HEALTH ECONOMICS
Health Econ. **23**: 397–409 (2014)
Published online 4 April 2013 in Wiley Online Library (wileyonlinelibrary.com). DOI: 10.1002/hec.2920

MULTILATERAL CONTRACTING AND PREVENTION

BRANDON POPE[a,*], ABHIJIT DESHMUKH[b], ANDREW JOHNSON[c] and JAMES ROHACK[d]

[a]*School of Industrial Engineering, Regenstrief Center for Healthcare Engineering, Purdue University, West Lafayette, IN, USA*
[b]*School of Industrial Engineering, Purdue University, West Lafayette, IN, USA*
[c]*Industrial & Systems Engineering, Texas A&M University, College Station, TX, USA*
[d]*Scott & White Health Care, Temple, TX, USA*

ABSTRACT

Incentives created through contracts can be used as a means of decentralized control in healthcare systems to ensure more efficient healthcare. In this paper, we consider an insurer contracting with a consumer and a provider. We focus on the trade-off between *ex ante* moral hazard and insurance, and consider both consumer and provider incentives in the insurer's contracting problem in the presence of unobservable preventive efforts. We study two cases of provider efforts: those that complement consumer efforts and those that substitute for consumer efforts. In the first case, our results show that the provider must have greater incentives when the consumer is healthy to induce effort and that inducing provider effort allows an insurer to offer a more complete insurance contract relative to the bilateral benchmark. In the second case, we state conditions under which these conclusions continue to hold. On the basis of our findings, we discuss the implications and challenges of multilateral contracting in practice. Copyright © 2013 John Wiley & Sons, Ltd.

Received 18 April 2011; Revised 20 December 2012; Accepted 15 February 2013

KEY WORDS: contracting; prevention; moral hazard; healthcare systems

1. INTRODUCTION

Healthcare costs are rising dramatically around the world. Between 2010 and 2020, costs are projected to grow from an average of 9.9% to 14.4% of gross domestic product (GDP) in Organisation for Economic Co-operation and Development (OECD) countries and from 5.4% to 6.2% of GDP in Brazil, Russia, India, and China (PricewaterhouseCoopers, 2010). US healthcare costs have risen at a much higher rate than that of other nations; in 2009, the costs were 17.3% of US GDP (Truffer *et al*., 2010). Although part of the cost increases can be attributed to the development of new medications, availability of advanced diagnostic and surgical procedures, and aging population bases, another major factor is related to the strategic behavior of healthcare system participants based on incentives embedded in healthcare systems (Porter and Teisberg, 2006). For example, the USA ranked ninth out of 11 countries[1] in the percent of primary care doctors who receive financial incentives for managing patients with chronic disease or complex needs, and 10th in enhanced preventive care activities. Incentives can perversely influence strategic behavior of participants in several ways; for example, insurance contracts can reduce incentives for preventive care by removing the burden of risk from individuals. Preventive behaviors have substantial impact on the burden of diseases. For example, the World Health Organization (2009) estimates that 80% of heart disease, stroke, and type 2 diabetes could be prevented by controlling the patient risk factors of diet,

*Correspondence to: School of Industrial Engineering, Regenstrief Center for Healthcare Engineering, Purdue University, West Lafayette, IN, USA. E-mail: popeb@purdue.edu
Supporting information may be found in the online version of this article.

[1]Australia, Canada, France, Germany, Italy, the Netherlands, New Zealand, Norway, Sweden, the UK, and the USA.

physical inactivity, and tobacco use. The estimated cost of heart disease and stroke in the USA alone was $503.2bn for the year 2010 (Lloyd-Jones *et al.*, 2010). However, the absence of incentives to encourage preventive measures means that potential cost savings will remain unrealized.

Incentives are garnering increasing interest as a means of decentralized control for healthcare systems (Valdez *et al.*, 2010; Institute of Medicine, 2010). To create a more efficient healthcare system, the interactions of incentives between consumers, providers, and insurers, and the resulting preventive behaviors must be studied. Thus, we seek to answer the question of how incentives for both consumers and providers can be created via multilateral insurance and remuneration contracts to induce preventive efforts. The interactions of behaviors and incentives for multiple agents can be managed through multilateral contracts, whereas classic bilateral contracts engage only two agents. In particular, the multilateral contracts we study are designed by an insurer to engage both healthcare consumers and providers, whereas classic bilateral contracts engage each of these entities with an insurer independently. Our primary focus is on the trade-off between *ex ante* moral hazard and insurance, considering the hidden preventive efforts of both consumers and providers. This paper first considers provider efforts that complement preventive efforts by consumers. Although it is assumed that providers are ethically motivated to keep consumers healthy, poorly designed incentives can in fact lead to conflict between a provider's ethical and financial incentives. Generally, provider efforts can be categorized into counseling, promoting the consumer's own preventive effort, explaining the benefits and consequences of prevention, prescription refill reminders, and educating the consumer about how to best implement preventive efforts.

We also study the conditions under which our results generalize to substitutive efforts. Our results show that if inducing the provider's effort is optimal, his or her incentives must be greater when the consumer is healthy. That is, the provider must be better off when consumers are healthy rather than ill. We also find that the interaction between provider and consumer incentives creates a distortion in the optimal insurance contract with consumers relative to the bilateral benchmark. Substituting provider effort for consumer risk allows the insurer to save on incentives paid to the consumer by offering a more complete insurance contract. The remainder of this paper is organized as follows. Section 2 reviews the related literature. Section 3 introduces the model, notations, assumptions, and the first-best result. Section 4 analyzes the insurer's multilateral contract problem with hidden efforts, discussing the incentives given to each agent and comparing with the multilateral first-best and bilateral second-best solutions. Section 5 analyzes a variation of the model whereby provider efforts substitute for consumer efforts. Section 6 draws conclusions and discusses implementation challenges and strategies.

