Persistently Inefficient? Organizational Fragmentation, Coordination Failures, and Common Agency Problems in the US Healthcare System

Brigham Frandsen Michael Powell James B. Rebitzer

August 15, 2015

Preliminary and Incomplete

Abstract

It is commonplace to argue that persistent inefficiency in the healthcare system is due to the presence of health insurance. In this paper, we consider another source of inefficiency: the failure of health insurers and other payers to write efficient cost-sharing contracts with providers. The absence of such contracts is widely believed to inhibit investments in integrated care and so to contribute to inefficient organizational fragmentation. Important public policy initiatives, such as Medicare's Accountable Care Organizationse (ACOs), aim to induce efficient contracting between insurers and providers, but the underlying market failure justifying intervention has not been well-articulated or analyzed.

This paper argues that common-agency problems may impede efficient contracting between insurers and providers. We find that common agency can lead to a coordination failure when incentive contracts aim to elicit organizational innovations involving lumpy investments or fixed costs. This coordination failure leads to an inefficient equilibrium in which contracts offer no incentives at all.

Although these results apply to many common-agency problems, they have specific implications for healthcare policy. Interventions such as Medicare's ACOs ameliorate the common-agency market failure in two ways: but subsidizing investments by agents and by jumpstarting more efficient contracting by private payers. Subsidies crowd out private incentives either partially or fully. Jumpstarting can do better, but only if the market is stuck in an inefficient equlibrium and only if the contracts ACOs offer are sufficiently high-powered and aggressive. Weak ACO contracts are likely to have no effect at all.

Keywords: Accountable Care Organizations, Common Agency, Moral Hazard

JEL classifications: D8, I10, I18

1 Introduction

The U.S. healthcare system is famously inefficient, but the causes are poorly understood (Baicker and Chandra, 2011). It is commonplace to argue that persistent inefficiency in the healthcare system is due to the presence of health insurance. In this paper, we consider another source of inefficiency: the failure of health insurers and other payers to write efficient cost-sharing contracts with providers due to a common-agency problem among insurers.

As we discuss below, the absence of such contracts is widely believed to inhibit investments in integrated care and so to contribute to inefficient organizational fragmentation. Important public policy initiatives, such as Medicare's Accountable Care Organizations (ACOs), aim to induce efficient contracting between insurers and providers, but the underlying market failure justifying intervention has not been well-articulated or analyzed. Although our work is motivated by a specific economic setting and a specific set of policy issues, the market failure we analyze is quite general and should apply to many other settings.

Our analysis is motivated by the problem of organizational fragmentation in the U.S. healthcare system. Organizational fragmentation refers to the inefficiencies that result when care delivery is spread out across a large number of poorly coordinated providers.[1] The inefficiencies from poor coordination could, in principle, be ameliorated by efforts to better integrate care. Integration, however, requires providers to make substantial investments in modern digital technologies as well as new managerial practices and new organizational structures (Burns and Pauly, 2002).[2]

The persistence of organizational fragmentation poses an important challenge for economics.[3] If fragmentation is truly associated with meaningful inefficiencies, one would expect

[1] For discussions of organizational fragmentation from an economic perspective see: Cebul, Rebitzer, Taylor, and Votruba (2008). For recent analyses of different aspects of fragmentation in the medical literature see Hussey, Schneider, Rudin, Fox, Lai, and Pollack (2014), Milstein and Gilbertson (2009), and Romano, Segal, and Pollack (2015).

[2] A partial list of these includes investments in health information technology infrastructure and electronic health record systems and clinical decision making support; investments in managerial and financial systems such as payment methods, prospective budgets and resource planning, measures of provider performance, methods of disbursing shared savings to providers and back office assistance; investments in new organizational processes such as care coordination across service settings, providers and specialties; and investments to create new standards of care and protocols that focus more on primary care physicians and non-physician providers as well as patient wellness and prevention . As this list makes clear, some of the investments required to create integrated delivery involve tangible assets, but others require much less-tangible efforts at restructuring relationships between providers.

[3] In 1933 the Committee on the Cost of Medical Care, a blue ribbon commission investigating the cause of the then-high cost of health care, recommended that comprehensive medical services be provided by structured groups of providers, preferably organized around hospitals (Ross, 2002). An analysis of the

the normal process of competition to compel providers to make investments that promote efficient integrated care (Stigler, 1958). Why does this not happen?

