In their analysis, only the Agent directly controls care; the Agent, moreover, is fully informed about the costs and benefits of his own actions and thus knows when he is negligent.

Grady (1988) discusses the possibility of inadvertent negligence, but treats this risk as exogenous and thus does not examine the impact of liability on incentives to reduce the likelihood of inadvertent error. Shavell (1992) examines an injurer's incentive to invest in information about the cost of his actions. In his model, injurers who invest in information become fully informed. Thus, when investment in information is optimal, optimal tort liability eliminates accidental negligence. Chu and Qian (1995) consider the consequence of incentives to collect information in a Principal-Agent context, where the Principal must invest in monitoring to reduce the likelihood of negligence. They find that strict liability results in the efficient outcome, yet negligence liability is

the standard rule for most torts. As Simon (1982) has shown, negligence is preferred to strict liability in the context of medical malpractice. In that case, Chu and Qian (1995) show that it may be optimal to lower the negligence standard or reduce entity level-liability.

Zeiler (2004) explicitly explores the consequence of the liability regime upon the contract that would be agreed upon between physicians and MCOs. Attention is restricted to the case in which the MCO can only contract over cost, and the physician is restricted to a binary choice - treat or not treat. Zeiler (2004) then finds that efficiency is enhanced when contracts are assumed to be publicly observed. In this case, she also finds, consistent with the earlier work of Kornhauser (1982) and Sykes (1984), that the liability regime has no effect upon the efficiency of the outcome.[1]

This paper extends this research by adding two new ingredients to the traditional Principal-Agent model. The first is to introduce the concept of "expertise." In the traditional analysis, the Agent selects care at the time she provides the service. At that point she trades off the cost of spending more time and effort on the case against the benefit of making a better decision. This analysis overlooks another aspect of treatment quality, namely the expertise of the physician in making a correct decision. This corresponds to the time and energy spent by the physician maintaining her skills and learning about new medical developments. Physician expertise is important because a physician does not necessarily know the expected costs and benefits of the available treatments. By investing in expertise, a physician can increase the likelihood that she will make a correct treatment choice, and correspondingly reduce her risk of inadvertent medical error.[2]

Second, this paper examines the impact of the Principal's use of authority to directly control the Agent's behavior on the efficiency of individual and entity-level liability. Given that the MCO reimburses the costs of treatment, the physician prefers to provide the patient with the treatment with the highest expected benefit, with little regard to costs. To counterbalance this effect, MCOs may require "utilization review" and refuse to cover the physician's recommended treatment in favor of a less expensive alternative treatment. The concern here is not with the details of the procedures involved in "utilization review," but rather we wish to explore how the tort system can complement the current system of utilization review to achieve a more efficient allocation of resources.[3] To examine this, we introduce a concept of "authority" to represent the likelihood that a utilization review leads to a treatment that is different from the one recommended by the physician.

The agenda of the paper is as follows. The next section discusses the institutional

---

[1] Proposition 6.

[2] This interpretation is particularly relevant for medical malpractice, where much of the effect of liability is upon diagnosis, as shown recently by Kessler and McClellan (2002)

[3] Dranove and Spier (2003) show that "utilization review" can enhance efficiency by improving the screening of cases, but do not explore the implications of their analysis for tort law. Malcomson (2004) explores the complications involved in writing an optimal contract for the services of a health gatekeeper whose job it is to select the apropriate physician, an issue we do not directly address here.

and legal background that motivates the addition of expertise and authority to the standard model. A key ingredient in this model is the assumption that neither authority nor physician expertise are contractible.[4] MCOs have the right to overrule physician decisions when the treatment is deemed "medically unnecessary," a term that is always open to interpretation.[5] Similarly, licensing boards set only minimum standards for physicians and patients have little information regarding physician activities that would enhance expertise beyond this minimum. The fact that authority and expertise are not contractible provides a nontrivial role for tort law.

The formal model is introduced in section 3. It adapts Aghion and Tirole (1997)'s theory of authority to the case of managed health care, and allows us to formally explore the impact of tort law upon the trade-off between the exercise of authority by the MCO and the acquisition of expertise by physicians. After a discussion of the various features of the model, we characterize the social optimum in section 3.3. Under the appropriate conditions, we find that it is efficient for the MCO to exercise some authority over treatment decisions, and that this authority decreases with physician ability. For very able physicians, a fee-for-service system is first-best. It is a feature of the model that the exercise of authority by the MCO unambiguously decreases quality, even though the exercise of some authority is part of an optimal allocation.

Ex ante individuals would be willing to decrease quality in exchange for a lower price. Of course, *ex post* individuals would always prefer the highest quality possible. This creates a problem when contracts are incomplete. Section 3.4 derives the equilibrium under the hypothesis that contracts are incomplete, markets are perfectly competitive, and there is no tort liability. The result is an inefficient equilibrium at which the level of authority may be too high and the level of expertise is too low.

In section 4 we examine the effect of individual negligence liability imposed for both physician negligence and negligent treatment decisions by MCOs. We find that physician expertise and MCO authority can be returned to their first-best levels with the introduction of tort law that allows patients to seek damages for injuries resulting from negligent treatments. In contrast with existing law, however, liability must be imposed for both physician and MCO negligence. In addition, the optimal damage rule is not set equal to harm, as in the traditional analysis, but is a complex function of the harm caused to the patient and probability that a negligent physician is detected.[6] Failure to

---

[4]   As Epstein and Sykes (2001) observe, if all actions were contractible then there would be little need for tort law.

[5]   The level of MCO authority is likely to be non-contractible because it is based on the notion of "medically necessary" treatment, which is not precisely defined. See Arlen and MacLeod (2003): 1970-1972 & note 148. Physician expertise also is likely to be non-contractible. Licensing boards can ensure physicians meet certain bare minimum standards, but neither licensing boards nor patient contracts regulate exactly how much effort and attention a physician puts into keeping abreast of the latest developments and improving her skills throughout her career. Gawande (2002) provides an insightful discussion of the importance of ongoing physician investment in expertise.

[6]   It is well known that the probability of detection can affect the optimal damage rule, and in some cases this implies that punitive damages may be efficient. For example, see Polinsky and Shavell (1998). We

hold MCOs liable for negligent treatment decisions results in both inefficient authority and inefficient physician expertise.

Negligence is efficient only if damages are set precisely equal to the optimal award. This result contrasts with that of the traditional model of negligence, where behavior is insensitive to excess damages because injurers can avoid liability by taking due care (See Cooter (1984)). In our model, the efficiency of negligence liability is very sensitive to excessive damages since each medical provider faces an omnipresent risk of inadvertent error, and thus medical providers inevitably face a risk of liability. Thus, excessive damages can distort behavior.

Section 5 extends the analysis to the case in which the MCO is able to use performance contracts. If the MCO and the physician cannot contract between themselves to impose reciprocal sanctions – for example, because of legal restrictions on such contracts – then tort damages must incorporate the cost to both the patient and the other medical care provider of inefficient authority or expertise.[7] In this case, the optimal damage rule is rather complex. If the MCO and the physician can contract to impose reciprocal sanctions on each other in the event of the other's negligence, then optimal damages only include the cost of negligence to the patient, consistent with the current legal standard.

Finally, assuming that damages are set equal to the cost of negligence to the patient, we consider whether liability for physician negligence is best imposed on the physician or the MCO. Consistent with Kornhauser (1982) and Sykes (1984), we find that MCO liability is equivalent to pure physician liability when physicians are not judgment proof, and is strictly preferred to individual liability when a physician's limited liability constraint is binding.[8] We find that MCO liability is efficient even though MCOs hire physicians as independent contractors. The final section of the paper contains a concluding discussion.

2. Institutional and Legal Background

The health care industry provides two basic services: medical care and insurance. At present, these services are provided through two different systems, pure indemnity (or fee-for-service) insurance and managed care organizations (MCOs).[9]

Under traditional indemnity insurance, insurers pay for all treatment costs (minus a

---

focus on the situation where negligent defendants may escape sanction because not all negligence results in *ex post* injury.

[7] Although Epstein and Sykes (2001) suggest informally that making the MCO and the physician personally liable for their own errors can result in the first-best, they do not explore the impact of MCO authority on optimal damages.

[8] A physician is judgment proof if she does not have sufficient assets to pay damages assessed against her.

[9] We use Managed Care Organization (MCO) to refer to any insurer which seeks to influence treatment choice, for example through utilization review. The term MCOs thus covers Health Maintenance Or-

5

deductible); physicians select and provide treatment without interference from insurers. Evidence suggests that fee-for-service generally results in excessively costly medical care because physicians and patients need not consider treatment costs in selecting between treatments.[10]

Medical care not only is high-cost, but treatment quality remains a concern. Evidence suggests that physicians often provide substandard, negligent, medical care.[11] Reducing medical negligence is difficult because physicians rarely provide substandard care knowingly.[12] Physicians generally err accidentally, often as a result of inadequate knowledge, training, or skill.[13] The traditional system vests physicians with primary authority over treatment decisions. Thus, improving treatment quality depends on inducing each individual doctor (or practice group) to obtain sufficient expertise to diagnose patients and assess all the available treatments.

Managed care organizations (MCOs) arose to address the problem of excessive health care costs, while also offering the potential to improve health care quality. They now dominate the medical insurer market.[14] Most MCOs control costs, at least in part, by requiring physicians to obtain prior approval for treatments through a process known as "utilization review." Under utilization review an MCO has the right to deny coverage for certain medical treatments before treatment is provided. Because medical care is complex and dynamic, MCOs do not pre-specify non-covered treatments in their contracts with patients, but instead assert the right to deny coverage for treatments that are either "experimental" or not "medically necessary and appropriate,"[15]

---

ganizations (HMOs) as well as ostensible fee-for-service insurers which employ utilization review. Our results also may apply to physician groups in California, which perform many of the same functions as MCOs.

[10] For example, Kessler and McClellan (1996) find that patient deductibles generally are capped at a given dollar amount and do not regulate choices between expensive treatments at the margin.

[11] Kohn, Corrigan, and Donaldson (2000) find that medical error results in countless injuries and approximately 98,000 deaths per year. For a discussion of other studies of error see, Arlen and MacLeod (2003), 1938-39.

[12] Patient care depends upon physicians maintaining ongoing investments in expertise. Medical care is a dynamic technology. Every day new treatments are invented and new evidence emerges of the effectiveness of existing practices. Gelijns, Zivin, and Nelson (2001), p. 914, find that in a single year, the Food and Drug Administration approved about 5,000 new and modified devices; each year approximately 35 percent of the 200 largest-selling prescription drugs are new.

[13] Greiner and Knebel (2003), p. 111, find that 20 to 50 percent of primary care physicians are not aware of (or are not using) new evidence on best medical practices. In addition, Krizek (2000), p. 1362, finds that 20 to 80 percent of surgical error is directly or indirectly caused by inadequate expertise (including medical personnel having inadequate knowledge or failing to employ knowledge they have, inadequate supervision of residents, and physician failure to update practice protocols). A single individual responsible for error could be identified in almost 38 percent of the cases of medical error (Id. at p. 1359).

[14] Glied and Zivin (2002) find that in 1997 only 2 percent of health plans conformed to traditional fee-for-service, while 98 percent of health plans are either managed care or fee-for-service with utilization review.