2. BACKGROUND AND LITERATURE REVIEW

Mechanism design theory and its two-agent special case, principal–agent theory (Laffont and Martimort, 2002), have produced results in a number of economic application domains, including healthcare (Fuloria and Zenios, 2001; Su and Zenios, 2006). A survey of the healthcare literature shows that the main objective for using principal–agent theory is to control the costs of agency via incentives. In situations of asymmetric information, the agency problems of adverse selection and moral hazard arise because of hidden information and hidden actions. Moral hazard, defined as the changes in an agent's behavior resulting from the changes in the risk the agent faces, can arise before or after the onset of disease in the context of health insurance. *Ex post* moral hazard arises when insurance contracts are written on the basis of expenditures because consumers who face reduced marginal costs of treatment will demand treatment at excessive levels. The seminal work of Zeckhauser (1970) studied the efficient trade-off between *ex post* moral hazard and risk reduction, concluding that the optimal contract should comprise a mix of risk spreading and incentives for efficient expenditures. Moreover, *ex ante* moral hazard occurs when insured consumers can take actions that reduce their risk of adverse events (Ehrlich and Becker, 1972). Without insurance, preventive actions would be taken by the consumer when the expected benefits exceed the costs. If these actions are contractible, the first-best solution of full insurance can be realized. However, if the preventive actions are noncontractible and costly, fully insured consumers will have no incentive to act. Ehrlich and Becker referred to such actions as self-protection and concluded that *ex ante* moral

DOI: 10.1002/hec

hazard will exist if the price of insurance is independent of self-protection. In their paradigm, the second-best solution leaves the consumer with some risk; for example, in 2012, insurance contracts with coinsurance rates of 10%, 20%, or 30% provide incentives for consumers to exert preventive efforts. These *ex ante* moral hazard circumstances are important because consumers have many actions (e.g., diet, smoking, and exercise) that insurers cannot fully observe, yet they affect consumer's risk of health outcomes. Zweifel *et al.* (2009, Chapter 6) provided a recent exposition of both modes of moral hazard. Within *ex ante* models, they focused on binary preventive efforts and conditions for which full insurance or copayments are optimal. Within *ex post* models, they concluded that copayments should be used to control moral hazard, with higher copayments for more price elastic services. Goldman and Philipson (2007) considered optimal insurance with a consumer under *ex post* moral hazard with multiple treatment goods. Their conclusions suggest that *ceteris paribus*, insured goods that are substitutable will have lower copays and those that are complements will have higher copays. Other recent extensions to moral hazard models include that of Ellis and Manning (2007), who derived different optimal coinsurance rates for prevention and treatment goods. Ellis and Manning reported comparative statics on the changes in optimal coinsurance rates for both goods as aspects of consumer preferences and uncompensated costs of treatment, prevention, and illness vary. In contrast, we model noncontractible preventive efforts and focus on the changes in consumer insurance completeness when supply-side (provider) moral hazard can be managed via multilateral contracts.

Although the literature initially focuses on demand-side incentives through the consumer's insurance contract, the strategic behavior of providers driven by supply-side incentives in their remuneration contract can create inefficiencies when providers acting as imperfect agents for their patients make quantity decisions (Ellis and McGuire, 1986) or have the ability to select patients (Ellis, 1998). The consensus of the literature is that when reimbursement is contingent on expenditures, retrospective schemes (such as fee-for-service) will lead to excessive services and upcoding (reclassifying patients into more lucrative diagnoses), and prospective schemes will lead to under-provision of services and avoidance of high-severity patients. Prospective schemes include capitation, whereby the provider is paid a set amount per period for each patient regardless of services delivered. Mixed schemes can be constructed to balance these trade-offs.

The literature also examines incentives that account for the multilateral interactions within a healthcare system. Ellis and McGuire (1990) studied incentives for a provider and consumer bargaining over utilization and concluded that the optimal incentive system gives generous insurance coverage to consumers and incentives to control costs to providers. Ma and McGuire (1997) considered moral hazard when both a provider and an insurer have noncontractible actions in the production of health, and the insurer sees only a report (possibly nontruthful) of treatment. Their results show that providers need cost-sharing incentives when their actions are substitutes and cost-plus incentives when they are complements. Ma and Riordan (2002) studied optimal contracts by using both demand-side and supply-side incentives and investigated the level of utilization incurred relative to the full information benchmark. Their work highlights the need to consider both agents' incentives in order to control inefficiencies in healthcare.

Although most literature has focused on the *ex post* moral hazard problem, we focus on *ex ante* moral hazard. One of the primary arguments for the diminished importance of *ex ante* moral hazard in health insurance has been that nonfinancial costs, for example, pain, discomfort, and suffering, associated with adverse health events are uninsurable; thus, even financially insured consumers will have reason to exert preventive efforts. If this argument is true, then the trend in medical research and technology to minimize or eliminate the nonfinancial burdens of disease could have the unintended consequence of increasing *ex ante* moral hazard. This logic echoes that of Kenkel (2000), who noted that the *ex ante* moral hazard problem becomes larger as prevention and cure become more perfect substitutes. In addition, there appears to be growing empirical evidence for *ex ante* moral hazard in healthcare. Using US data, Stanciole (2008) found evidence of *ex ante* moral hazard in the choice of heavy smoking, lack of exercise, and obesity. In the work most closely related to ours, Dave and Kaestner (2009) found *ex ante* moral hazard regarding physical activity and tobacco consumption and provided evidence that providers influence consumer decisions regarding prevention. de Preux (2011) also found some evidence in the physical activity choices before consumers turn 65 years old and begin receiving

DOI: 10.1002/hec

Medicare. In general, both *ex ante* and *ex post* moral hazards lead to the same conclusion that the consumer should be left with some risk over health outcomes to manage the trade-off between efficiency and insurance. However, the role of the provider's incentive contract in inducing prevention and the effect of multilateral contracting on the incentives and preventive efforts of the consumer are open topics on which we focus in this paper.