One possible answer is that the claimed inefficiencies resulting from fragmentation are overstated. Another possibility, however, is that there are factors inhibiting the operation of normal market forces.[4] We analyze this second possibility via a theoretical model of market failures in contracting between the payer and provider. Payers such as Medicare and commercial insurers have an obvious interest in having providers deliver cost-efficient care, but it is the providers who have the knowledge and ability to direct investments that can achieve that goal. In addition, the investments required for integration are so multi-faceted, complex and hard to observe that it is infeasible for insurers to condition payments directly on these investments. Traditional fee-for-service payment arrangements provide little incentive for providers to make these investments, as all savings accrue to payers. If instead, payers write incentive contracts that allow providers to keep a fraction of the savings they generate from improving the efficiency of their operations, providers have some incentive to make investments. From this perspective, the problem of organizational fragmentation can be analyzed as a principal-agent problem between payers and providers.

The core of our argument—that integration in a fragmented healthcare delivery system requires substantial up-front investments and that these may not be forthcoming when the health insurance system has weak incentives to control costs—will be familiar to many physicians, health services researchers, health economists and other observers of the health care market.[5] Indeed, concern about the role of incentives is the prime motivator for the

modern healthcare system by Baker, Bundorf, and Royalty (2014) points to a continued paucity of integrated care delivery. They estimate that as late as 1998, 29% of physicians worked in solo practices and 55% in practices of 9 or fewer physicians. A great deal of care is still delivered via small practices. According to the 2010 National Ambulatory Care Survey, 31.5% of office visits were to solo practices, and 67.5% were to offices with five or fewer physicians. Only 22.6% of office visits were to multi-specialty groups (Centers for Disease Control, 2010, Table 2). There is some evidence of a sharp increase in the fraction of physicians who are office-based but affiliated with hospitals since the year 2000 (Moses, Matheson, Dorsey, George, Sadoff, and Yoshimura, 2013), but it is not yet clear whether growing hospital affiliations reflect true integration in the delivery of care (Burns and Pauly, 2002).

[4] In earlier eras one could have pointed to legal strictures that sustained fragmentation—especially the corporate practice of medicine laws that made it illegal for physicians to be employed by other organizations, especially hospitals (Robinson, 1999). The legal impact of these laws, however, has nearly disappeared over time—so other explanations are required (Cebul, Rebitzer, Taylor, and Votruba, 2008; Rebitzer and Votruba, 2011).

[5] Crosson (2009), for example, makes a similar argument informally in Health Affairs. Another noteworthy example is Burns and Pauly (2002) who argue that the decline of the ACO-like Integrated Delivery Networks of the 1990s was the due in part to the absence of sufficient up-front investments combined with the fact that the networks entered capitated contracts on a piecemeal basis with only a few payers. Blumenthal (2011)

widespread application of principal-agent models to analyzing health care delivery.

Our analysis differs from conventional principal-agent analysis, however, in one key respect. Providers typically have multiple payers and so their incentives for making investments in integrated care depend on the shared savings contracts they have with *all* their payers. This complication creates a "common agency" problem because multiple principals are simultaneously trying to influence the behavior of a common agent.

Prior analyses of common agency have found that the market failures are more severe than the distortions produced by having a single principal and a single agent. This literature has highlighted the free-riding problem that emerges when multiple principals attempt to elicit costly effort from a common agent. In our setting this free-riding problem emerges if payer A decides to share cost savings with providers to reduce costs, as some of the providers' cost-saving efforts may accrue to other payers. In equilibrium, this free-riding leads to inefficiently weak incentives.[6]

Our research has uncovered an additional distortion that emerges when incentive contracts elicit actions involving "lumpy" investments and fixed costs. These sorts of contracts can produce a previously overlooked market distortion: a coordination failure in which some equilibria are highly inefficient and have essentially no incentives at all. Under these conditions common agency poses an especially problematic barrier to organizational innovation.

To see the issue, consider the problem of a payer who would like a provider to invest in a health IT system to reduce costs. An effective system involves a large fixed-cost investment in hardware and software and then other variable costs resulting from the new processes, routines, and protocols required to effectively operate the system. The payer is not in a good position to understand the right ways to invest or implement such a system, so it instead offers the provider a share of the savings such investments would generate. The provider's incentive to undertake the investment, however, is determined by the collective effect of all the incentive contracts they have with their payers. If each payer believes that it will be the only one to offer a cost-sharing contract with meaningful incentives, then it is not willing to shoulder the entire burden of getting the provider to invest and so it will

argues that the fee-for-service payment system fails to reward improved efficiency that investments in health information technology enable.