[15] The term "medically necessary and appropriate" is sufficiently vague to enable an MCO to categorize a treatment as not medically necessary even when most medical experts would disagree. See Korobkin (1999), p.31, for a discussion of the wide variation in insurance companies' views of what constitutes

and determine whether a treatment is covered only once that treatment is requested.[16]

Utilization review effectively gives MCOs substantial authority to determine what treatments they will pay for (and, in turn, what expensive treatments patients receive) after contracting with the patient. If properly implemented, this process can improve health quality by providing a centralized decision maker with the capacity to collect and analyze complex national data on best medical practices. Yet MCO authority may result in patients receiving suboptimal care either because the MCO explicitly denies coverage for expensive cost-justified treatment in favor of suboptimal treatment, or because the utilization review process itself injures the patient by delaying the provision of time-sensitive treatment.[17]

Evidence on the actual effect of utilization review suggests that while it has been effective at reducing treatment costs, its effects on expected treatment quality have been more variable. A review of empirical studies on MCOs' cost-control mechanisms suggests that mechanisms such as utilization review either lower treatment quality or, at best, leave it unchanged.[18] There is also evidence to suggest that MCOs have used their authority over which hospitals their patients use to direct patients to lower-cost, but also lower-quality, hospitals.[19]

The expected cost and quality of patient care thus depends directly on actions taken by both physicians and MCOs. The present system regulates treatment quality in part

---

"medically appropriate" or "experimental" treatments.

[16] See Hall and Anderson (1992), pp.1638-41. Most MCOs assert the right to determine coverage free from external review of treatment coverage denials. Increasingly, states now mandate independent external review of treatment denials. Yet even with external review, utilization review directly affects treatment. Many patients do not appeal denials (in part because some MCOs have gag clauses that preclude physicians from effectively objecting to MCO treatment denials). See, for example, Havighurst (2001). Second, patient appeals cannot reverse the negative effects of an initial treatment denial if the appeals process introduces sufficient enough delay to render the recommended treatment ineffective.

[17] The utilization review process itself may alter treatment even if the MCO wants to provide optimal care. MCOs necessarily make initial treatment coverage decisions based on summary information that is less nuanced and patient-specific than the information available to a physician. This may lead an MCO to deny medically necessary treatment because the MCO is less well-informed. MCOs cannot always remedy these initial errors through patient appeals because the appeal process may introduce sufficient delay that the patient can no longer benefit from the recommended treatment (or any other treatment). E.g., Cicio v. Vytra Healthcare, 208 F. Supp. 2d 288, 290 (EDNY, 2001), aff'd in part and vacated in part, 321 F.3d 83 (2nd Cir. 2003) (alleging that delay between the MCO's initial denial, and subsequent approval of, physician-recommended cancer treatment rendered the treatment ineffective); see Roark v. Humana, Inc., 307 F.2d 298, 303 (5th Cir. 2002) (alleging that Aetna's initial refusal to approve the arthritis drug Vioxx – which carries a low ulcer risk – forced the patient to take an alternative drug, which gave the patient bleeding ulcers, rendering the patient unable to take Vioxx or any other medication absorbed through the stomach).

[18] See Sullivan (1999) for a review of existing studies that suggest that MCOs' cost-containment measures result in patients getting either inferior or equal health quality. There also is evidence that both seriously ill and poor patients suffer worse outcomes under managed care. See Duggan (2002) and Aizer, Currie, and Moretti (2004) for some more recent evidence.

[19] See Arlen and MacLeod (2003) 1977 for further discussion of this problem.

through tort liability imposed for medical negligence.[20] A physician who provides negligent medical treatment – as measured by the customary medical care – faces liability in tort should that treatment injure a patient. Yet, MCOs generally avoid liability for both their own negligence in denying coverage for medically appropriate treatments and for negligence by their affiliated physicians.[21]

Patients are seeking the right to pursue two different types of claims against MCOs. The first are direct actions against MCOs for negligent coverage determinations that result in the patient being injured by the denial of medically necessary and appropriate medical care. At present, both state and federal laws impose impediments to these suits. Many states do not permit recovery for negligent refusal to provide insurance coverage, even when it results in the patient receiving negligent treatment. Moreover, in those states permitting such actions, patients' right to recover from MCOs under state law is precluded by the federal Employee Retirement Income Security Act of 1974 (ERISA).[22]

The second class of claims are indirect claims seeking to hold MCOs liable for negligence by their affiliate physicians. While ERISA is likely to pose less of a problem for these actions, state law governing entity-level liability generally insulates MCOs. Under state law governing entity-level liability, MCOs are not liable for physician negligence if the physician is an independent contractor and the MCO ensured that its subscribers knew that the physician was not an employee.[23] Most MCOs thus should be able to avoid liability for physician negligence because they do not actually employ physicians, but instead offer health care services through a network of independent contractors.

Congress is considering legislation that would significantly change the legal landscape for MCOs. Several proposals would amend federal law to permit state law tort actions against MCOs for personal injury or death against any person resulting from the denial of treatment coverage. Legislatures also are considering MCO liability for injuries resulting from physician negligence. This article evaluates whether imposing negligence liability for MCOs' negligent treatment coverage decisions or for physician negligence would improve the joint welfare of patients, MCOs, and physicians.

3. The Model and Preliminary Analysis

Consider a three-Agent model consisting of a patient, an insurer, and a physician. The risk-neutral and liquidity-constrained patient seeks two types of medical services:

---

[20] For a discussion of the limits of direct regulation of physician quality through licensing requirements and other regulations, see, e.g., Arlen and MacLeod (2003), notes 108, 109, 123, and 128.

[21] See Arlen and MacLeod (2003) for a more extensive discussion of the legal issues.

[22] Employee Retirement Income Security Act of 1974, 29 USC § 1001-1461 (1974).

[23] The MCO may be liable under ostensible agency if the MCO holds the physician out as an employee and the patient relies on this. MCOs can avoid liability under ostensible agency by making sure patients are informed that physicians are independent contractors and not employees.

medical treatment and insurance.[24] The patient seeks medical treatment from a physician who is the exclusive provider of medical services. Insurance is provided by a specialized MCO who reimburses the physician for treatment costs.

The MCO is assumed to operate in a competitive market making zero profits. Thus, the patient's premium, P, equals the MCO's expected costs of providing insurance. The MCO also contracts with the physician, agreeing both to pay all treatment costs and to pay the physician an *ex ante* wage of W. The physician is assumed to earn her reservation utility of U^0.

For any given illness, there are a variety of possible treatments, $t \in T = \{1, ..., n\}$, with different consequences for the patient and the MCO. The expected benefit of any given treatment is b_t, which is net of any expected adverse consequences, l_t. As b_t is an expectation taken over a distribution of potential outcomes resulting from treatment t, *ex ante* expectations can differ from *ex post* outcomes. The cost of providing the treatment is c_t, which is borne by the insurer. The concept of a "treatment" is very general. It includes differences in treatment type (for example, medication or surgery), as well as differences in how and when a given procedure is performed. For example, a procedure performed immediately is deemed a different treatment than the same procedure performed a week later if the delay affects the expected benefit (or costs) of the treatment.[25]

The patient does not have sufficient information to either diagnose himself or select treatment, and thus delegates authority to select treatment to a medical professional (either the physician or the MCO). A patient who selects MCO insurance vests the MCO with the right to determine who will select the treatment he receives. If the MCO does not assert authority over treatment choice, the physician selects the treatment, as discussed below.

The MCO determines its authority level based on the relative costs and benefits to it of authority. Vesting the physician with authority over treatment potentially raises a moral hazard problem because neither the physician nor the patient bears the cost of any treatment provided. The net expected benefit of physician-selected treatment also depends on the probability that the physician errs. To select the treatment correctly, the physician must be informed about both the proper diagnosis and the net benefit of available treatments. To obtain this information the physician must invest in expertise (both pre- and post-contract). This in turn determines the probability that she errs and inadvertently provides erroneous treatment. A physician who errs not only risks injury to the patient, but also potentially imposes costs directly on the MCO (to the extent that additional treatment is needed to remedy the error). The physician can reduce, but not eliminate, the probability that she errs by investing in expertise post-contract. As discussed in the previous section, it is assumed that it is not possible to contract directly

---

[24] The patient and the physician are assumed to be liquidity constrained to generate a market for insurance provided by someone other than the physician.

[25] Thus utilization review can affect treatment choice if it introduces delay that materially alters either expected patient outcomes or treatment costs.

over the level of expertise.

The MCO cannot directly control the physician's recommendation, but can alter the choice of treatment by asserting authority over treatment choice when it has sufficient information to suggest that a lower-cost treatment exists that is acceptable (given the loose constraints placed by its contract with the patient). In some cases, this treatment lowers costs while still providing the patient with the same benefits as informed physician treatment. In other cases, it lowers treatment costs by providing suboptimal quality treatment that injures the patient. The MCO determines its authority level *ex post* after contracting with the patient. This authority is formally defined by the probability that a decision is overruled, and is assumed to be unobservable and non-contractible.

The decision sequence is summarized as follows:

1. The patient contracts with an insurance company that is assumed to be in a perfectly competitive industry, and thus earns zero profits.

2. The insurance company then contracts with the physician, who is also assumed to be in a perfectly competitive market, with a default utility of U^0.

3. The MCO and the physician make non-contractible investment decisions: the level of authority, in the case of the MCO, and the level of expertise, in the case of the physician.

4. If the patient falls ill, she contacts a physician for treatment. For simplicity, this occurs with probability 1.

5. The physician recommends treatment, which may be overruled by the MCO in favor of an alternative treatment.

6. The patient is treated, and faces a chance of inadequate care.

7. Treatment outcomes are revealed. The patient files suit if care is inadequate, and damages are positive.

Physician Treatment Choice

Ex ante – at the moment of contracting – the patient would like to contract to receive the treatment that maximizes the total expected benefit of treatment minus treatment costs. When ill, however, the insured patient would like to receive the treatment with the maximum expected benefit regardless of cost, because he does not bear treatment costs. Thus, the patient would like to receive treatment t^*, where $t^* = \arg\max_{t \in T}\{b_t\}$. The expected benefit and cost of treatment t^* are given by b^* and c^*, respectively.

The patient delegates initial authority over treatment choice to the physician. The physician is assumed to care directly about the welfare of her patients, in addition to caring about any pecuniary rewards. Thus, the physician benefits directly from good treatment outcomes (and bears a direct cost of adverse outcomes). To capture this, it is assumed that the physician obtains a direct benefit from providing treatment t given

by αb_t, where α reflects either physician compassion, the impact of norms (e.g., the Hippocratic oath), or the effect of patient outcomes on physician reputation.[26] While treatment outcomes affect the physician, it is assumed that they affect her less than they affect the patient. In other words, $1 > \alpha \geq 0$. Thus, a physician who kills a patient suffers a loss, but the loss is less than that suffered by the patient himself. The absolute value of α is referred to as the physician's level of "compassion." A physician with $\alpha = 0$ is referred to as a non-compassionate physician.

Whenever $\alpha > 0$, the physician wants to provide the treatment that maximizes b_t. Thus, she endeavors to provide treatment $t^* = \arg\max_{t \in T}\{b_t\}$, which is the patient's *ex post* preferred treatment. Although the physician wants to provide treatment t^*, in practice she may fail to do so because she is "uninformed." Providing medical care is a complex task involving uncertainty about the patient's condition, as well as the nature of the available treatments and their expected outcomes. A physician who is "uninformed" about the patient's condition or the true net benefits of the available treatments may inadvertently err and provide erroneous medical treatment, given by $\hat{t}$.

The physician can reduce her risk of error by investing in "expertise." We focus on one particular type of expertise: post-contractual investments that affect whether the physician is correctly informed regarding the best treatment for any and all of her patients.[27] This expertise includes a physician's post-contractual efforts to improve her diagnostic skills, her understanding of the expected benefits of all available treatments, and her ability to provide treatments – for example, by doing rounds in a hospital, reading the latest medical journals, participating in continuing medical education, and attending conferences.