3. MODEL

We consider a risk-averse consumer facing an uncertain health state. We assume that the consumer's health state will be either healthy (h) or ill (i), that is, there are no varying degrees of illness or at least there is a single treatment option that restores the consumer to his or her full measure of health and does not vary with the level of illness. Although this may appear to be a strong assumption, we believe that it is approximately valid for acute illness episodes, and we prefer to focus on the *ex ante* dimension of moral hazard. The consumer obtains insurance to alleviate risk stemming from his or her uncertain health state and healthcare expenditures. In the case of illness, the consumer goes to the provider for treatment, which costs d to administer. Both the consumer and the provider choose efforts ($e_C, e_P \in [0,1]$) that are hidden from the insurer but are relevant for determining the level of prevention utilized. We consider the provider's effort as advocating or promoting prevention and the consumer's as taking preventive action. When the provider exerts effort he or she incurs a cost, $c_P(e_p)$, which reflects the time and resources required to exert the effort. We assume that this cost is increasing and convex, with $c_P(0) = 0$, on the basis of an opportunity cost argument that the provider must sacrifice increasingly valuable activities as his or her level of effort increases. The provider's effort serves to lower the disutility experienced by the consumer when the consumer takes preventive action. A consumer exerting effort experiences disutility $\psi(e_C, e_P)$. We assume that this disutility is increasing and convex in his or her own effort and decreasing both absolutely and marginally in the provider's effort $\left(\frac{\partial \psi}{\partial e_P} \le 0, \frac{\partial^2 \psi}{\partial e_C \partial e_P} \le 0\right)$. These reductions in consumer disutility can be thought of as the benefits of inducing the provider's effort from the insurer's perspective; we assume that these benefits are marginally decreasing $\left(\frac{\partial^2 \psi}{\partial e_P^2} \ge 0, \frac{\partial^3 \psi}{\partial e_C \partial e_P^2} \ge 0\right)$. We assume that the provider's and consumer's utilities are separable in income and effort. Let $v(\cdot)$ and $u(\cdot)$ denote the provider and consumer Bernoulli utility functions over wealth, which are strictly increasing, and further assume the consumer is strictly risk averse.

Figure 1 shows the timeline of the contracting problem. In the first stage, the insurer offers a contract to the provider and consumer, which is accepted or rejected. In the second and third stages, the provider and consumer choose their efforts, with the consumer observing the effort level of the provider before exerting his or her own effort. In the fourth stage, nature determines the consumer's health state, and in the fifth stage, the contract is executed by the delivery of care (if necessary) and transfer of payments. We abstract away from the issue of adverse selection, taking the consumer to be representative of the population.

The consumer's effort impacts the probability distribution over his or her health states. Let $\pi(e_C) \in (0,1)$ denote the probability that the consumer is healthy, with $\pi'(\cdot) > 0, \pi''(\cdot) < 0$. Observe that the provider's effort and the consumer's effort are not substitutes in the sense that no amount of effort from the provider can directly impact the probability over health states. This modeling assumption attempts to capture the actions taken by providers to influence consumer prevention. These actions are in contrast to preventive actions providers may take, which directly substitute for the consumer's effort such as vaccinations. We consider the effect of

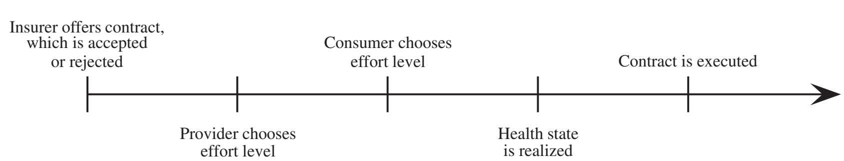

Figure 1. Contract timeline

DOI: 10.1002/hec

substitutive efforts in Section 5. After receiving provider effort e_P, exerting his or her own effort e_C, and facing incomes y_h or y_i when he or she is healthy or ill, respectively, the consumer's expected utility is

$$U(y_h, y_i, e_C, e_P) = \pi(e_C)u(y_h) + [1 - \pi(e_C)]u(y_i) - \psi(e_C, e_P). \quad (1)$$

We assume the consumer is an expected utility maximizer regarding his or her prevention decision. Under this assumption, the consumer will exert effort up to the point where the marginal benefit of prevention is equal to the marginal cost, defined by

$$\pi'(e_C)\Delta u = \frac{\partial \psi}{\partial e_C}(e_C, e_P), \quad (2)$$

where $\Delta u = u(y_h) - u(y_i)$ is the risk the consumer faces, his or her marginal value of staying healthy. Denote the associated optimal level of consumer prevention by $e_C(e_P, \Delta u)$. Clearly, only partially or uninsured consumers will exert positive effort. Intuitively, the model allows the insurer to increase the consumer's effort by inducing more provider effort, which lowers the consumer's marginal disutility of prevention, or increasing the risk the consumer faces, which raises the consumer's marginal benefit of prevention. However, both controls have costs. Exposing the consumer to more risk will limit the transfers the insurer can extract (consumers are willing to pay less for less-complete insurance), whereas inducing the provider's effort requires compensation via increased payments.

3.1. Private insurer's problem

The issue of the relationship between the patient, the provider, and the entity that pools funds to pay for care is similar for the private insurer (for-profit, nonprofit, or self-insured employer) or a public insurer (Medicare, Medicaid, or Federal Employee Health Benefit Program). We will discuss the private sector insurer for simplicity of exposition. Private insurers may have contracting power that can rival a federal program such as Medicare because in many markets, the largest two insurers control the bulk of the market; data from 44 states show the average market share of the top two insurers is 70% (American Medical Association, 2007). Under this assumption, the insurer offers a contract to maximize expected profit (or positive margin in the case of a nonprofit insurer). The contract specifies a set of transfers from the consumer and to the provider contingent upon whether the consumer is healthy or ill, and the effort levels for both agents $\left(\left\{t_h^C, t_i^C, t_h^P, t_i^P, e_C, e_P\right\}\right)$, and provides treatment to the consumer in the case of illness. In traditional indemnity insurance, t_i^P would have to exceed the cost of treatment to ensure provider participation, but increasing the provider's utility when the consumer is healthy through t_h^P allows the insurer to relax this constraint while still maintaining the provider's reservation utility. Although we make the strong assumption that transfers can be made contingent on health states, our assumption is weakened by our minimal health state space ($\{h,i\}$). Under the mild conditions that consumers seek treatment if and only if they are ill, observing expenditures is equivalent to observing the binary health state. We note that our assumption presents little conflict with the geographic variations literature, which casts doubt on the correlation between spending and illness on the basis of data encompassing many individuals, diseases, and other complex factors. Rather, our assumption relates expenditures and health for a single condition and individual. From the insurer's perspective, the contract establishes uncertain payments to be made to the provider and to be received from the consumer, both contingent on the health outcome. Under the terms of the contract, and with initial wealth w, let $u_h = u\left(w^C - t_h^C\right)$, $u_i = u\left(w^C - t_i^C\right)$, $v_h = v\left(w^P + t_h^P\right)$, and $v_i = v\left(w^P + t_i^P - d\right)$ denote the utilities from income of the consumer and provider when the consumer is healthy or ill. Also, let $f(\cdot)$ and $g(\cdot)$ respectively denote the consumer and provider inverse utility functions, for example, $f(u(x)) = x$.