[6]The "common agency" models we present have been studied since the mid-1980s and are well suited to analyzing provider-insurer relationships, but have rarely been applied in the health care setting (for an important exception see Glazer and McGuire, 2002). The seminal paper is Bernheim and Whinston (1986b). Much of the subsequent literature on common-agency models focuses on problems of lobbying and influence in political settings. See for example Dixit, Grossman, and Helpman (1997), Besley and Coate (2001); and Kirchsteiger and Prat (2001).

be better off sticking with a traditional fee-for-service contract. In contrast, if each insurer believes that others will offer contracts that reward cost-savings, its incentive to break from traditional fee-for-service contracts increases. The result is a coordination problem in which some equilibria involve significantly less cost-sharing incentives than others. We find that a necessary condition for coordination failures to arise under common agency is that the actions contracts seek to involve "lumpy" investments or fixed costs.

Our analysis of common-agency also has important implications for health care policy, especially for Medicare's Accountable Care Organization (ACO) program. ACOs are a high-profile public policy initiative designed to promote efficient integrated care. ACOs are entities comprised of hospitals and/or other providers that contract with the Center for Medicare Services (CMS) to provide care to a large bloc of Medicare Patients (5,000 or more). Although the details vary and are complex, ACOs that come in under their specified cost benchmarks keep a fraction of the savings conditional on meeting stringent quality standards. The intention of these incentives is twofold. The first is to subsidize new investments in technology and organizational processes that promote more efficient, integrated care. The second effect is to jump-start similar contracts with private insurers—thereby spreading cost-effective, integrated care throughout the healthcare system.[7]

The theory behind ACOs has not been well articulated. Our model enables us to identify conditions under which these subsidies will be effective and jump-starts will occur. Regarding subsidies, we find that ACO contracts do in fact have the effect of subsidizing investments in integrated care, but they crowd out private contracts either partially or fully.

With respect to jump-start effects, we find that these only appear when incentive contracts elicit actions involving "lumpy" investments or fixed costs, as would be true for investments in the sort of technologies and management practices required for moving from fragmented to integrated care. In this case, coordination failures can cause payers to become stuck in a highly inefficient equilibrium with fee-for-service type contracts that offer only limited incentives to control costs. From such a starting point, ACOs can indeed jump start a change in contracting throughout the private sector. But our model finds that this does not have to happen. Jump-start effects require that the contracts written by Medicare ACOs be sufficiently high-powered and aggressive. Weak interventions are likely to have no effect, and there will be no jump starting at all.

[7]Moses, Matheson, Dorsey, George, Sadoff, and Yoshimura (2013) report that there are more than 300 ACOs established in most regions of the United States with 8% of Medicare patients eligible to be served, with a goal to have one-third of the Medicare recipients enrolled by 2018.

The paper proceeds as follows. In Section 2, we present an extended example, which illustrates the central applied lessons of our paper. Section 3 describes our general framework and develops necessary and sufficient conditions for an action choice by an agent to be an equilibrium action. In section 4, we examine properties of the equilibrium action set, and in section 5, we show how public policy interventions affect the equilibrium action set. Section 6 concludes.

Relationship to the Theoretical Literature on Common Agency. Our model of common agency builds upon the seminal work of Bernheim and Whinston (1986a,b), but it differs from the models commonly used in the literature in several ways. Within the common-agency literature, ours is a *complete-information moral-hazard* model with *public contracting* variables, *risk neutrality*, and *limits on transfers*. In particular, we restrict the contracting space to nonnegative, nondecreasing contracts. The agent's outside option yields a payoff of zero, so he would never reject any subset of the contracts he is offered. As a result, our model can be viewed as either one of *delegated* or *intrinsic common agency*.

In the special case of our model with a single principal, the contracting friction is one that arises in limited liability models due to limits on transfers—the resulting trade-off is between incentive provision and rent extraction (Sappington, 1983; Innes, 1990). In the terminology of Martimort and Stole (2012), our contracting space is not *closed under subtraction*,[8] which requires us to depart from the standard analytical tools that rely on a *bilateral efficiency* requirement. Our objective is to characterize the entire set of equilibrium action choices by the agent within a specific class of games, rather than to describe the distributional properties of a subset of equilibria in a general class of games, which is the common approach in the literature.