Expertise reduces the probability of accidental medical error. To model this, we employ the approach of Aghion and Tirole (1997) and assume that investment in expertise increases the probability that the physician is "informed" and able to accurately assess the costs and benefits of the available treatments. The probability that the physician is informed is given by e; the probability that she is uninformed (and errs) is given by $(1-e)$. The cost to the physician of acquiring expertise is given by $V_D(e)/\theta_D$, where θ_D represents the physician's innate ability and $V_D(0) = V_D'(0) = 0$, $V_D''(e) > 0$, and $\lim_{e \to 1} V_D(e) = \infty$. Expertise is assumed to be non-contractible, as it is either non-observable or non-verifiable. When selecting treatment, the physician knows her level of expertise – and thus her probability of error – but she does not know whether she is informed or uninformed in any particular case. Error, thus, is truly inadvertent.

It is assumed that erroneous treatment $\hat{t}$ on average produces lower gross – and net

---

[26] The parameter α may be employed to model physician reputation when the benefit to the physician of good treatment outcomes (and the cost to the physician of bad outcomes) is less than the direct benefit (or cost) to the patient of treatment.

[27] Patient care depends on physicians maintaining on-going investments in expertise. Medical care is a dynamic technology. Every day new treatments are invented and new evidence emerges on the effectiveness of existing practices. E.g., Gelijns, Zivin, and Nelson (2001), p.914, (In a single year, the Food and Drug Administration approved about 5,000 new and modified devices; approximately 35 percent of the 200 largest-selling prescription drugs are new each year).

– patient outcomes than expected treatment t^*, namely $\hat{b} < b^*$ and $\hat{b} - \hat{c} < b^* - c^*$, where $\hat{c}$ is the cost of erroneous treatment and $\hat{b}$ is the *expected* benefit of erroneous treatment. The expected benefit of the treatment provided by the physician is therefore $eb^* + (1-e)\hat{b}$, which is less than the expected benefit of the patient's preferred treatment t^*. Under these assumptions, injurious treatment costs $\hat{c}$ might be greater or less than than noninjurious treatement costs c^*.

Although the expected benefit of erroneous treatment is less than the expected benefit of the patient's preferred treatment ($b^* > \hat{b}$), a patient provided erroneous treatment $\hat{t}$ may, in fact, emerge unharmed. A treatment may be negligent *ex ante,* and yet not result in any legally recognizable *ex post* injury. For example, a physician who fails to treat a patient presenting symptoms of a serious life-threatening disease may provide negligent care *ex ante*, and yet not cause an injury *ex post* if, in fact, the patient's condition was temporary and did not need treatment.[28] To model the possibility of error that does not cause *ex post* injury, it is assumed that when a physician errs and provides treatment $\hat{t}$ there is a probability $\pi^D < 1$ that the error injures the patient and a probability $1 - \pi^D$ that the error does not injure the patient. The expected benefit of "injurious" error is given by b^e, where $b^e < b^*$. The expected benefit of non-injurious errors is assumed to be b^*. Thus, the expected benefit of erroneous treatment is given by $\hat{b} = \pi^D b^e + (1 - \pi^D)b^*$.

MCO Treatment Choice

The MCO agrees to provide the patient with necessary medical services in exchange for a fixed fee. The insurance motive is not explicitly modeled, rather it is assumed that the MCO reimburses the patient for all out-of-pocket *approved* medical expenses for a fixed insurance premium, given by P.

The MCO-physician relationship is subject to a moral hazard problem because the physician would like to select the treatment that maximizes patient outcomes without regard to cost.[29] The MCO, however, bears the full treatment costs of any approved treatment provided and thus, all else equal, would prefer that the physician provide low cost treatment.

The MCO regulates treatment choice by asserting authority over treatment selec-

---

[28]   For example, a physician is negligent if she fails to test or treat a patient with the classic symptoms of meningitis because bacterial meningitis is easily detected (and treated) and can kill a patient in a matter of days. Nevertheless, the negligent treatment will not injure any untreated patient lucky enough to have viral meningitis, which is self-limiting.

[29]   The moral hazard problem explored here is different from the traditional moral hazard problem, as in, for example, Holmstrom (1979) or Shavell (1979). In the standard problem, the Principal has an external measure of performance and makes compensation an increasing function of this performance measure to provide appropriate performance incentives. Here, the nature of medical treatment is such that critical information regarding the patient's illness and the evaluation of acceptable treatments come from the physician herself. With the physician in substantial control of the signal, the MCO cannot rely on this mechanism to control moral hazard.

tion through a process called utilization review.[30] The MCO reviews some portion of the physician's treatment recommendations and overrules the physician when its information indicates that a lower cost treatment is available that arguably satisfies the contractual requirement that the treatment be medically appropriate.[31] This treatment is given by t^0.

Although the MCO has broad discretion to overrule the physician, it does not always do so because asserting authority is expensive. As in Aghion and Tirole (1997), to assert authority, the MCO must invest in an information system at a cost $V_I(a)$, where $V_I(0) = V_I'(0) = 0$, $V_I''(a) > 0$, and $\lim_{a \to 1} V_I(a) = \infty$. This investment determines the probability that the MCO asserts authority to alter a decision by the physician. This probability is given by a. The MCO cannot credibly precommit to a particular level of authority in its contract with the patient, and thus determines its authority level post-contract, after the premium is paid.[32]

The MCO determines which treatment it will select post-contract, after the premium is paid. Absent intervening forces, the MCO thus would have an incentive to select the cost-minimizing treatment, even if suboptimal, because the MCO bears treatment costs but, post-contract, does not benefit directly from improved patient outcomes.[33] The MCO is constrained (imperfectly) from pure pursuit of cost-minimization in its treatment choice by the MCO-patient contract (and by reputation), and thus selects treatment t^* in some cases. In other cases, the MCO selects the cost-minimizing treatment. In some cases the cost-minimizing treatment provides the best patient outcomes ($t^* = t^0$) (for example, providing aspirin to a heart attack victim). Nevertheless, in other cases MCO authority results in the patient receiving suboptimal treatment either because the MCO is not fully informed about the patient's condition or because its assertion of authority introduces sufficient enough delay that it results in the patient receiving suboptimal treatment.

To model this, it is assumed that the MCO asserts authority when it can recommend a treatment t^0 that lowers treatment costs relative to the costs of *expected* physician-

---

[30] The dual moral hazard problem between the physician and the MCO regarding treatment costs and the physician and the patient regarding expertise (and potential treatment) distinguishes our analysis of entity-level liability from Kornhauser (1982) and Sykes (1984).

[31] In fact, in actual practice, MCOs can, and do, deny coverage for treatments that are "experimental" even if no alternative treatment exist.

[32] MCOs cannot contract over authority by pre-specifying all approved treatments because there are too many illnesses and too many treatments. Also, too many patient-specific factors are relevant to treatment choice. This is compounded by the dynamics of medical technology that make it impossible for an MCO to precommit to a particular treatment for each and every illness. Moreover, MCOs cannot share explicit utilization review guidelines with physicians or patients for fear that physicians will alter the information they provide to MCOs in order to help patients obtain coverage.

[33] Market forces (e.g., reputation) may ensure that the MCO obtains some post-contractual benefit from patient outcomes. Our results do not turn on the assumption of zero-benefit. They depend only on the milder assumption that the MCO obtains less post-contractual benefit from good patient outcomes than does either the patient herself or the physician.

selected treatment: thus, $c^0 < ec^* + (1 - e)\widehat{c}$.[34] When the MCO asserts authority to select t^0, with probability $1 - \pi^I$, its substitute treatment provides the patient with the same expected benefit as informed-physician treatment, b^*, but at lower cost. With probability π^I, treatment t^0 lowers expected costs to c^0, but injures the patient by providing suboptimal benefits b^e, equivalent to injurious erroneous physician treatment. Thus, the expected benefit from the MCO decision is $b^0 = \pi^I b^e + (1 - \pi^I) b^*$, and satisfies $b^* > b^0 > \hat{b} > b^e$.

To analyze the general case, it is assumed that the expected benefit of MCO-selected treatment, $b^0 = \pi^I b^e + (1 - \pi^I) b^*$, may be greater than or less than the expected benefit of physician-selected treatment, $eb^* + (1-e)\hat{b}$, depending in part on physician expertise. Where the MCO asserts authority, we assume that the patient receives the MCO-selected treatment, even if $b^0 < eb^* + (1 - e)\hat{b}$. Adhering to the MCO's choice is both *ex ante* and *ex post* rational for the patient, provided that the benefit of MCO-approved treatment b^0 exceeds the *net* benefit to the patient of recommended treatment minus treatment costs, given that the patient need not pay for covered treatment but must pay the entire cost of uncovered treatment herself: $b^0 > eb^* + (1-e)\hat{b} - \{ec^* + (1 - e)\hat{c}\}$.

The Social Optimum

The expected payoffs of the MCO, the physician, and, the patient can now be defined as a function of the main choice variables: the level of expertise of the physician (the probability e, of being informed) and the level of authority exerted by the MCO (the probability a that the MCO overrules the physician's decision). The direct expected benefit to the patient of treatment is $B(a, e)$, and is defined by:

$$B(a, e) = a \cdot b^0 + (1 - a) [e \cdot b^* + (1 - e)\hat{b}].$$

The corresponding expected cost of treatment is $C(a, e)$ and is defined by:

$$C(a, e) = a \cdot c^0 + (1 - a) [e \cdot c^* + (1 - e)\hat{c}].$$

The patient has income I, and pays a premium P to the MCO, and hence has utility:

$$U_P(P, a, e) = I - P + B(a, e).$$

The MCO receives payment P as income, pays compensation W to the physician, and incurs investment cost $V_I(a)$ and $C(a, e)$ for any additional medical services ordered by either the physician or the MCO. The MCO's payoff is:

$$U_I(P, W, a, e) = P - W - C(a, e) - V_I(a).$$

Finally, the physician earns nonpecuniary benefits $\alpha B(a, e)$ from treating the patient, and pecuniary benefits W, and bears cost $V_D(e)/\theta$ to maintain a level of expertise e.

---

[34] When the MCO decides whether to assert authority, it does not know the cost of physician-selected treatment with certainty. The actual cost of physician-selected treatment will differ from c^* if the physician erred in her diagnosis and treatment selection.

The physician's utility is given by:

$$U_D(W, a, e) = W + \alpha B(a, e) - V_D(e)/\theta.$$

Both the MCO and the physician are assumed to be in competitive markets, with the profits of the MCO normalized to zero, while the alternative utility of the physician is U^0.