DOI: 10.1002/hec

3.1.1. Complete information. If the efforts of consumers and providers were observable, only the individual rationality (IR) constraints of the agents would be active in the insurer's problem. The IR constraints ensure participation in the contract by offering each agent at least the level of expected utility they could obtain outside the contract. We denote these reservation utilities by V^0 and U^0 for the provider and consumer, respectively. The incentive compatibility (IC) constraints are not active in the complete information case because the insurer can specify effort levels in the contract, observe the efforts exerted, and heavily penalize the consumer if the specified effort levels are not followed. Thus, the insurer's complete information problem is

$$\max_{\{u_h,u_i,v_h,v_i,e_C,e_P\}} \pi(e_C)\left[w^C + w^P - f(u_h) - g(v_h)\right] + (1-\pi(e_C))\left[w^C + w^P - f(u_i) - g(v_i) - d\right]$$

subject to

$$\pi(e_C)v_h + (1-\pi(e_C))v_i - c_P(e_P) \geq V^0 \tag{3}$$

$$\pi(e_C)u_h + (1-\pi(e_C))u_i - \psi(e_C,e_P) \geq U^0. \tag{4}$$

Under our assumptions, the insurer's objective, to maximize his or her expected profit function, is written as the sum of his or her net payments from the consumer and his or her net payments to the provider in each state of consumer health, weighted by the probability of the health states. Letting λ and μ denote the Lagrange multipliers of (3) and (4) and taking derivatives of the Lagrangian with respect to $u_{\mathrm{h}}, u_{\mathrm{i}}, v_{\mathrm{h}}$, and v_{i} yields the following conditions:

$$\lambda = g'(v_i) = g'(v_h) > 0 \quad \Rightarrow \quad \pi(e_C)v_h + (1-\pi(e_C))v_i = V^0 + c_P(e_P), \tag{5}$$

$$\mu = f'(u_i) = f'(u_h) > 0 \quad \Rightarrow u_i = u_h = U^0 + \psi(e_C,e_P). \tag{6}$$

We make two observations: the consumer obtains full insurance because we assume his or her marginal utility from income is identical in all states of health, and the provider's expected payments and costs of treatment just cover his or her reservation utility plus the cost of effort induced. Note that if the provider is risk averse, his or her contract would make him or her equally well-off in each state of nature, similar to the consumer. Such a contract could be accomplished by a zero cost-sharing scheme (fee-for-service), with $t_i^P = t_h^P + d$. Substituting in these expressions into the insurer's objective function, the first-order conditions (which are sufficient when $\frac{\partial^2\psi\partial^2\psi}{\partial e_C^2\partial e_P^2} \geq \left[\frac{\partial^2\psi}{\partial e_C\partial e_P}\right]^2$) prescribe effort levels that equate the marginal benefits and costs,

$$\pi'(e_C)d = \frac{\partial\psi}{\partial e_C}(e_C,e_P)\cdot f'\left(U^0 + \psi(e_C,e_P)\right) \tag{7}$$

$$g'\left(V^0 + c_P(e_P)\right)\cdot c'_P(e_P) = -\frac{\partial\psi}{\partial e_P}(e_C,e_P)\cdot f'\left(U^0 + \psi(e_C,e_P)\right). \tag{8}$$

The first-best efforts equate the marginal savings in treatment payments from increased consumer effort with the increased cost of inducing consumer effort and the increased cost of inducing provider effort with the decreased cost of inducing consumer effort. We now turn to the more interesting and realistic case of incomplete information on behalf of the insurer.

4. INCOMPLETE INFORMATION

When efforts are unobservable, the insurer must write the contract to ensure that the agents exert the intended level of effort. The IC constraints ensure that the agents' efforts are best responses to the incentives created by the contract. The provider's and consumer's IC constraints are

DOI: 10.1002/hec

$$e_P \in \operatorname{arg}max_{\hat{e}_P}\, \pi(e_C(\hat{e}_P;\Delta u))v_h + [1-\pi(e_C(\hat{e}_P;\Delta u))]v_i - c_P(\hat{e}_P)$$
$$e_C \in \operatorname{arg}max_{\hat{e}_C}\, \pi(\hat{e}_C)u_h + [1-\pi(\hat{e}_C)]u_i - \psi(\hat{e}_C,,e_P).$$

Our assumptions ensure that the consumer's IC constraint is concave. A sufficient condition for concavity of the provider's IC constraint is that the provider's effort has decreasing marginal ability to induce consumer effort ($e''_C(\cdot) \leq 0$). The concavity of both IC constraints is considered in detail in the Supporting information online. When these IC constraints are concave programs, they can be replaced by their first order conditions (FOCs),

$$\pi'(e_C(e_P,\Delta u))e'_C(e_P,\Delta u)\Delta v = c'_P(e_P) \tag{9}$$

$$\pi'(e_C)\Delta u = \frac{\partial \psi}{\partial e_C}(e_C,e_P), \tag{10}$$

where $\Delta v = v_h - v_i$ is the provider's marginal value of keeping the consumer healthy. From (9), we observe that the provider must be better off when the consumer is healthy ($\Delta v > 0$) in order to exert any effort. Intuitively, the provider's best response will be to exert zero effort when greater utility is linked to illness. This highlights the need for prospective and mixed incentive schemes, in contrast to cost-plus and fee-for-service schemes when providers have significant influence over consumers' preventive behaviors. Taking derivatives of (9) and (10) shows how an insurer can induce preventive efforts via the risk each agent faces.