In contrast to existing common agency models of public contracting, multiplicity of equilibrium actions is not ubiquitous and does not result from parties' flexibility in specifying off-path contractual payments. (Bernheim and Whinston, 1986a; Kirchsteiger and Prat, 2001; Besley and Coate, 2001; Martimort and Stole, 2009) This is because in our setting, all contractible outcomes are generically reached with positive probability on the equilibrium path due to a full-support assumption. Multiplicity cannot be refined away in the standard way by focusing on *truth-telling equilibria* (Bernheim and Whinston, 1986a; Dixit, Grossman, and Helpman, 1997; Martimort and Stole, 2009). Such equilibria generically do not exist in our framework, because ours is a setting in which bilateral efficiency cannot be separated

[8] A contracting space $\mathcal{W}$ is closed under subtraction if given $w, w' \in \mathcal{W}$, $w - w' \in \mathcal{W}$.

from bilateral distribution. In our model, there is always at least one equilibrium action and sometimes more than one. One of our objectives is to establish necessary conditions on the Agent's cost function for there to be more than one.

2 An Illustrative Example

In this section we develop a model of a common-agency market failure, and we use this model to illustrate the effects of policy interventions. To enable a visual representation of the model we present a special case where agent actions have a fixed cost and quadratic variable costs. Section 3 further develops, in a more general setting, the insights we illustrate in this section.

Previous work analyzing common-agency moral hazard problems have typically highlighted a single equilibrium outcome in which incentives are dampened as a result of a free-rider problem among payers. Our approach in this section differs in that we also identify conditions under which common agency leads to multiple equilibria, some of which Pareto dominate others. In this case a coordination failure can emerge in which payers become stuck in an inefficient equilibrium where contracts offer no incentives at all. Not only does this coordination failure result in distinctly worse outcomes than conventional common-agency games, it also has distinct implications for the effects of policy interventions. The possibility of coordination failures, it turns out, depends on a number of factors including: the cost to providers of actions required to implement integrated care and whether these actions involve fixed costs or "lumpy" investments; the efficiency gains from integrated care; and the number of distinct payers interacting with the provider.

2.1 Setup

In our model, an agent's incentives to take an action depend on the contracts they have with all the principals, and the principals set their contracts non-cooperatively and simultaneously. In our case this means that a provider's incentives to invest in efficient integrated care depend on the cumulative effect of all the contracts they have with all their insurers. The actions taken by the agent include investments in infrastructure, measurement and management systems. These are sufficiently complex and hard to observe that contracting directly on them is infeasible. Instead contracts are based upon successful outcomes, corresponding to patients' healthcare costs or quality. We express the cost the provider incurs to achieve a probability of success $a \in A = [0, 1]$ as $c(a) = F + da^2/2$. The term F captures the fixed cost component, while $da^2/2$ term captures increasing marginal costs.

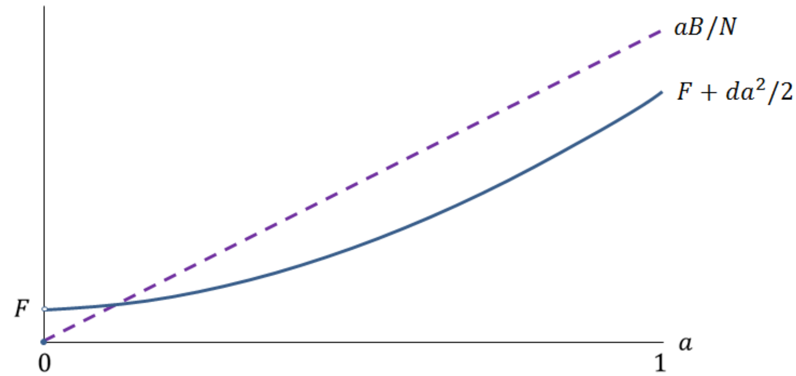

Figure 1: Provider cost function (solid curve) and payer benefit function (dotted line) as a function of the action taken by the provider.

A provider's successful implementation of integrated care improves outcomes for all her patients. Assuming that a provider's patients are equally distributed across N payers (the principals, indexed by i), each payer receives a benefit B/N in the event of a successful outcome and zero otherwise. Figure 1 illustrates the provider's cost function and a payer's benefit function as a function of the action, a, taken by the provider. We assume $B \geq F + d/2$ so that it would be efficient for the provider to choose $a = 1$ and guarantee success. The payers cannot contract directly on the action; instead, each payer i (non-cooperatively and simultaneously) offers a non-negative bonus payment b_i paid to the provider conditional on a success, and no payment otherwise. After receiving the contract offers, the provider takes an action based on the total incentives provided by the aggregate contract, $b = \sum_i b_i$.