The optimal allocation under the hypothesis that expertise and authority are contractible maximizes the patient's utility subject to the constraints that both the MCO and the physician receive at least their alternative payoffs:

$$\max_{P,W,a,e} U_P(P, a, e) = I - P + B(a, e), \tag{1}$$

subject to:

$$U_I(P, W, a, e) = P - W - C(a, e) - V_I(a) \geq 0, \tag{2}$$

$$U_D(W, a, e) = W + \alpha B(a, e) - V_D(e)/\theta \geq U^0. \tag{3}$$

At the optimum, the individual rationality constraints 2 and 3 are binding. The individual rationality constraint for the physician implies $W = V_D(e)/\theta - \alpha B(a, e) + U^0$, while for the MCO one has $P = W + C(a, e) + V_I(a)$. Substituting these into the payoff for the patient, one obtains the following expression for net social welfare as a function of authority and expertise:

$$SW(a, e) = I + (1 + \alpha) B(a, e) - C(a, e) - V_I(a) - V_D(e)/\theta - U^0. \tag{4}$$

Differentiating, we can solve this problem using reaction functions. Given a level of authority a the optimal level of expertise $E(a)$ solves:

$$V_D'(E(a))/\theta = (1 + \alpha)B_e - C_e \tag{5}$$

where

$$B_e(a, e) = (1 - a)(b^* - \hat{b}), \tag{6}$$

$$C_e(a, e) = (1 - a)(c^* - \hat{c}). \tag{7}$$

Observe that expertise increases the net benefit to the parties of treatment whenever authority is less than perfect ($a < 1$): $(1 + \alpha)B_e - C_e > 0$. Since this expression is independent of expertise, then equation (5) uniquely defines expertise as a function of authority. Given that the right-hand side of (5) is decreasing in authority, and $V_D'' < 0$, then expertise decreases with authority. In addition, the reaction curve $E(a)$ shifts upwards with an increase in physician ability. This is illustrated in figure 1.
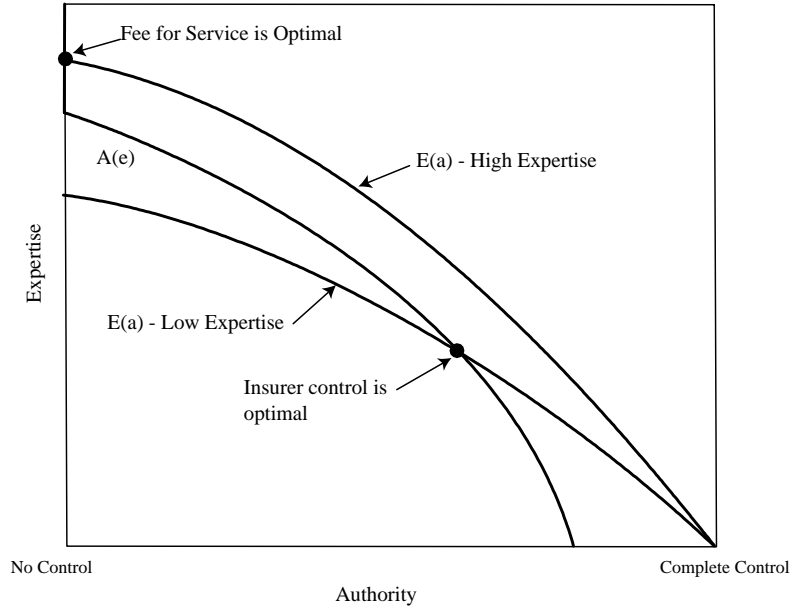
15

Figure 1

The optimal level of authority as a function of expertise, $A(e)$, solves:[35]

$$V_I'(A(e)) = \begin{cases} (1+\alpha)B_a - C_a, & \text{if } (1+\alpha)B_a - C_a > 0, \\ 0, & \text{if not.} \end{cases} \tag{8}$$

where

$$B_a(a,e) = b^0 - \{eb^* + (1-e)\widehat{b}\}, \tag{9}$$
$$C_a(a,e) = c^0 - \{ec^* + (1-e)\hat{c}\}. \tag{10}$$

Observe that when $e = 0$, then $(1+\alpha)B_a - C_a = (1+\alpha)\left(b^0 - \hat{b}\right) - \left(c^0 - \hat{c}\right) > 0$, while it is less than zero when $e = 1$. Therefore, when expertise is low, it is optimal to exercise some authority, and the level of authority satisfies the first line of equation (8). When expertise is sufficiently high that $(1+\alpha)B_a - C_a \leq 0$, then exercising no authority is efficient, as given by the second line of equation (8). This is illustrated in

---

[35] Notice that the level of ability does not affect the reaction curve for authority, while b^0 and c^0 do not affect the reaction curves for expertise, and hence we can view these parameters as primitives describing the behavior of the physician and the insurer, respectively. In particular, notice that increasing physician ability shifts reaction curve $E(a)$ up, and hence for sufficiently high θ it is the case that it is efficient to have no insurer control, while the converse is true when θ is low.

figure 1. This figure also illustrates the solution to program (1) by the intersection of the two reaction curves for authority and expertise.

Notice the effect of physician ability on the optimal solution. When physician ability is low, the expected quality of physician treatment is lower, and the optimal solution entails some positive level of authority by the MCO. When physician ability increases with all other parameters held fixed, optimal physician expertise (and expected treatment quality) increases and the optimal level of authority falls. When expertise is sufficiently high, then exerting no authority is optimal and fee-for-service insurance is preferable to MCO insurance.

The term "ability" needs to be interpreted broadly as measuring the relative cost to the MCO and physician of making informed treatment decisions. The modern rise of MCO authority is, in part, a response to technological improvement, particularly in information technology, that allows health care providers to compile information and respond in a timely fashion to treatment requests – something that would not have been possible before the advent of modern computer technology. For frequent illness where the MCO has extensive data, and where the optimal treatment varies little across patients, it is likely to be efficient for the MCO to exert authority (when it has optimal incentives). In contrast, it will often not be optimal for the MCO to assert authority with respect to illnesses where the physician can more accurately assess optimal treatment at lower cost – as is likely when the determination of the optimal treatment depends heavily upon patient-specific characteristics that the MCO cannot easily assess from afar.

Equilibrium with Incomplete Contracts

As discussed above, it is not possible for the patient and the MCO to contract *ex ante* over what treatments the MCO or the physician will select for any given illness once the patient gets ill. To model this, it is assumed that at stage 3 in the decision sequence the MCO and the physician simultaneously choose the level of authority and expertise, respectively. Neither the MCO nor the physician can observe the actual *ex post* choice of the other, but each has rational expectations regarding these choices. The patient cannot observe the actual *ex post* choice of either the MCO or physician, but has rational expectations regarding these choices.

Therefore, the optimal incomplete contract is given by the solution to the following problem:

$$\max_{P,W,a,e} I - P + B(a,e). \tag{11}$$

subject to:

$$P - W - C(a, e) - V_I(a) \geq 0, \tag{12}$$

$$W + \alpha B(a, e) - V_D(e)/\theta \geq U_D^0, \tag{13}$$

$$a \in \arg \max_{a' \in [0,1]} P - W - C(a', e) - V_I(a'), \tag{14}$$

$$e \in \arg \max_{e' \in [0,1]} W + \alpha B(a, e') - V_D(e')/\theta. \tag{15}$$

As in the Principal-Agent literature (e.g., Holmstrom (1979)), we can model contract selection with non-contractible authority and expertise as an optimization problem with individual rationality constraints, (12) and (13), and incentive compatibility constraints, (14) and (15), for the MCO and the physician, respectively. As in the standard Principal-Agent model, this approach implicitly supposes that when the MCO or the physician is indifferent between two actions, then the action that is favorable to the patient is selected.

This problem can be solved as follows. Observe that once the contract has been signed, the MCO and the physician select authority and expertise simultaneously, under the hypothesis that they have correct expectations regarding the choice of the other party. In essence, the contract defines a two-player game between the MCO and the physician, with the outcome given by the Nash Equilibrium of this game. The solution can be found by characterizing the reaction function for each party, and then finding the equilibrium.

Let $A^c(e)$ denote the MCO's optimal choice of authority given the physician's expertise e. This function satisfies the first-order conditions for (14):

$$-C_a(e) = V_I'(A^c(e))$$

where $C_a(e) = c^0 - \{ec^* + (1-e)\hat{c}\}$, as given by (10). Similarly, the physician's optimal choice of expertise, given the level of authority, is denoted by $E^c(a)$, and satisfies the corresponding first-order condition for (15):

$$\alpha B_e(a) = V_D'(E^c(a))/\theta$$

where $B_e(a) = (1-a)(b^* - \widehat{b})$, as in (6).

Observe that the first-order conditions do not depend upon either the price P nor the wage W, and hence, the choice of authority and expertise is independent of the outside options for the MCO and the physician. Formally, the solution to the post-contract game is characterized as follows:

Definition A *Nash Equilibrium* for the medical services game is a pair $\{a(\omega), e(\omega)\}$ solving:

$$\begin{aligned} a(\omega) &= A^c(e(\omega)), \\ e(\omega) &= E^c(a(\omega)), \end{aligned}$$

18

where ω is the vector of exogenous parameters.

This solution is illustrated in figure two. Observe that at the incomplete-contracts equilibrium with no sanctions, the MCO invariably asserts authority, *regardless* of the level of expertise, even if the fee-for-service system would be the first-best. Thus, the MCO asserts authority to reduce expected treatment costs even when doing so reduces the expected net benefit of medical care: $(1+\alpha)B_a - C_a < 0$. The slope of the MCO's reaction function depends upon the sign of $c^* - \hat{c}$. When this is positive, namely when informed physicians provide more costly treatment than uninformed physicians, then in equilibrium high-ability physicians face more control. This is because the gain in cost savings is greatest with this group. In contrast, when $c^* < \hat{c}$ the MCOs exert more control over lower-quality physicians.

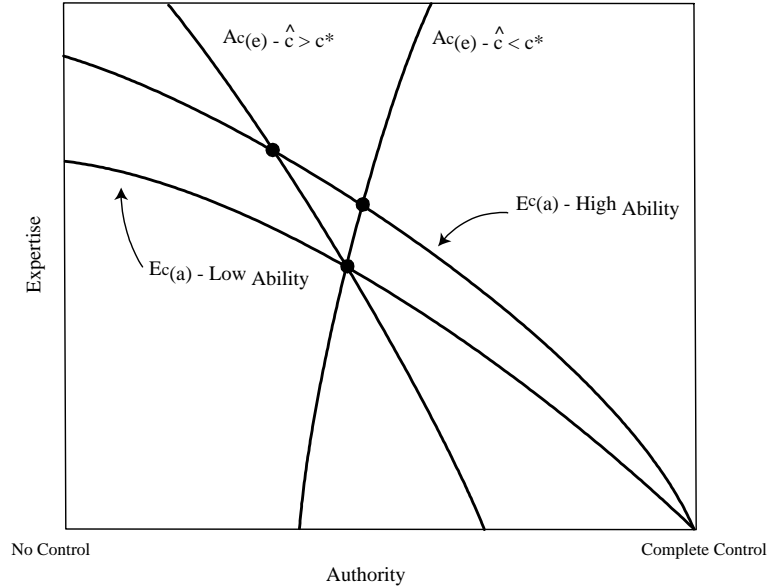
Figure 2

The properties of the complete solution to the provision of medical services under incomplete contracts are summarized in the next proposition.

Proposition 1   *There is a solution to program* 11 *such that every equilibrium* (a^c, e^c)

has the properties:

1. *Both authority and expertise are strictly positive, but there is less-than-perfect control, $a^c, e^c \in (0,1)$.*

2. *At an equilibrium $SW_e\left(a^c, e^c\right) > 0$, and hence increasing expertise at an equilibrium always increases social welfare. Welfare is also increasing with physician ability θ.*

3. *$SW_a\left(a^c, e^c\right) < 0$ if and only if $b^0 < \{eb^* + (1-e)\,\widehat{b}\}$, decreasing authority at an equilibrium increases social welfare whenever the benefit from MCO treatment is less than the expected benefit of physician-selected treatment.*

Proof. See Appendix. ∎

If authority and expertise are non-contractible, then the equilibrium absent liability is suboptimal notwithstanding the assumption that all parties know each other's payoff functions and accurately predict each other's behavior in equilibrium. Behavior by both the physician and the MCO is suboptimal absent sanctions because each affects care post-contract, and does not bear the full impact of their actions on either each other or on the patient.