Proposition 1:

(i) The consumer's effort increases as the consumer's risk increases, $\frac{\partial e_C}{\partial \Delta u} \geq 0$.
(ii) The consumer's and provider's efforts increase as the provider's risk increases, $\frac{\partial e_C}{\partial \Delta v} \geq 0$ and $\frac{\partial e_P}{\partial \Delta v} \geq 0$.
(iii) The provider's ability to induce prevention decreases according to the consumer's risk, $\frac{\partial^2 e_C}{\partial e_P \partial \Delta u} \leq 0$.
(iv) The provider's effort decreases in the consumer's risk, $\frac{\partial e_P}{\partial \Delta u} \leq 0$.

The proof is found in the Supporting information. The intuition underlying the results of Proposition 1 is discussed in order. (i) The consumer's effort increases as his or her risk increases because his or her financial risk over his or her health outcome is linked directly to his or her marginal benefit of preventive effort. Facing greater financial loss when ill, the consumer will find it worthwhile to increase preventive efforts. (ii) Linking the provider's risk to the efforts of the consumer and provider is the observation that increasing provider effort increases consumer effort. Facing higher income when the consumer is healthy, the provider would like the consumer to exert more preventive effort, which increases the likelihood of obtaining the higher income. Understanding the effect of his or her own effort on the consumer's effort, the provider will increase his or her effort to increase the likelihood of a healthy consumer. This increase in provider effort produces a subsequent increase in the consumer's effort. Results (iii) and (iv) in Proposition 1 are related. The provider's ability to induce prevention from the consumer decreases in the risk the consumer faces because of the expected utility maximization assumption on behalf of the consumer, meaning that consumers facing great risk already have strong intrinsic incentives to exert preventive efforts. Exerting more preventive effort has less impact on the probability of illness and will result in greater disutility because of the diminishing marginal returns and rising marginal disutility of prevention. Similarly, consumers facing little risk, have little intrinsic incentive to exert preventive effort, meaning that the consumer's effort has more impact on the chance of illness and does not cause excessive disutility. Hence, the provider's effort has more impact. In other words, the provider's risk and the consumer's risk are substitutes in the insurer's task of inducing prevention from the consumer. In terms of insurance completeness, result (iii) says that consumers with more incomplete insurance, for example, via higher cost sharing, have stronger intrinsic incentives for prevention and therefore will exert less incremental prevention when encouraged by the provider. Because the provider's ability to induce preventive effort from the consumer is reduced, the provider will exert less effort for a consumer facing high risk.

The insurer's incomplete information problem is

$$\max_{\{u_h,u_i,v_h,v_i,e_C,e_P\}} \pi(e_C)\left[w^C + w^P - f(u_h) - g(v_h)\right] + (1-\pi(e_C))\left[w^C + w^P - f(u_i) - g(v_i) - d\right],$$

subject to (3), (4), (9), and (10). Let λ, μ denote the multipliers of the IR constraints, and let γ, δ denote the multipliers of the IC constraints (9) and (10). Forming the Lagrangian and taking derivatives with respect to u_h and u_i yields

$$\frac{1}{u'\left(w^C - t_h^C\right)} = \mu + \delta\frac{\pi'(\cdot)}{\pi(\cdot)} + \gamma\left(\frac{\Delta v}{\pi(\cdot)}\right)\left[\pi''(\cdot)e'_C(\cdot)\frac{\partial e_C(\cdot)}{\partial \Delta u} + \pi'(\cdot)\frac{\partial e'_C(\cdot)}{\partial \Delta u}\right] \tag{11}$$

$$\frac{1}{u'\left(w^C - t_i^C\right)} = \mu - \delta\frac{\pi'(\cdot)}{1-\pi(\cdot)} - \gamma\left(\frac{\Delta v}{1-\pi(\cdot)}\right)\left[\pi''(\cdot)e'_C(\cdot)\frac{\partial e_C(\cdot)}{\partial \Delta u} + \pi'(\cdot)\frac{\partial e'_C(\cdot)}{\partial \Delta u}\right]. \tag{12}$$

The first two terms on the right sides of (11) and (12) are standard terms found in the second-best bilateral contract between an insurer and consumer in the presence of *ex ante* moral hazard. The final term represents a distortion from the bilateral second-best result due to the interaction between the provider and consumer. The direction of the distortion depends on the provider's risk attitude. The provider's first-order conditions, which determine the incentive spreading needed to induce his or her effort, look very similar to bilateral insurance distortions, moderated by the provider's ability to induce consumer effort, $e'_\mathrm{C}(\cdot)$,

$$\frac{1}{v'\left(w^P + t_h^P\right)} = \lambda + \gamma\frac{\pi'(\cdot)\cdot e'_C(\cdot)}{\pi(\cdot)} \tag{13}$$

$$\frac{1}{v'\left(w^P + t_i^P - d\right)} = \lambda - \gamma\frac{\pi'(\cdot)\cdot e'_C(\cdot)}{1-\pi(\cdot)}. \tag{14}$$