The provider responds to the aggregate contract offered by the payers by choosing an action a to solve:

$$a^*(b) = \underset{a}{\operatorname{argmax}}\, ba - c(a)\,.$$

Because of the fixed cost F, the provider chooses $a = 0$ unless the incentives exceed a certain threshold, $\underline{b} = \sqrt{2dF}$, as the following expression for the provider's best response shows:

$$a^*(b) = \begin{cases} 0 & , \quad b < \underline{b} \\ b/d & , \quad b \geq \underline{b} \end{cases}. \tag{1}$$

This expression also implies what the aggregate contract must be to induce a given action:

$$b^*(a) = \begin{cases} \{b : 0 \leq b < \underline{b}\} & , \quad a = 0 \\ da & , \quad a \geq \underline{a} := \sqrt{2F/d} \end{cases}.$$

The expression for $b^*(a)$ shows what the aggregate contract b must be in order to induce the provider to take action a. It implies that when there are fixed costs, only actions $a = 0$ and $a \geq \underline{a}$ are incentive compatible.

To find the set of equilibria in this setting, consider first payer i's choice of contract given b_{-i}, the sum of the contracts offered by all payers besides i. Payer i's choice of contract b_i can be equivalently cast as choosing the action a to induce from the provider. Given b_{-i}, the following expression gives the contract i must offer in order to induce a given action a:

$$b_i^*(a) = \begin{cases} \{b_i : 0 \leq b_i < \underline{b} - b_{-i}\} & , \quad a = 0 \\ da - b_{-i} & , \quad a \geq \underline{a} \end{cases}.$$

This expression implies the cost to payer i of implementing a given b_{-i}:

$$C_i(a) := ab_i^*(a) = \begin{cases} 0 & , \quad a = 0 \\ da^2 - ab_{-i} & , \quad a \geq \underline{a} \end{cases}.$$

Payer i's optimal choice of contract, or, equivalently, choice of action to implement, solves

$$\max_a aB/N - C_i(a).$$

The action payer i chooses to implement depends on how the marginal benefit of the action, B/N, compares to the marginal costs (or savings) from increasing (or decreasing) the action. In the absence of fixed costs, the marginal savings from reducing a is the same as the marginal cost of increasing a. With fixed costs, however, the left and right derivatives of costs are not the same when $a = 0$ and when $a = \underline{a}$. Specifically, payer i chooses $a = 0$ if

$$0 \leq B/N \leq \frac{C_i(\underline{a})}{\underline{a}} = \sqrt{2dF} - b_{-i}. \tag{2}$$

The left-hand side of expression (2) reflects the fact that a cannot fall below zero. The right-hand side reflects the marginal cost of the incentive payment required to move from $a = 0$ to $\underline{a}$ and no higher. Because of the discontinuity resulting from fixed costs, this marginal cost is the same as the average cost of moving from a to $\underline{a}$.

Payer i instead chooses $a = \underline{a}$ if

$$\sqrt{2dF} - b_{-i} \leq B/N < 2d\underline{a} - b_{-i} = 2\sqrt{2dF} - b_{-i}. \tag{3}$$

Finally, payer i chooses $a > \underline{a}$ to satisfy the first order condition

$$C_i'(a) = B/N \tag{4}$$

if

$$B/N > 2\sqrt{2dF} - b_{-i}.$$

Conditions (2), (3), and (4) are, in fact, evaluations of the optimality condition that the marginal benefit to inducing higher action, B/N, be bounded by the left and right marginal costs. Where the cost function $C_i(a)$ is non-differentiable, namely at $a = 0$ and $a = \underline{a}$, the left and right marginal costs differ, as the left and right sides of (2) and (3) show. Where $C_i(a)$ is differentiable ($a > \underline{a}$) the left and right marginal costs coincide and the usual first-order condition (4) determines the optimum.

Equilibrium means all payers choose their b_i to implement the same action $a = \bar{a}$ given all other players' contracts. A consequence of Corollary 2 in the Appendix is that for the purposes of analyzing equilibrium actions $\bar{a}$, we can without loss of generality focus on symmetric equilibria in which all payers offer identical contracts. With this simplification, expression (2) implies that $a = 0$ is an equilibrium action if

$$B/N \leq \sqrt{2dF}. \tag{5}$$

The intuition for this condition is that if all the other payers anticipate $a = 0$, then the marginal cost to a single payer of moving from $a = 0$ to $\underline{a}$ must be born entirely by the single payer. In contrast, the marginal cost when all payers anticipate $a = \underline{a}$ is split equally among the N payers. For this reason, expression (3) implies $a = \underline{a}$ is an equilibrium action if

$$\frac{\sqrt{2dF}}{N} \leq \frac{B}{N} < \frac{\sqrt{2dF}\,(N+1)}{N}. \tag{6}$$

Comparing (5) and (6), it is clear that there exists a value of B/N under which there are two equilibrium actions, the first being much more inefficient than the second because it offers no incentives. This is the essence of the coordination failure produced under common agency. A necessary condition for this coordination failure is that the organizational innovation that the payers wish to elicit from providers involves fixed costs. We develop this point more generally in section 4.