The conclusion that the no-sanction equilibrium is inefficient, even when patients accurately predict the parties' behavior, contrasts with the result of the classic model of accidents. In that model, when injurers and victims are in a market relationship the market alone can induce optimal caretaking when victims accurately predict the risks imposed upon them.[36] Proposition 1 illustrates that despite accurate customer expectations and competitive markets, if expected accident costs depend upon non-contractible, post-contractual actions by the injurer, then the result may be an inefficient allocation of resources. In such situations, a victim's *ex ante* ability to accurately "price" expected accident costs is not sufficient to ensure that, after the contract price is paid, the injurer faces optimal incentives to take victim costs into account. In this case, the injurer will take too little "care," even though the parties jointly would be better off if the injurer could make a credible *ex ante* commitment to take due care.

4. Optimal Damages with Individual Liability

This section considers whether social welfare could be improved through the use of sanctions for medical malpractice. Only negligence liability is considered because this rule currently governs medical malpractice cases, and it is the basis of the leading proposals in Congress for MCO liability. This section considers a regime of individual liability for negligence under which the physician and the MCO are governed by a negligence liability rule as to their treatment decisions, but neither is liable for negligent treatment provided by the other. Thus the physician is potentially liable for treatments she selects and provides. The MCO, in contrast with existing law, is potentially liable for treatments that it selects.[37]

---

[36] See Shavell (1980), pp. 4-5, pp. 20-22, and Sykes (1984), p. 1256.

[37] This is also the rule suggested by Epstein and Sykes (2001) to address the problem of negligent acts committed by the MCO.

Under negligence liability, a medical care professional is liable if she selected a "negligent" treatment that injured the patient. Consistent with existing law, it is assumed that negligence liability depends upon the quality of the treatment provided, and not the quality (expertise) of the medical provider. We also assume that negligence is determined with respect to medical custom. Thus, a treatment is negligent if it provided the patient with lower expected benefits than the "customary" treatment provided by an informed physician (t^*). It is assumed that courts can assess whether the medical professional was negligent, but that they cannot determine optimal expertise or authority.[38]

Negligence is actionable only if the patient is injured. Thus, an uninformed physician is not liable every time she provides erroneous treatment $\hat{t}$; instead she is liable only if the treatment injured the patient (which occurs with probability π^D). The MCO faces potential liability only when it selected negligent treatment that injured the patient, which occurs with probability π^I.[39]

The present analysis extends the traditional model of negligence by permitting formal consideration of accidental, or inadvertent, negligence, the probability of which can be reduced, but not eliminated, through expertise. The traditional model of accidents assumes that injurers (here, medical professionals) know when they are being negligent. Thus, they will not be negligent if damages are set sufficiently high to induce due care (see Shavell (1980)). This framework does not appear to capture an essential feature of medical malpractice because physicians rarely knowingly decide to provide substandard care. The present model permits consideration of such errors through the assumption that even compassionate physicians can err and injure a patient with probability $(1-e)\pi^D$.[40]

The present analysis examines negligence liability relative to a system where there is no intervention at all. The relative merits of negligence and quality regulation are not formally considered. Nevertheless, while quality regulation has an important role to play, there still may be benefits from providing additional incentives via the tort system because regulations must be set in advance, and cannot dynamically adjust to new and changing conditions. Thus, the level of expertise considered here should be interpreted as the amount of expertise *above* the minimum level set by regulatory agencies.

---

[38] In order to focus on expertise and authority, we do not examine the impact of tort liability on treatment choice. We evaluate the role of medical negligence, accepting the constraint that due care is based on t^*. Yet even when due care is set too high (at t^*), we find that the tort system can improve the parties' joint welfare.

[39] In order to focus on the authority decision, we assume that the probability that the MCO is negligent if it asserts authority is exogenously given by π^I. In other words, we focus on patient injuries that result from the assertion of authority itself. MCO authority may injure the patient even if the MCO wants to select t^*, either because the MCO is less well-informed than the physician (and does not know it), or utilization review delays time-sensitive treatment. Extending the model to permit MCOs to intentionally select negligent treatments absent tort liability would strengthen our argument for imposing liability on MCOs.

[40] See text preceding note 1 discussing the prior literature on inadvertent error and learning about risk.

Optimal Damages for Individual Negligence

The first question to be addressed is the optimal level of damages that induce the MCO and the physician to select the socially optimal levels of authority and expertise, respectively. Let the expected damage award for physician negligence be L^D, and the damage award for MCO negligence be L^I. It is assumed that the patient, MCO, and the physician all know the current liability rule before entering into a contract. Accordingly, the optimal incomplete contract under individual liability solves the following program:

$$\max_{P,W,a,e} I - P + \left\{ B\left(a,e\right) + (1-a)(1-e)\pi^D L^D + a\pi^I L^I \right\}$$

subject to:

$$P - W - C\left(a,e\right) - a\pi^I L^I - V_I\left(a\right) \geq 0$$

$$W + \left\{ \alpha B\left(a,e\right) - (1-a)(1-e')\pi^D L^D \right\} - V_D\left(e\right)/\theta \geq U_D^0$$

$$a \in \arg\max_{a' \in [0,1]} P - W - C\left(a',e\right) - a'\pi^I L^I - V_I\left(a'\right)$$

$$e \in \arg\max_{e' \in [0,1]} \left\{ \alpha B\left(a,e'\right) - (1-a)(1-e')\pi^D L^D \right\} - V_D\left(e'\right)/\theta.$$

This problem is the same as the incomplete contract problem in the previous section, with the MCO facing an additional cost L^I whenever it exerts authority and injures the patient, and the physician facing an expected liability L^D when she provides negligent (uninformed) treatment that injures the patient.

In the case of the MCO, the marginal impact of an increase in liability on profits is $-a\pi^I$ and hence increasing its liability decreases the authority it exerts (as long as $a > 0$), while increasing the liability of the physician results in an increase in expertise. Also notice that the marginal impact of expertise decreases with authority.

Proposition 2 *The following expected damage rules result in the optimal level of authority and expertise:*

$$L^D = (b^* - b^e) + \frac{(\hat{c} - c^*)}{\pi^D}, \tag{16}$$

$$L^I = (1+\alpha)\left\{ (b^* - b^e) - \frac{(1-e^*)\pi^D(b^* - b^e)}{\pi^I} \right\}, \tag{17}$$

where e^* is the optimal level of expertise.

Proof. See Appendix. ∎

Imposing liability for both MCO and physician negligence increases the parties' joint welfare relative to the no-sanction equilibrium by inducing optimal expertise and

authority. Liability enhances the parties' welfare by enabling the physician and the MCO to credibly commit to undertaking optimal behavior: the promise to invest optimally in expertise and to assert optimal authority is credible because the imposition of optimal liability renders optimal behavior *ex post* incentive compatible for the MCO and the physician.

Proposition 2 shows that optimal damages for medical negligence place more exacting requirements on courts than the standard analysis of accidents would suggest. In the standard model, negligence liability forms a lower bound on damages because any potential injurer can completely avoid liability by taking due care (see Cooter (1984)). Thus, excessive damages do not distort behavior.[41] In this model, damages must be set precisely at the optimal level, L^D and L^I, because excessive damages would cause physicians to over invest in expertise.[42]

Proposition 2 also shows that the standard rule of damages equal to the plaintiff's harm is not necessarily efficient in this agency context. As is well-established, the goal of tort liability is to internalize external costs. In some cases this is accomplished by setting damages equal to the victim's loss. This does not always apply in an agency setting when both the Principal (the MCO) and the Agent (the physician) take actions that affect the quality of the service to be provided. In this case, in order to induce optimal expertise when the MCO does not sanction the physician itself, damages for physician error must equal the expected harm to both the patient $(b^* - b^e)$, and the MCO $(\hat{c} - c^*)$, arising from the provision of uninformed treatment.[43]

Thus, the patient's damage award must be adjusted by the term $\frac{(\hat{c}-c^*)}{\pi^D}$, which represents the expected difference in costs divided by the probability of a harm being committed when the physician is uninformed.[44] Expected optimal damages exceed the patient's expected harm $(b^* - b^e)$ if physician negligence increases expected treatment costs $(\hat{c} > c^*)$ and are less than the patient's expected harm if physician negligence

---

[41]   In the traditional framework, excessive damages can distort behavior if courts err when determining negligence. See Craswell and Calfee (1986). Our analysis shows the importance of accuracy in assessing damages even when courts accurately assess liability.

[42]   To use Professor Robert Cooter's terminology, negligence liability operates as a "sanction" in its regulation of treatment choice, but operates to "price" physician expertise. See Cooter (1984).

[43]   Observe that the component of damages targeted to the patient's harm, $(b^* - b^e)$, is based on the expected *ex ante* cost of receiving injurious erroneous treatment, and not on the *ex post* value of the harm suffered. This damage measure is equivalent to the harm suffered if the patient necessarily would have recovered fully if provided treatment b^* and was destined to suffer injury from treatment $\hat{t}$. Damages are less than the *ex post* harm suffered, however, if treatment t^* entailed a risk of injury even if properly performed (as is the case with any surgery). In this case, damages for medical error should be discounted for the expected harm the patient could have suffered from nonnegligent care (e.g., the background risk of death from any surgery).

[44]   The damage for increased treatment cost is adjusted by the probability that the physician harms the patient on the assumption that negligent treatment invariably affects expected treatment costs, but this negligence is actionable only if the treatment injures the patient (which occurs with probability π^D). The use of a multiplier to adjust for the probability that injurious negligence (to the MCO) may not result in a lawsuit is consistent with Becker (1968), Cooter (1982), and Polinsky and Shavell (1998).

reduces expected treatment costs ($\hat{c} < c^*$). This result illustrates the ambiguous effect that utilization review can have upon optimal damages. The standard rule is that the physician reimburses the patient for the harm caused, namely $(b^* - b^e)$.[45] This result shows that depending upon the cost structure, the optimal damage may be greater or less than the standard rule now used in court.

The proposition also shows that if the physician is held liable for her treatment decisions, then the MCO must also be held liable for its negligent treatment decisions made pursuant to utilization review in order to implement the optimal allocation. To induce optimal MCO authority, damages for MCO negligence must be based on the expected direct harm to the patient and the physician of the patient receiving injurious treatment from the MCO instead of the expected benefit of physician-selected treatment. As physician-selected treatment includes a risk of error, optimal damages thus equal the cost to the patient and the physician of the patient receiving injurious treatment instead of nonnegligent treatment $(1 + \alpha)(b^* - b^e)$, adjusted to reflect the probability-weighted expected cost to them of injurious physician error:

$$(1 + \alpha) \frac{(1 - e^*) \pi^D (b^* - b^e)}{\pi^I}.$$

Thus, in general, the MCO should face damages that are *less* than the harm caused by providing injurious treatment instead of nonnegligent treatment.

Cost-Sharing Contracts

In practice, the MCO is not limited to the use of authority to affect physician decision-making. The MCO also can make physician reimbursement dependent upon treatment costs. For example, the MCO may pay the physician a lump sum for each patient, but only reimburse her for a portion of the cost of any treatment provided. In this section we briefly consider the effect of cost-sharing on optimal damage rules.

Initially, suppose that the *treatment decision* does not depend upon the cost-sharing rule. That is, the informed physician is assumed to choose the treatment that is in the best *ex post* interests of the patient. Let the amount of cost sharing be represented by the parameter $\beta \in [0, 1]$, where the MCO reimburses $(1 - \beta)$ fraction of the costs, and the physician pays the rest. A straightforward extension of proposition 2 implies the following corollary:

Corollary 3 *When the physician pays a fraction β of the costs of treatment, then, assuming the negligence rule constrains an informed physician to select treatment t^*,*

---

[45]   Observe that the costs $\hat{c}$ and c^* are paid by the MCO at the time service is provided. Hence, it would not normally be part of the award. Costs of further treatment and maintanance not paid by the MCO would be part of a tort award, and would be included in the calculation of b^e.

the optimal damage rules are:

$$\bar{L}^D = (b^* - b^e) + \frac{(1-\beta)(\hat{c} - c^*)}{\pi^D},$$

$$\bar{L}^I = (1+\alpha)\left\{ (b^* - b^e) - \frac{(1-e^*)\pi^D(b^* - b^e)}{\pi^I} \right.$$

$$\left. + \beta\frac{\left(c^0 - (e^*c^* + (1-e^*)\hat{c})\right)}{\pi^I} \right\}$$

where e^* is the optimal level of expertise.