Here, we are interested in investigating the provider's risk attitude because even though a single provider may see thousands of patients, the provider's practice style and ability, as well as the local health conditions from which the provider draws his or her patient base, influence the outcomes of these patients. Therefore, a provider would not benefit from risk spreading in the same fashion as a large insurer would. When the provider is risk neutral, the insurer can induce his or her effort at no informational cost, only increasing his or her expected payments to cover his or her cost of effort, $c_\mathrm{P}(\cdot)$. Using the intuition that the provider's effort and the consumer's risk are substitutes for inducing prevention, when the provider is risk neutral, the insurer will use the maximum amount of provider effort in any multilateral contract. At the other extreme, as the provider becomes infinitely risk averse, no provider effort is induced, and the optimal bilateral contract is implemented. In between these two extremes, we observe three effects in the efforts and incentives of the agents. First, as the provider becomes more risk averse, inducing effort from the consumer becomes more expensive, and thus less prevention is induced. Second, as the provider becomes more risk averse, his or her effort is increasingly costly to induce, and hence the insurer substitutes consumer risk for provider effort in inducing prevention. Third, because the provider responds to both his or her own risk, and the consumer's risk in choosing a level of effort (Proposition 1), as the provider becomes more risk averse, the insurer will use less provider risk to induce his or her effort and compensate by decreasing the consumer's risk. We summarize these results in the following proposition:

Proposition 2:

- (i) The interaction of consumer and provider risk allows a second-order distortion of the consumer's risk back toward full insurance.
- (ii) The level of consumer effort in the optimal multilateral contract is decreasing in the provider's risk aversion.
- (iii) As the provider becomes increasingly risk averse, the insurer substitutes consumer risk for provider effort to induce consumer effort.

DOI: 10.1002/hec

Result (i) stems from the additional terms in (11) and (12) and the previous result from Proposition 1 that the provider decreases his or her effort as the consumer faces increased risk. Results (ii) and (iii) of this proposition are clear from Proposition 1 and the increased cost to induce effort from the provider as he or she becomes more risk averse. By inducing the provider's effort, the insurer needs less risk to induce the consumer's effort and can offer a more complete insurance contract.

Just as distorting incentives away from full insurance is costly to give to a risk-averse agent, distortions back toward full insurance provide a savings. This can be seen in Figure 2, in which u_h, u_i represent a first-best full insurance contract, u_h^*, u_i^* show a bilateral distortion, and u_h^{**}, u_i^{**} show a multilateral distortion for a given effort level. The figure shows that providing the same level of expected utility from income, the distortions in the multilateral contract allow the insurer to save on the incentives created for the consumer. Despite the insurer's ability to offer more complete insurance relative to bilateral contracts, strictly speaking, the consumer will be no better off in the private insurer's multilateral contract, because in either case, the contract provides his or her reservation level of utility. A public or social welfare maximizing insurer could return some of the savings in the multilateral contract to the consumer.

5. SUBSTITUTIVE EFFORTS

Although the types of provider efforts discussed so far complement consumer efforts, other activities on the part of the provider would likely substitute for consumer effort. Examples of these activities include free samples of medicine and elements of treatment not reported. Rather than reducing the consumer's disutility from effort, we model these efforts as directly impacting the consumer's probability of illness ($\pi(e_C,e_P)$). The provider's effort is preventive $\left(\pi'_{e_P}(\cdot) > 0\right)$ but with decreasing effectiveness $\left(\pi''_{e_P^2}(\cdot) < 0\right)$. We model the interaction of these efforts as substitutes by assuming that each agent's ability to prevent illness is decreasing in the other agent's effort $\left(\pi''_{e_C e_P} \leq 0\right)$. We briefly examine the insurer's contracting problem to identify conditions under which our primary conclusions still hold; that is, the provider must be better off when the consumer is healthy to induce effort, and inducing provider effort allows consumer incentives to be shifted back toward the first best of full insurance.

Facing risk Δu and provider effort e_P, the consumer chooses effort e_C, which maximizes $\pi(\hat{e}_C,, e_P)\Delta u - \psi(\hat{e}_C)$. This objective is again concave, and the consumer's IC first-order condition imposes that

$$\pi'_{e_C}(e_C, e_P)\Delta u = \psi'(e_C). \tag{15}$$

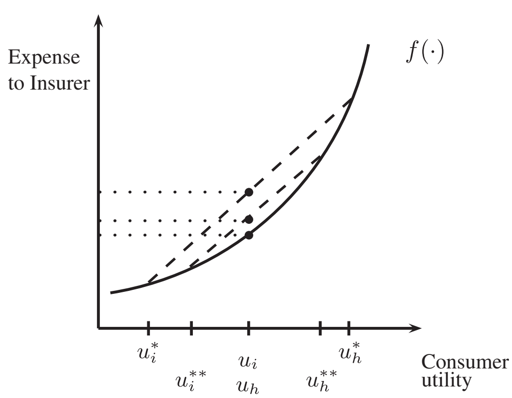

Figure 2. Consumer incentives

On the basis of the aforementioned assumptions, the consumer again increases effort in response to higher risk $\left(\frac{\partial e_C}{\partial \Delta u} \geq 0\right)$ but now decreases effort in response to provider effort $\left(\frac{\partial e_C}{\partial e_P} \leq 0\right)$. The provider chooses effort e_P to maximize $v_i + \Delta v \cdot \pi(e_C(\hat{e}_P; \Delta u), \hat{e}_P) - c_P(\hat{e}_P)$. Again focusing on the case when the provider's IC constraint is concave, the first-order condition gives that

$$\Delta v \left(\frac{\partial \pi(\cdot)}{\partial e_C} \frac{\partial e_C}{\partial e_P} + \frac{\partial \pi(\cdot)}{\partial e_P} \right) = c_P'(e_P). \tag{16}$$

When the total derivative of $\pi(\cdot)$ with respect to e_P is positive, or intuitively, that increasing provider efforts leads to a higher likelihood of a healthy consumer, the provider must be better off when the consumer is healthy to exert effort. A sufficient condition for this to hold is that as the consumer's effort increases, the consumer's ability to prevent illness decreases relative to the provider's ability to prevent illness.[2]