Finally, the first-order condition (4) implies $a = \frac{B}{d(N+1)}$ is an equilibrium action if

$$\frac{B}{N} \geq \frac{\sqrt{2dF}\,(N+1)}{N}. \tag{7}$$

These equilibrium possibilities are illustrated graphically in Figure 2, which plots the marginal-cost correspondence faced by payer i in a symmetric equilibrium at each candidate equilibrium action $\bar{a}$. The sections of the curve composed of vertical segments occur where $C_i(a)$

is nondifferentiable and span the range between the left and right marginal costs. The point at the origin and the thick solid line along the horizontal axis show the set of implementable actions at $a = 0$ and $a \geq \underline{a}$. The dotted horizontal line shows a possible value of B/N corresponding to multiple equilibrium actions at $a = 0$ and $a = \underline{a}$. Equilibrium actions correspond to points where the horizontal line intersects the marginal-cost correspondence. The figure illustrates that a single equilibrium action at $a = 0$ is possible for B/N sufficiently small, or a single equilibrium at $a \geq \underline{a}$ for B/N sufficiently large.

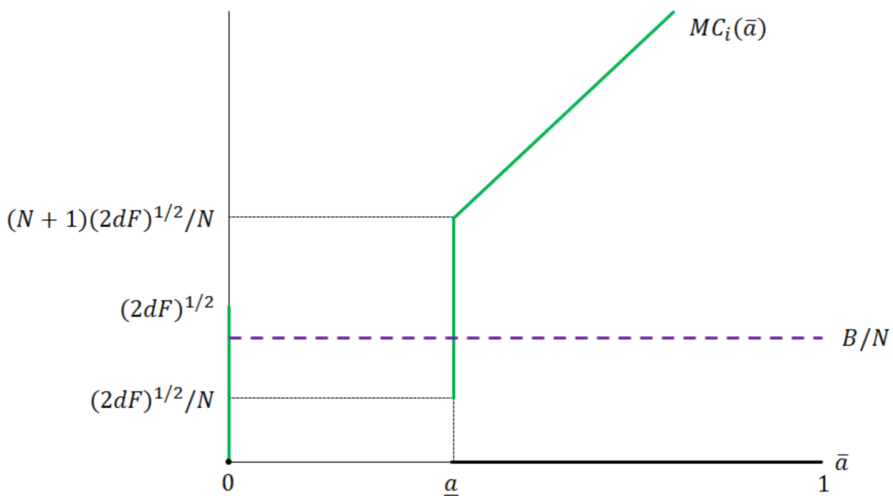

Figure 2: Equilibrium conditions in common agency example. The thick solid line along the horizontal axis shows the support of implementable actions. The thick solid curve labeled $MC_i(\bar{a})$ shows the marginal cost correspondence faced by payer i in a symmetric equilibrium at each candidate equilibrium action level $\bar{a}$. The dotted horizontal line shows a possible value of B/N corresponding to multiple equilibrium actions a $a = 0$ and $a = \underline{a}$.

These conditions illustrate two important consequences of common agency for our setting. First, with several principals, both $a = 0$ and $a = \underline{a}$ are equilibrium actions if

$$\frac{\sqrt{2dF}}{N} \leq \frac{B}{N} \leq \sqrt{2dF},$$

since both (5) and (6) are satisfied in this case. Thus, payers may coordinate on an inefficient equilibrium ($a = 0$) although a Pareto-superior equilibrium exists—that is, coordination failures are possible. Figure 2 depicts an example of possible coordination failure. Second, even when the benefit of the investment is large enough to escape the possibility of coordination failure (that is, $B/N > \sqrt{2dF}$) the equilibrium choice of action is declining in N. This illustrates the free-riding source of inefficiency in common agency settings.