Observe that as the share of costs borne by the physician increases, this moves the optimal liability rule in the direction of the current law. In the extreme case in which the physician bears all treatment costs, liability is simply $(b^* - b^e)$, the harm to the patient due to negligent treatment. Under our maintained hypothesis that MCO treatment costs are lower, this unambiguously *decreases* optimal MCO liability relative to no cost sharing.

The Effect upon Treatment Decision

Cost-sharing not only affects authority and expertise, it also alters the physician's treatment decision. Though cost-sharing is often lauded as a way to improve physician treatment decisions, in fact it does not necessarily do so. When the MCO fully insures the patient and the physician against costs, an informed physician selects overly expensive treatments because she only considers the *ex post* benefit from treatment, and ignores treatment costs: $t^* = \arg\max \alpha b_t$. When there is cost sharing, *ex post* the physician's benefit from a treatment decision is $\alpha b_t - \beta c_t$. Accordingly, absent liability, cost sharing will induce an informed physician to select the treatment:

$$t(\beta) = \arg\max_{t \in T} \alpha b_t - \beta c_t,$$
$$= \arg\max_{t \in T} \frac{\alpha}{\beta} b_t - c_t.$$

Cost-sharing can induce the physician to select the optimal treatment:

$$t^{**} = \arg\max_{t \in T} (1+\alpha) b_t - c_t,$$

by setting $\beta = \frac{\alpha}{(1+\alpha)} < 1$. Under full capitation $\beta = 1$, in which case one has suboptimal treatment quality, absent liability since $b_{t(1)} < b^*$. Notice that when α is small, then one needs only a small amount of cost sharing to induce optimal treatment choice. However, given that α is likely to vary greatly from physician to physician it is difficult, if not impossible, for the the MCO to choose an optimal cost sharing parameter.

This implies that when physicians are subject to cost-sharing, negligence may play

25

a dual role, regulating not only expertise but also treatment choice. If β is moderately large, then negligence liability is needed to ensure that informed physicians do not provide suboptimal care in an effort to cut costs. In this case, the negligence standard is more likely to place a binding constraint on the physician's treatment choice.

In conclusion, the use of cost sharing can, in principle, result in the physician making a more efficient decision, but it can do so reliably only if accompanied by negligence liability for negligent treatment outcomes. Indeed, the use of cost sharing further increases the importance of the tort system to ensure quality treatment. For those physicians who place a lower weight α upon good outcomes, the negligence standard in effect defines the standard of care provided by these physicians when they are informed.

5. Optimal Damages with Performance Pay and Entity Liability

The present section expands the previous analysis of individual liability to consider optimal negligence liability where the MCO and physician can contractually agree that, if either is negligent, the negligent provider must make payments to the other. The section compares optimal damage awards under such a regime with optimal damages where such contracts are not possible. We also consider whether welfare would be improved by holding the MCO, rather than the physician, liable for physician negligence.

The existing rule of vicarious liability only holds a Principal liable for its Agent's negligence if the Principal employs the Agent; Principals generally are not liable for torts committed by independent contractors. Most MCO contracts are structured so that affiliated-physicians are independent contractors, and hence MCOs often avoid liability for physician negligence even when physicians are insolvent (and unable to pay optimal damages). We capture the independent-contractor nature of the physician relationship by assuming that the MCO cannot observe the physician's behavior, and thus cannot directly control physician care-taking.

Performance Pay

Consider the situation where the MCO is liable for its own treatment decisions, and either the MCO or the physician may be liable for physician negligence, with the physician liable under "individual" liability and the MCO liable under "entity-level" liability. Assume further that the MCO can sanction the physician whenever she would be found negligent in court, and, in turn, that the physician can sanction the MCO for negligence.[46]

To model this, suppose that the MCO bears a fraction $\gamma \in [0, 1]$ of the physician's liability, with $\gamma = 1$ corresponding to the case of entity-level liability (EL for short)

---

[46]   This does not necessarily entail a court decision. Some cases may decided without the plaintiff ever going to court, either through settlement or arbitration. In that case, the MCO may still penalize the physician should it (or the arbitrator) determine the physician was negligent.

and $\gamma = 0$ corresponding to individual liability (IL for short). In this case the liability of the MCO and the physician are:

$$\mathcal{L}_I = a\pi_I L^I + \gamma (1-a)(1-e)\pi_D L^D, \tag{18}$$
$$\mathcal{L}_D = (1-\gamma)(1-a)(1-e)\pi_D L^D, \tag{19}$$

where L^I are damages awarded for harms caused by MCO negligence and L^D are damages for harm caused by physician negligence. For simplicity, the explicit dependence of L_t on a,e,L^I,L^D, and γ is suppressed.

The sequence of decisions in the relationship is as follows:

1. The patient offers a contract to the MCO that entails a payment P, and damage rules, L^I and L^D. The MCO can either accept the contract or reject it and obtain 0. The MCO reimburses γ of the physician's liability.

2. The MCO then offers a contract to the physician, given the damage rules, the contract terms for a wage, W, and penalties k_I and k_D, depending upon who selected the treatment. The penalties k_I and k_D depend upon the same conditions that trigger a court case, and hence they are substitutes for $\mathcal{L}_I$ and $\mathcal{L}_D$, with the difference that the payments go to the MCO and not to the patient.

3. The MCO and the physician simultaneously set authority and expertise.

4. The patient falls ill and receives treatment from the physician.

As before, it is assumed that the contract terms between the physician and the MCO are agreed upon after the patient has purchased insurance. This is consistent with the observation that the MCO has no obligation to inform the patient of any changes in the terms of the contract with the physician. The purpose of the sanction k_D is to provide the MCO with control over the actions of the physician. The optimal contract is therefore the solution to the following constrained optimization problem:

$$\max_{P,W,L^D,L^I,a,e} I - P + \{B(a,e) + \mathcal{L}_I + \mathcal{L}_D\} \tag{20}$$

subject to the MCO's participation constraint:

$$P - W + a\pi^I k_I - \mathcal{L}_I + (1-a)(1-e)\pi^D k_D - C(a,e) - V_I(a) \geq 0,$$

and the MCO's incentive constraint:

$$W, a, e, k_I, k_D \in \arg\max_{W,a,e,k_I,k_D} \left\{ \begin{array}{c} P - W + a\pi^I k_I - \mathcal{L}_I + (1-a)(1-e)\pi^D k_D \\ -C(a,e) - V_I(a) \end{array} \right\}.$$

The MCO in turn must offer a contract that satisfies the physician's participation con-

straint :

$$\left\{ \begin{array}{c} W + \alpha B\left(a,e\right) - a\pi^I k_I - (1-a)(1-e)\pi^D \\ -k_D - \mathcal{L}_D - V_D\left(e\right)/\theta \end{array} \right\} \geq U_D^0.$$

Finally, the fact that the contract is incomplete implies that the MCO and the physician choose optimal investments *ex post:*

$$a \in \quad \arg \max_{a' \in [0,1]} \left\{ \begin{array}{c} P - W + a'\pi^I k_I - \mathcal{L}_I + (1-a')(1-e)\pi^D k_D \\ -C\left(a',e\right) - V_I\left(a'\right)\} \end{array} \right\},$$

$$e \in \quad \arg \max_{e' \in [0,1]} \left\{ \begin{array}{c} \alpha B\left(a,e'\right) - a\pi^I k_I - (1-a)(1-e')\pi^D k_D \\ -\mathcal{L}_D - V_D\left(e'\right)/\theta \end{array} \right\}.$$

The MCO's ability to employ performance pay fundamentally changes the role of the tort system. In the previous section, where performance pay was not available, tort damages lay where they fell. Thus tort liability has to ensure that, post-contract, the MCO and the physician each have optimal incentives. By contrast, when the MCO employs performance pay it is the MCO that determines post-contractual incentives to invest in expertise and assert authority. Accordingly, the role of the tort system shifts from providing optimal *ex post* incentives to ensuring that, *ex ante*, MCO profits are maximized when expertise and authority are both efficient. These *ex ante* incentives induce the MCO to design performance contracts that in turn induce both the physician and itself to behave efficiently post-contract.

Solvent Physicians

When physicians are solvent with respect to optimal damages, the allocation of liability between the MCO and the physician does not affect either expected liability or MCO incentives. The MCO bears the full cost of its own liability directly, and bears the cost of any liability imposed on the physician through its obligation to pay the physician a sufficiently large wage to ensure that being affiliated with the MCO always is as good as her next best market alternative. Thus, the MCO bears the same expected liability for physician negligence, whether it is imposed directly – in the form of entity-level liability – or indirectly, through the effect of individual liability on physician wages. Thus, we get the following neutrality result:

Proposition 4 *(Neutrality)When performance pay is possible, then the distribution of liability between the physician and the MCO does not affect authority nor expertise, regardless of the damage award.*

This result follows from Proposition 5 below, and so the proof is omitted. It extends the earlier results of the Kornhauser (1982) and Sykes (1984) for a standard Principal-

Agent model to the case of an authority relationship.[47] Performance contracts allow the MCO to internalize all the costs and benefits of medical care to both itself and the physician. Thus, tort liability need only ensure that the MCO internalizes the benefit of treatment to the patient; liability no longer need ensure that each provider bears the cost of her negligence to the other. Once the MCO has incentives to induce e^* and a^*, the MCO will ensure that each provider has the appropriate *ex post* incentives through private sanctions k_I and k_D. The exact form that damages take depends upon the allocation of liability between the MCO and the physician as follows:

Proposition 5 *When $a^* > 0$ then, under either EL or IL, optimal expected damage rules are:*

$$L^{D*} = (b^* - b^e),$$
$$L^{I*} = (b^* - b^e),$$

The optimal contract offered by the MCO entails penalties:

$$\pi^I k_I^* = \alpha \left(b^0 - \left(eb^* + (1-e)\hat{b} \right) \right) + (1 - e^*) \left\{ (b^* - c^*) - \left(\hat{b} - \hat{c} \right) \right\},$$
$$\pi^D k_D^* = \gamma \left(b^* - \hat{b} \right) + \hat{c} - c^*.$$

The proof of this result is in the Appendix.

With performance contracts, optimal liability for physician negligence simply equals the realized harm to the patient, $b^* - b^e$. The MCO employs a private sanction k^D to ensure that the physician also considers the cost of her negligence to the MCO. Thus, the total expected sanction faced by a negligent physician is: $(b^* - c^*) - \left(\hat{b} - \hat{c} \right)$, regardless of whether one has individual or enterprise-level liability.

Similarly, optimal damages for MCO negligence are based only on the patient's expected losses. They no longer must include the cost to the physician of MCO authority because the MCO bears the cost *ex ante* through physician wages, and can provide optimal *ex post* incentives to take physician welfare into account through k_I. Whereas absent incentive contracts, the court would need to adjust liability for MCO negligence by a term to reflect the probability (e) and cost of physician error (see Equation 17), performance contracts simplify the task of computing court- imposed damages. This leaves the task of estimating *expected* physician expertise to the MCO.

Thus, the present analysis suggests that courts and legislatures should not interfere with the MCO and physician performance-pay contracts, providing both MCOs and physicians are subject to optimal negligence liability.[48]

---

[47] Zeiler (2004) shows that neutrality with repect to the liablity regime holds in the context of a game with information disclosure.