The insurer's incomplete information contracting problem is

$$\max_{\{u_h, u_i, v_h, v_i, e_C, e_P\}} \pi(e_C, e_P)\left[w^C + w^P - f(u_h) - g(v_h)\right] + (1 - \pi(e_C, e_P))\left[w^C + w^P - f(u_i) - g(v_i) - d\right],$$

subject to (15) and (16) and the IR constraints

$$u_i + \pi(e_C, e_P)\Delta u - \psi(e_C) \geq 0 \tag{17}$$

$$v_i + \pi(e_C, e_P)\Delta v - c_P(e_P) \geq 0. \tag{18}$$

In the case of complementary efforts, we found that an insurer could offer the consumer more complete insurance relative to the bilateral contract by inducing the provider's effort. Here, we discuss when these results hold for the case of substitutive efforts. In the optimal multilateral contract, the insurer will substitute provider effort for consumer effort. Similar as before, this substitution is largest when the provider is risk neutral, and there is no added incentive required to induce his or her effort. Intuitively, inducing less effort from the consumer should require less incentives and therefore more complete insurance; however, in general, this may not be the case. Examining the consumer's IC constraint (15), substituting provider effort for consumer effort will have three effects. First, reducing the consumer's effort will reduce the consumer's marginal disutility of effort, the right-hand side of (15). Second, reducing consumer effort will also increase the marginal effect of the consumer's effort on prevention. All else equal, these first two effects would decrease the incentives required (Δu) to induce the lesser amount of consumer effort. The third effect is that increasing the provider's effort will decrease the marginal effect of the consumer's effort on prevention by our assumption of substitutive efforts. This third effect makes the total effect of substituting provider effort for consumer effort on the consumer's ability to prevent ambiguous. We state in the following proposition a condition that ensures the consumer will bear less risk in the optimal multilateral contract when efforts substitute.

Proposition 3:
When $\left|\pi''_{e_C e_P}\right| \leq \pi''_{e_C^2} \frac{\partial e_C}{\partial e_P}$, the insurer's optimal multilateral contract with substitutive efforts will offer the consumer more complete insurance than the bilateral benchmark.

The proof is found in the Supporting information. This condition could actually be weakened, albeit to a less interpretable condition. This result shows that our findings are not unique to a single type of provider effort and interaction between the patient and provider.

For consistency, we have maintained our assumption that the provider is a first mover and considers the effect of his or her effort in inducing effort from the consumer. This assumption also agrees well with the

[2] $\frac{\partial}{\partial e_C}\left(\frac{\pi'_{e_C}(\cdot)}{\pi'_{e_p}(\cdot)}\right) \leq 0 \Rightarrow$ is a sufficient condition, shown in the Supporting information.

DOI: 10.1002/hec

intuition and evidence (Town *et al*., 2005) that consumers look to and respond to provider guidance and direction. However, it is also plausible that the consumer at least partially allocates effort before the provider moves. For example, this scenario could explain providers with strong incentives for patient health allocating more effort to consumers who are observed to be lax in their own self-care. Because it seems unlikely that consumers internalize the effect of their own preventive efforts on their provider's decision, a more complex model of this scenario could look like a blend of a classic bilateral model of consumer effort followed by a version of the multilateral interaction seen here. In this case, it seems likely that the provider will still require incentives tied to patient health in order to exert effort, although the reduced interaction of the patient and provider risk may mitigate an insurer's ability to offer more complete insurance.

6. CONCLUSIONS

This paper considered a healthcare system consisting of an insurer contracting with a consumer and a provider whose efforts interact to stochastically produce the health outcome experienced by the consumer. The trade-off between *ex ante* moral hazard and insurance vis-à-vis the insurer's contracting problem to induce preventive efforts was described in detail. The provider's effort interacting with the consumer's effort in both complementary and substitutive fashions was modeled to capture the various actions performed by providers. The model produced two options for controlling the consumer's preventive efforts: managing the consumer's risk related to the health outcome and inducing the provider's effort.

Our results show that when provider effort is induced in the optimal contract, both agents must be better off when the consumer is healthy. This feature echoes suggestions of Antos *et al*. (2009) as a feature of new payment systems to improve system incentives. The provider's incentives must be designed such that his or her payments ensure his or her reservation level of utility and also make the desired level of effort his or her best response. Achieving the proper balance of expected payments and risk requires a mix of prospective payments, retrospective payments, and bonuses for good health outcomes of consumers. When little provider effort is optimal, little risk is needed, and a cost-sharing hybrid of fee-for-service and capitation may be sufficient, but when greater risk is needed to exert the optimal level of effort, full capitation and bonuses for healthy consumers are useful. Exposing a risk-averse provider to greater risk requires higher expected payments, which can be controlled by increasing the fixed (capitated) component of remuneration. Our multilateral contracting findings are of particular interest for increasingly popular organizational structures such as accountable care organizations and patient-centered medical homes where the association between consumers and providers makes capitated remuneration and health bonuses feasible. On the consumer side, this characteristic is common in policies with copays and coinsurance rates. Comparing with the second-best bilateral benchmark, inducing provider effort allows the optimal multilateral consumer incentives to be shifted back toward the first-best contract of full insurance.

Implementation of multilateral contracts to increase prevention still faces challenges. Our static model speaks to preventive efforts and illness that are enacted and realized within a single contracting period, typically one-year in practice. However, many preventive behaviors and illnesses have multiyear consequences and relationships and would need to be considered in a dynamic model, which we leave for future research. Also, increased specialization in healthcare may eventually split the tasks of treatment and prevention to specialists and primary care physicians. Although such a development would complicate the type of contracts we study in this paper, it would also alleviate some of the provider's moral hazard concern by separating the roles of the prevention and treatment. Alternatives to contracting solutions, including the use of professional ethics or regulation to ensure prevention, face challenges as well. Notably, the resources required for accurate measurement, reporting, and review of data such as the Healthcare Effectiveness Data and Information Set.