2.2 Policy Interventions

This example can also illustrate how policy interventions such as incentives offered to ACOs influence equilibrium outcomes. Medicare's ACO policy provides incentives to providers to invest in efficient care by sharing savings with the providers when cost targets are met subject to quality thresholds. We capture ACO policy interventions in this example by augmenting the provider's objective function to be $(b+t)\,a - c\,(a)$. The quantity t represents the shared savings distributed to the provider from Medicare if a success occurs, defined in terms of cost and quality outcomes. The ACO policy influences equilibrium outcomes in several important ways. First, a sufficiently aggressive policy can remove the possibility of coordination failure. The equilibrium conditions for $a = 0$ to be an equilibrium action show this:

$$B/N + t \leq \sqrt{2dF}. \tag{8}$$

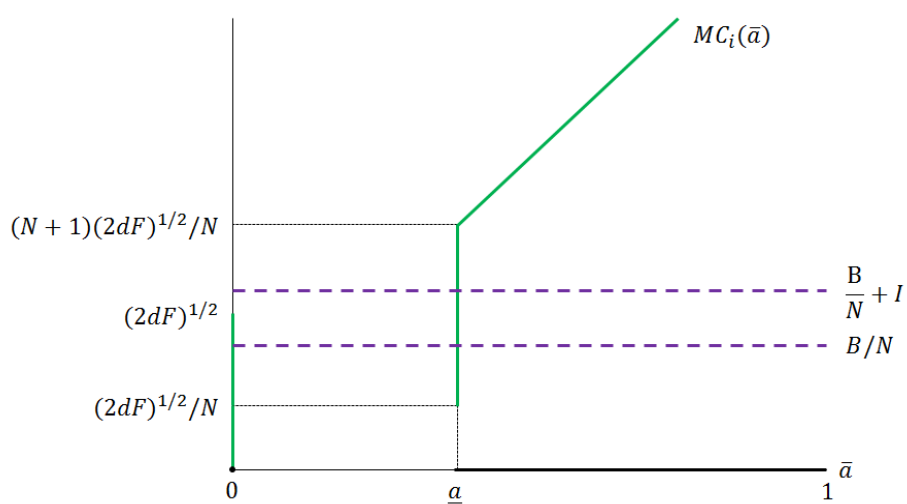

Figure 3: Equilibrium conditions in common agency example with ACO intervention I. The thick solid line along the horizontal axis shows the support of implementable actions. The thick solid curve labeled $MC_i\,(\bar{a})$ shows the marginal cost correspondence faced by payer i in a symmetric equilibrium at each candidate equilibrium action level $\bar{a}$. The lower dotted horizontal line shows a possible value of B/N corresponding to multiple equilibrium actions at $a = 0$ and $a = \underline{a}$. The upper dotted horizontal line shows that with ACO incentive of size I equilibria with $a = 0$ are eliminated.

Thus, when $t > \sqrt{2dF} - B/N$ there cannot be an equilibrium with $a = 0$. Figure 3 illustrates graphically how this condition implies that ACO incentives can eliminate the possibility of coordination failure. The figure shows that an ACO incentive of size t has the same effect as increasing the marginal benefit of the action by t, which can shift the point of intersection with the marginal cost correspondence out of the region where both $a = 0$ and $a = \underline{a}$ are possible equilibrium actions. Thus, an ACO payout sufficiently large ensures that $a = 0$ is not an equilibrium action, and eliminates the possibility of coordination failure. In this way, Medicare's ACO intervention can also spur private payers to offer more powerful incentives for efficient care, jump-starting a widespread move towards investments in integrated care throughout the healthcare delivery system.

Expression (8) also offers a more general framework for thinking about the supply side effects of disparate dimensions of health care policy. One of the major health care initiatives early in the Obama presidency was a large-scale subsidy of investments in health information technology, the Health Information Technology for Economic and Clinical Health Act, which set aside up to $29 billion over ten years to support the adoption and meaningful use of electronic health records (Blumenthal, 2011).

At the time these investments were justified by the observation that the state of health care information technology badly lagged the rest of the economy and that small physician offices could not, on their own, afford the substantial fixed costs such investments entail. Public subsidies of health IT were also seen as part of the larger fiscal stimulus required to combat the very early stages of the severe 2008 recession (Blumenthal, 2011).