[48] Efforts abound to limit the ability to MCOs to obtain indemnification from negligent physicians and

Judgment Proof Physicians

The purpose of this section is to explore the extent to which enterprise liability is optimal in the independent-contractor context when the Agent's assets are not sufficient to cover optimal damages and when the Principal can alter care by regulating the scope of the agency relationship.

For simplicity's sake suppose that the wealth constraint is given by W, the physician's income. This assumption is made for notational simplicity only; the more general case yields the same results. Accordingly, the payoff to the physician is given by:

$$U^\alpha (a, e, \mathcal{C}) = W - a\pi^I k_I - (1 - a)(1 - e)\pi^D \min\left\{k_D + (1 - \gamma) L^D, W\right\} - V_D(e)/\theta,$$

with the sanctions subject to the constraints:

$$k_I \leq W,$$
$$k_D \leq \max\left\{W - (1 - \gamma) L^D\right\}.$$

The wealth constraint implies that it is not possible to fully control the physician via monetary sanctions. This constraint can be relaxed by moving liability from the physician to the MCO, who would then fully internalize the effects of authority and expertise on patient outcomes. When the wealth constraint is binding, this can result in a Pareto improvement.

Proposition 6 *Suppose damages are set equal to harm:*

$$L^{D*} = (b^* - b^e), \tag{21}$$

$$L^{I*} = (b^* - b^e). \tag{22}$$

Then, if the MCO can contract with the physician using the same information as the courts, the rule of entity liability implements the optimal allocation given the physician's wealth constraint, and is strictly preferred to individual liability when this constraint is binding.

Proof. The liability rules (21) and (22) imply that the MCO's liability for negligence is $\mathcal{L}(a, e, L^{I*}, L^{D*}) = B(a, e) + \text{constant}$. Hence, the proof of proposition 5 implies $\Pi^{EL} = (1 + \alpha)B(a, e) - C(a, e) - V_I(a) + /- \text{constant}$. Therefore the MCO has exactly the same objective function as the patient, and hence the MCO chooses a contract that maximizes patient welfare, subject to the physician's IR constraint and the IC constraints for the MCO and the physician, as given by the appropriately modified participation and incentives constraints for the optimization problem 20. Given that

---

to shift liability for MCO negligence. The present analysis demonstrates the value of permitting such contracts.

the set of contracts offered by the MCO include all possible liability regimes that the court can impose as a special case, this implies that the optimal allocation subject to the physician's wealth constraint is implemented.

If the physician's wealth constraint is binding under individual liability, then $W < b^* - \hat{b}$, and one has $\partial \Pi^{IL}/\partial e = \alpha B_e(a,e) + (1-a)W - C_e < \partial \Pi^{EL}/\partial e = (1 + \alpha)B_e(a,e) - C_e$. Since the constraints are continuous and differentiable in expertise, it follows that entity-level liability is superior to individual liability. ∎

This result demonstrates that entity-level liability, when damages are equal to harm, can always ensure the second-best optimum. This does not imply that there does not exist some individual liability rule that implements the second-best, only that such a liability rule would have to be a function of parameters characterizing the MCO-physician relationship, including the level of physician wealth, as we demonstrate in proposition 2. The point is that the MCO affects the quality of medical care via the use of utilization reviews *ex post,* and through the use of cost-sharing contracts with the physician. By moving all liability to the MCO, the courts need only ensure that the MCO internalizes the consequences of its decisions upon patient outcomes, as measured by $b^* - b^e$. The case for MCO liability is further strengthened if we suppose that they have superior information regarding physician performance.

MCO Has Superior Information

Insolvency presents a problem because in serious, permanent injury cases $b^* - b^e$ is very large and is likely to exceed the physician's wealth. Thus the physician's insolvency insulates her from the full cost of her error. The MCO may have the capacity to impose expected sanctions on physicians equal to the full cost of physician error if the MCO can increase the probability that a negligent physician is sanctioned (thereby reducing the magnitude of the optimal sanction). While a court only learns about physician negligence through the victim's lawsuit, the MCO may be able to detect (and sanction) physician negligence even when the patient does not sue, and thereby increase the probability of sanction. For example, in our model physician error yields an expected harm of only $(b^* - \hat{b})$, but optimal damages must be set at the cost of injurious error (at $b^* - b^e$ which is larger than $b^* - \hat{b}$) because courts only sanction physicians if the patient is injured. An MCO able to detect all physician negligence – i.e., all instances when treatment costs equal $\hat{c}$ – could reduce the insolvency problem by imposing a sanction for expected patient losses of $b^* - \hat{b}$ for each instance of negligence, even if the the patient suffered no lasting injury. Through these reduced, but more frequent, sanctions, the MCO could provide the physician with the requisite *ex ante* incentives to take care with a much lower actual damage award, thereby reducing the likelihood that the physician is judgment proof.

Entity-level liability provides the MCO with the optimal incentive to use its information to reduce the problem of physician insolvency because this regime ensures that the MCO maximizes its profits by inducing optimal expertise and authority. Individual liability does not provide the MCO with an incentive to employ this alternative sanc-

tion regime, because reducing physician insolvency only increases its own expected costs: the greater the physician's own expected liability, the greater the MCO's wage obligation to the physician.

To explore this, suppose that the courts impose a penalty only when physician negligence injures the patient – which occurs with probability $(1-a)(1-e)\pi^D$ – but that the MCO can sanction the physician whenever the physician is negligent, which occurs with probability $(1-a)(1-e)$. Hence, if the courts impose a sanction L^D upon the physician, then the MCO can achieve the same deterrence with a sanction of $k_D = \pi^D L^D < L^D$, and hence the MCO faces a less- binding wealth constraint. Therefore, we may conclude:

Proposition 7 *Suppose that the MCO can observe whenever the physician makes an uninformed decision, then if α is sufficiently small, and physician expertise is sufficiently large, entity-level liability is preferred to individual liability. When $\left(b^* - \hat{b}\right)/\pi^D + (\hat{c} - c^*) > W^* > \left(b^* - \hat{b}\right) + (\hat{c} - c^*)$, where W^* is the optimal wage, then entity-level liability results in the first-best, while individual liability is inefficient.*

This result captures the basic reason why vicarious liability is the preferred rule when the Principal has better information than the courts on *ex post* outcomes, and thus can modify dangerous or inefficient behavior on the part of the Agent, even if it does not result in a tort action.[49] This result is consistent with Sykes (1984), who finds that entity-level liability can improve care if damages may be large or small because the entity can impose excess damages in small cases, and thereby increase incentives to take care. Our results extend this insight to the situation where third parties sue whenever they are injured, but where negligence does not always injure the third party.

6. Discussion

Current law treats physicians who provide services for MCOs as independent contractors and hence MCOs are shielded from torts arising from physician negligence. An important feature of MCO contracts is the right to overrule physician decisions and thus to deny the patient coverage for some treatments via a system of utilization review. Dranove and Spier (2003) show that utilization review, especially when the MCO can commit itself to a review strategy, can enhance efficiency through the appropriate screening of cases. Using the model of Aghion and Tirole (1997), we illustrate that utilization review can also lower the incentives for physicians to acquire expertise. This creates a trade-off between the cost reducing effects of an increase in the exercise of authority by the MCO, and the negative consequences for a physician's incentive to

---

[49] The proof is similar to the previous proposition and hence is omitted.

acquire expertise.

In the context of this model, utilization review unambiguously reduces the quality of health care, a result that is consistent with the empirical findings of Duggan (2002) and Aizer, Currie, and Moretti (2004).[50] However, there are situations under which the cost savings can outweigh the impact upon outcomes, and hence it is efficient to have such a system. When the levels of authority and expertise are not contractible and the MCOs are shielded from tort liability, as is currently the law, then the market equilibrium is not efficient. We have provided two scenarios under which tort law can be modified to ensure efficiency.

If the MCO implements only a system of utilization review, and does not otherwise regulate the behavior of physicians, then efficiency can be restored if MCOs are held liable for torts arising from the system of utilization review. The optimal damages in this case differ from the standard damages under a negligence and depend upon the costs of services, as well as the likelihood that a negligent decision is detected.

If the MCO writes a contract with physicians that is conditional upon their performance and the costs that they incur, then holding the MCO liable for all torts and using patient's harm as the basis for the computation of damages results in an efficient allocation. In some cases this allocation is strictly superior to a system with individual-level liability.

Even though our model highlights the potentially adverse consequences of utilization review, it is still the case that the efficient allocation may entail some denial of treatment. In practice, as Aizer, Currie, and Moretti (2004) discuss, the introduction of MCOs may have other beneficial effects. For example, if they are required to cover a patient for a long period of time, then they have an incentive to invest in preventive care to lower future costs. Moreover, they are in a position to collect information regarding the consequences of different treatments from a large population. Potentially, this puts them in a position to make better decisions than individual physicians, whose experience is derived from a much smaller sample of patients. More generally, MCOs are in the unique position of being able to coordinate the many providers that a modern health care system might use to treat a difficult (and expensive) medical case. This in turn suggests that future research should fully explore the possibility of changing current tort law toward allocating responsibility to a single actor, such as the managed care organization.

---

[50] Duggan (2002) also finds that the introduction of an MCO does not appear to lower costs, while Gaynor, Rebitzer, and Taylor (2004) provide evidence that physicians can make decisions that reduce expenditures. These empirical findings are consistent with our model and suggest that the cost saving from less treatment may be outweighed by the fixed costs of setting up the monitoring of treatment decisions.

7. References

AGHION, P., AND J. TIROLE (1997): "Formal and Real Authority in Organizations," *Journal of Political Economy*, 105(1), 1–29.

AIZER, A., J. CURRIE, AND E. MORETTI (2004): "Competition in Imperfect Markets: Does It Help California's Medicaid Mothers?," Discussion Paper w10429, NBER.

ARLEN, J., AND W. B. MACLEOD (2003): "Malpractice Liability for Physicians and Managed Care Organizations," *New York University Law Review*, 78(6), 1929–2006.

BECKER, G. S. (1968): "Crime and Punishment: An Economic Approach," *Journal of Political Economy*, 76(2), 169–217.

CHU, C. Y. C., AND Y. QIAN (1995): "Vicarious Liability under a Negligence Rule," *International Review of Law and Economics*, 15(3), 305–22.

COOTER, R. D. (1982): "Economic-Analysis of Punitive Damages," *Southern California Law Review*, 56(1), 79–101.

COOTER, R. D. (1984): "Prices and Sanctions," *Columbia Law Review*, 84(October), 1523–60.

CRASWELL, R., AND J. E. CALFEE (1986): "Deterrence and Uncertain Legal Standards," *Journal of Law, Economics, and Organization*, 2(2), 279–303.

DRANOVE, D., AND K. E. SPIER (2003): "A Theory of Utilization Review," *Contributions to Economic Analysis and Policy*, 2(1).

DUGGAN, M. G. (2002): "Does Contracting out Increase the Efficiency of Government Programs? EVidence from Medicaid Hmos," Discussion Paper 9091, NBER.

EPSTEIN, R. A., AND A. O. SYKES (2001): "The Assault on Managed Care: Vicarious Liability, Class Actions and the Patient's Bill of Rights," *Journal of Legal Studies*, 30(2), 625–660.

GAWANDE, A. (2002): "The Learning Curve," *The New Yorker*, 77(45), 52–61.

GAYNOR, M., J. B. REBITZER, AND L. J. TAYLOR (2004): "Physician Incentives

in Health Maintenance Organizations," *Jounal of Political Economy*, 112(4), 915–931.