We note that two assumptions made to retain relative tractability may limit our results: (1) the lack of selection effects among heterogeneous consumers and (2) the lack of *ex post* moral hazard. Writing contracts with providers with risk over patient outcomes raises the concern that unhealthy patients will be selected against. To

DOI: 10.1002/hec

combat this effect, the incentives in the provider contract, for example, the cost-sharing or health bonus parameters, should be risk adjusted for classes of patients to make all patients financially viable for providers and to induce the proper levels of effort. The contracts we recommend for combating *ex ante* moral hazard have mitigating effects on *ex post* moral hazard. In other words, offering patients more complete insurance tends to exacerbate *ex post* moral hazard, whereas decreasing or eliminating the cost-sharing component of provider treatment tends to alleviate *ex post* moral hazard. When unobservable quality, rather than quantity, of treatment is an *ex post* concern, we suggest that pay-for-performance schemes with incentives for positive treatment outcomes and observable process measures can be used in conjunction with the contracts we recommend.

In conclusion, we believe that optimal alignment of incentives to affect strategic preventive behaviors in healthcare systems is still a relatively unexplored source of potential cost savings. Modeling various system participants and their impacts on a healthcare system, researchers will be able to identify policies by which such incentives can be transmitted.

ACKNOWLEDGMENTS

This research has been funded in part by NSF grant CMMI-0944670. The first author acknowledges the financial support of the U.S. Department of Education through GAANN P200A060241. This paper has greatly benefited from the insights and critique of two anonymous reviewers.

REFERENCES

American Medical Association. 2007. Competition in health insurance: a comprehensive study of U.S. markets. *Technical report*. Available from: http://www.ama-assn.org/ama1/pub/upload/mm/368/compstudy_52006.pdf. (Accessed November 10, 2009)

Antos J, Bertko J, Chernew M, Cutler D, Goldman D, McClellan M, McGlynn E, Pauly M, Schaeffer L, Shortell S. 2009. Bending the curve: effective steps to address long-term health care spending growth. *Technical report*. Available from: http://www.brookings.edu/~/media/Files/rc/reports/2009/0826_btc/0826_btc_fullreport.pdf. (Accessed November 10, 2009)

Dave D, Kaestner R. 2009. Health insurance and ex ante moral hazard: Evidence from Medicare. *International Journal of Health Care Finance and Economics* **9**(4): 367–390.

Ehrlich I, Becker G. 1972. Market insurance, self-insurance, and self-protection. *Journal of Political Economy* **80**(4): 623–648.

Ellis R. 1998. Creaming, skimping and dumping: provider competition on the intensive and extensive margins. *Journal of Health Economics* **17**(5): 537–555.

Ellis R, Manning W. 2007. Optimal health insurance for prevention and treatment. *Journal of Health Economics* **26**(6): 1128–1150.

Ellis R, McGuire T. 1986. Provider behavior under prospective reimbursement - cost-sharing and supply. *Journal of Health Economics* **5**(2): 129–151.

Ellis R, McGuire T. 1990. Optimal payment systems for health-services. *Journal of Health Economics* **9**(4): 375–396.

Fuloria P, Zenios S. 2001. Outcomes-adjusted reimbursement in a health-care delivery system. *Management Science* **47**(6): 735–751.

Goldman D, Philipson T. 2007. Integrated insurance design in the presence of multiple medical technologies. *American Economic Review* **97**(2): 427–432.

Institute of Medicine. 2010. The Healthcare Imperative: Lowering Costs and Improving Outcomes: Workshop Series Summary. The National Academies Press: Washington, DC.

Kenkel D. 2000. Prevention. In Handbook of Health Economics, Culyer AL, Newhouse JP (eds). North Holland: Amsterdam.

Laffont J, Martimort D. 2002. The Theory of Incentives: The Principal-agent Model. Princeton University Press: Princeton, NJ.

Lloyd-Jones D, Adams R, Brown T, Carnethon M, Dai S, De Simone G, Ferguson T, Ford E, Furie K, Gillespie C *et al*. (2010). Heart disease and stroke statistics-2010 update a report from the American Heart Association. *Circulation* **121**(7): E46–E215.

Ma C, McGuire T. 1997. Optimal health insurance and provider payment. *American Economic Review* **87**(4): 685–704.

Ma C, Riordan M. 2002. Health insurance, moral hazard, and managed care. *Journal of Economics and Management Strategy* **11**(1): 81–107.

Porter M, Teisberg E. 2006. Redefining Health Care: Creating Value-based Competition on Results. Harvard Business School Press: Boston.
de Preux LB. 2011. Anticipatory ex ante moral hazard and the effect of Medicare on prevention. *Health Economics* **20**(9, SI): 1056–1072.
PricewaterhouseCoopers. 2010. Build and beyond: the (r)evolution of healthcare PPPs. Available from: http://www.pwc.com/us/en/health-industries/publications/build-and-beyond.jhtml [January 7, 2011].
Stanciole A. 2008. Health insurance and lifestyle choices: identifying ex ante moral hazard in the US market. *Geneva Papers on Risk and Insurance-Issues and Practice* **33**(4): 627–644.
Su X, Zenios S. 2006. Recipient choice can address the efficiency-equity trade-off in kidney transplantation: a mechanism design model. *Management Science* **52**(11): 1647–1660.
Town R, Kane R, Johnson P, Butler M. 2005. Economic incentives and physicians' delivery of preventive care: a systematic review. *American Journal of Preventive Medicine* **28**(2): 234–240.
Truffer C, Keehan S, Smith S, Cylus J, Sisko A, Poisal J, Lizonitz J, Clemens M. 2010. Health spending projections through 2019: the recession's impact continues. *Health Affairs* **29**(3): 522–529.
Valdez R, Ramly E, Brennan P. 2010. Industrial and systems engineering and health care: critical areas of research–final report. *Technical report*. Available from: http://healthit.ahrq.gov/portal/server.pt/gateway/PTARGS_0_3882_948220_0_0_18/IndustrialAndSystemsEngineeringAndHealthCare. (Accessed October 15, 2010)
World Health Organization. 2009. Ten facts about chronic disease. Available from: http://www.who.int/features/factfiles/chp/10_en.html [November 10, 2009].
Zeckhauser R. 1970. Medical insurance: a case study of tradeoff between risk spreading and appropriate incentives. *Journal of Economic Theory* **2**(1): 10–26.
Zweifel P, Breyer F, Kifmann M. 2009. Health Economics. Springer: New York.

DOI: 10.1002/hec