GELIJNS, A. C., J. G. ZIVIN, AND R. R. NELSON (2001): "Uncertainty and Technological Change in Medicine," *Journal of Health Politics Policy and Law*, 26(5), 913–924.

GLIED, S., AND J. G. ZIVIN (2002): "How Do Doctors Behave When Some (but Not All) of Their Patients Are in Managed Care?," *Journal of Health Economics*, 21(2), 337–53.

GRADY, M. F. (1988): "Why Are People Negligent? Technology, Nondurable Precautions, and the Medical Malpractice Explosion," *Northwestern University Law Review*, 82(Winter), 293–334.

GREINER, A., AND E. KNEBEL (eds.) (2003): *Health Professions Education: A Bridge to Quality*, Quality Chasm Series. National Academies Press, Washington, D.C.

HALL, M. A., AND G. F. ANDERSON (1992): "Health Insurers Assessment of Medical Necessity," *University of Pennsylvania Law Review*, 140(5), 1637–1712.

HAVIGHURST, C. C. (2001): "The Backlash against Managed Health Care: Hard Politics Make Bad Policy," *Indiana Law Review*, 34(2), 395–417.

HOLMSTROM, B. (1979): "Moral Hazard and Observability," *Bell Journal of Economics*, 10(1), 74–91.

KESSLER, D., AND M. MCCLELLAN (1996): "Do Doctors Practice Defensive Medicine?," *Quarterly Journal of Economics*, 111(2), 353–90.

KESSLER, D. P., AND M. B. MCCLELLAN (2002): "How Liability Law Affects Medical Productivity," *Journal of Health Economics*, 21(6), 931–956.

KOHN, L. T., J. CORRIGAN, AND M. S. DONALDSON (2000): *To Err Is Human: Building a Safer Health System*. National Academy Press, Washington, D.C.

KORNHAUSER, L. A. (1982): "An Economic Analysis of the Choice between Enterprise and Personal Liability for Accidents," *California Law Review*, 70(December), 1345–92.

KOROBKIN, R. (1999): "The Efficiency of Managed Care "Patient Protection"

Laws: Incomplete Contracts, Bounded Rationality, and Market Failure," *Cornell Law Review*, 85(1), 1–88.

KRIZEK, T. J. (2000): "Surgical Error - Ethical Issues of Adverse Events," *Archives of Surgery*, 135(11), 1359–1366.

MALCOMSON, J. M. (2004): "Health Services Gatekeepers," *Forthcoming Rand Journal of Economics*.

POLINSKY, A. M., AND S. SHAVELL (1998): "Punitive Damages: An Economic Analysis," *Harvard Law Review*, 111(4), 869–962.

SHAVELL, S. (1979): "Risk Sharing and Incentives in the Principal and Agent Relationship," *Bell Journal of Economics*, 10, 55–73.

——— (1980): "Strict Liability Versus Negligence," *Journal of Legal Studies*, 9, 1–25.

——— (1992): "Liability and the Incentive to Obtain Information About Risk," *Journal of Legal Studies*, 21, 259–270.

SIMON, M. J. (1982): "Diagnoses and Medical Malpractice: A Comparison of Negligence and Strict Liability Systems," *Bell Journal of Economics*, 13(1), 170–80.

SULLIVAN, K. (1999): "Managed Care Plan Performance since 1980: Another Look at 2 Literature Reviews," *American journal of public health*, 89(7), 1003.

SYKES, A. O. (1984): "The Economics of Vicarious Liability," *The Yale Law Journal*, 93(June), 1231–80.

ZEILER, K. (2004): "Medical Malpractice and Contract Disclosure: An Equilibrium Model of the Effects of Legal Rules on Behavior in Health Care Markets," Discussion Paper 539224, Georgetown University.

Appendix A

Proof of Proposition 2: Given that the payoff function for each agent is continuous, and quasiconcave in the agent's own strategy, and the strategy space is compact and convex, implies the existence of a Nash equilibrium. The compactness of the constraints ensures the existence of a solution. The fact that $-C_a(e)$, $\alpha B_e(a) > 0$ for $a, e \in [0,1)$, combined with the maintained assumptions on $V_D(a)$ and $V_I(e)$, ensures that both strategies are positive.

Inequality 2 follows from the fact that at the equilibrium one has:

$$
\begin{aligned}
SW_e(a^c, e^c) &= (1+\alpha) B_e(a^c, e^c) - C_e(a^c, e^c) - V_D'(e^c)/\theta^D \\
&= B_e(a^c, e^c) - C_e(a^c, e^c) \\
&= (1 - a^c)\left[\left(b^* - \hat{b}\right) + c^* - \hat{c}\right] > 0.
\end{aligned}
$$

In the case of authority, one has:

$$
\begin{aligned}
SW_a(a^c, e^c) &= (1+\alpha) B_a(a^c, e^c) - C_a(a^c, e^c) - V_I'(a^c) \\
&= (1+\alpha) B_a(a^c, e^c) \\
&= (1+\alpha)[b^0 - \{eb^* + (1-e)\hat{b}\}]
\end{aligned}
$$

from which the final inequality follows.

Proof of Proposition 3: Under the negligence rule, the physician's behavior, given authority $E^N(a)$, solves:

$$
\begin{aligned}
V_D'\left(E^N(a)\right)/\theta &= \alpha B_e(a, e) + (1-a)\pi^D L^D, \\
&= \alpha(1-a)\left(b^* - \hat{b}\right) + (1-a)\pi^D L^D.
\end{aligned}
$$

The first-order condition for optimal expertise for the physician is given by:

$$
(1-a)\left\{(1+\alpha)\left(b^* - \hat{b}\right) + \hat{c} - c^*\right\} = V_D'(e)/\theta
$$

Since V_D is convex, physician expertise is uniquely identified from this expression, and, therefore, at the optimum:

$$
\alpha(1-a)\left(b^* - \hat{b}\right) + (1-a)\pi^D L^D = (1-a)\left\{(1+\alpha)\left(b^* - \hat{b}\right) + \hat{c} - c^*\right\}
$$

and hence:

$$
\pi^D L^D = (b^* - \hat{b}) + (\hat{c} - c^*).
$$

and, therefore, given $\hat{b} = \pi^D b^e + (1 - \pi^D) b^*$:

$$
L^D = (b^* - b^e) + \frac{(\hat{c} - c^*)}{\pi^D}
$$

In the case of the MCO, the first-order conditions the damages must satisfy are given

by:

$$SW_a\left(a,e\right) = -C_a\left(a,e\right) - \pi^I L^I - V_I'\left(a\right),$$
$$\left(1+\alpha\right)B_a\left(a,e\right) - C_a\left(a,e\right) - V_I'\left(a\right) = -C_a\left(a,e\right) - \pi^I L^I - V_I'\left(a\right),$$
$$\left(1+\alpha\right)B_a\left(a,e\right) = -\pi^I L^I.$$

This implies expression 17.

Proof of Proposition 6: Notice that the incentive constraints for the MCO and the physician are concave programs, and hence described by their first-order conditions. Hence, the Lagrangian for the MCO's problem, given the damage rules, is given by:

$$\begin{aligned}
L = {} & P - W - C\left(a,e\right) + a\pi^I k_I + (1-a)(1-e)k_D - \mathcal{L}_I - V_I\left(a\right) \\
& + \lambda \left\{ \begin{array}{c} W + \left\{\alpha B\left(a,e\right) - \mathcal{L}_D - a\pi^I k_I - (1-a)(1-e)\pi^D k_D\right\} \\ -V_D\left(e\right)/\theta - U_D^0 \end{array} \right\} \\
& + \mu_I \left\{\pi^I k_I - (1-e)\pi^D k_D - \partial\mathcal{L}_I/\partial a - C_a\left(a,e\right) - V_I'\left(a\right)\right\} \\
& + \mu_D \left\{\alpha B_e\left(a,e\right) + \partial\mathcal{L}_D/\partial e + (1-a)\pi^D k_D - V_D'\left(e\right)/\theta\right\},
\end{aligned}$$

where λ is the multiplier for the physician's IR constraint, μ_I is the multiplier for the MCO's IC constraint, and μ_D is the corresponding multiplier for the physician's IC constraint.

The first-order condition $\partial L/\partial W = 0$ implies that $\lambda = 1$, hence the Lagrangian is now:

$$\begin{aligned}
L = {} & P + \alpha B\left(a,e\right) - C\left(a,e\right) - \left(\mathcal{L}_D + \mathcal{L}_I\right) - V_I\left(a\right) - V_D\left(e\right)/\theta - U_D^0 \\
& + \mu_I \left\{\pi^I k_I - (1-e)\pi^D k_D - \partial\mathcal{L}_I/\partial a - C_a\left(a,e\right) - V_I'\left(a\right)\right\} \\
& + \mu_D \left\{\alpha B_e\left(a,e\right) + \partial\mathcal{L}_D/\partial e + (1-a)\pi^D k_D - V_D'\left(e\right)/\theta\right\}
\end{aligned}$$

Now notice that $0 = \partial L/\partial k_D = \mu_I(1-a)\pi^D$, and since $a < 1$, then this implies that $\mu_I = 0$. Similarly, $\mu_D = 0$. Therefore, the incentive constraints are not binding. That is, the MCO can select, at no cost, the level of k_I and k_D to control the desired level of authority and expertise *ex ante*. Consequently, the MCO chooses a and e to satisfy:

$$\alpha B_a\left(a,e\right) - C_a\left(a,e\right) - \partial\left(\mathcal{L}_D + \mathcal{L}_I\right)/\partial a - V_I'\left(a\right) = 0, \qquad \text{(A-1)}$$
$$\alpha B_e\left(a,e\right) - C_e\left(a,e\right) - \partial\left(\mathcal{L}_D + \mathcal{L}_I\right)/\partial e - V_D'\left(e\right)/\theta = 0. \qquad \text{(A-2)}$$

Observe that $\mathcal{L}_D + \mathcal{L}_I = a\pi^I L^I + (1-a)\left(1-e\right)\pi^D L^D$, hence demonstrating the first claim, namely that the MCO's decision is independent of γ.

Now the patient's problem is to solve:

$$\max_{P,W,\mathcal{L}_I,\mathcal{L}_D,a,e} I - P + \left\{B\left(a,e\right) + (1-a)(1-e)\pi^D L^D + a\pi^I L^I\right\},$$

subject to the individual rationality and incentive constraints. Substituting in for W and P, using the individual rationality constraints, one arrives at the problem:

$$\max_{P,W,\mathcal{L}_I,\mathcal{L}_D,a,e} I + \left(1+\alpha\right)B\left(a,e\right) - C\left(a,e\right),$$

subject to A-1 and A-2. Since L^D and L^I no longer enter the objective function, then they are selected such that:

$$
\begin{aligned}
-\partial \left(\mathcal{L}_D + \mathcal{L}_I \right) / \partial a &= -\pi^I L^I + (1 - e) \, \pi^D L^D = B_a \left(a^*, e^* \right), \\
-\partial \left(\mathcal{L}_D + \mathcal{L}_I \right) / \partial e &= (1 - a) \, \pi^D L^D = B_e \left(a^*, e^* \right),
\end{aligned}
$$

which implies the optimal-damage rules.

The IC constraints can be used to derive k_I and k_D, which must solve:

$$
\pi^I k_I - (1 - e) \pi^D k_D - \partial \mathcal{L}_I / \partial a - C_a \left(a, e \right) = (1 + \alpha) B_a - C_a
$$

$$
\alpha B_e + (1 - a) \pi^D k_D - \partial \mathcal{L}_D / \partial e = (1 + \alpha) B_e - C_e,
$$

from which one obtains the expressions in the proposition.