Whatever the economic merits of these explanations, expression (8) highlights another economic aspect of these subsidies. Subsidies have the effect of reducing the fixed or marginal cost of providers' investments in integration, (captured by the parameters F and d), which causes the right hand side of this expression to fall. To the extent that information technology expenditures are an important component of F and d, subsidizing provider expenditures on health IT will make it easier to transform the "least" equilibrium into one in which providers write contracts that promote cost-efficient care. The costs captured by F and d ought also to include the costs of negotiating the rules, procedures and quality metrics that providers and payers must agree upon to implement cost-sharing incentives. To the extent that Medicare's leadership and legitimacy in formulating rules and regulations makes these private negotiations easier, they also reduce F or d and also make it easier to encourage efficient incentives in the "least" equilibrium

Our analysis also highlights a second implication for policy: Medicare's ACO intervention can crowd out private incentives for efficient care. This crowd out is most stark when the initial (prior to policy intervention) equilibrium action is $a = \underline{a}$. A condition for an equilibrium at $a = \underline{a}$ is that the average payer benefit, B/N, be less than $\frac{(N+1)\sqrt{2dF}-t}{N}$.

Thus implementing policies with $t < (N+1)\sqrt{2dF} - B$ may not move the equilibrium action taken by the provider at all, completely crowding out private incentives. Note that this condition for complete crowd out is more likely to hold as the payer side of the game becomes more fragmented (i.e. as N increases). Even when the equilibrium action is beyond the sticking point at $a = \underline{a}$ policies still have a crowding out effect. In this regime the provider's and payers' first-order conditions imply that the equilibrium action does respond to the policy, since the equilibrium action taken by the provider is increasing in t, as the following expression shows:

$$a(t) = \frac{B+t}{d(N+1)}.$$

Total private incentives, however, are partially crowded out, since the equilibrium incentive contract offered by payers is decreasing in t, as the following expression shows:

$$b(t) = \frac{B - Nt}{N+1}.$$

The illustrative model we have developed in this section is limited in two important ways. First, it restricts attention to a particular functional form of fixed and marginal costs. This restriction greatly simplifies the exposition of coordination failures and free-riding because it concentrates the non-differentiable "sticking points" at $a = 0$ and $a = \underline{a}$, allowing a stark distinction between regimes where coordination failures and free riding drive the inefficiencies. We show in the next section, however, that the main insights conveyed by this example hold in a more general setting.

In the following sections, we present a generalized version of the common-agency model we considered in this section. Our analysis differs from prior analysis in that we identify a set of nontrivial necessary conditions for coordination failures to emerge in a common-agency moral hazard setting. As we detail below, these conditions require that the actions agents take have a discrete component such as a fixed cost. In the healthcare context, organizational innovations involving new management structures, care processes, or information technology systems likely meet these criteria. These criteria are not satisfied in most prior common-agency moral hazard models—when the agent's cost function is well-behaved. Thus, coordination failures should not arise when principals are trying to elicit greater effort and attention from a common agent.

3 The Model

There are N risk-neutral Principals (denoted $P_1, \ldots, P_N$) and a single risk-neutral Agent (denoted A). Principal i receives benefit B_i if a binary outcome $y = 1$ and zero of $y = 0$. The probability of a successful outcome is determined by A's action choice, a: $\Pr[y = 1|a] = a \in \mathcal{A} \subset [0, 1]$, where $\mathcal{A}$ is compact. The action is costly to the Agent: choosing a costs $c(a)$, where c is lower semicontinuous and nonnegative. Principals simultaneously and noncooperatively offer contracts $w_i \in \mathcal{W}$, where $\mathcal{W} = \{(w(0), w(1)) : w(1) \geq w(0) \geq 0\}$ is the set of nonnegative, nondecreasing contracts. We will often represent feasible contracts equivalently as a nonnegative salary component $s_i \geq 0$ and a nonnegative bonus component $b_i \geq 0$, so that $s_i = w_i(0)$ and $s_i + b_i = w_i(1)$.

The Agent can decide whether to accept a subset of the contracts, and if he accepts no contracts, he receives 0. Since contracts must pay a nonnegative amount to the Agent, we can without loss of generality assume he accepts all contracts. As a result, the Agent cares about the **aggregate contract** $w = w_1 + \cdots + w_N$. The game is therefore a delegated common-agency game with a public contracting variable and risk-neutral participants. We assume throughout that if the Agent is indifferent among several action choices, he chooses the highest action he is indifferent among.

The timing of the game is as follows:

1. $P_1, \ldots, P_N$ simultaneously offer $w_i \in \mathcal{W}$;
2. A chooses $a \in \mathcal{A}$ at cost $c(a)$;
3. $y \in \{0, 1\}$ is realized, and $w_i(y)$ is paid from P_i to A.

A **subgame-perfect equilibrium** of this game is a vector of contracts $(w_1^*, \ldots, w_N^*)$, where $w_i^* \in \mathcal{W}$ and an action-choice function $a^* : \mathcal{W} \to \mathcal{A}$ such that: (1) given w_{-i}^* and a^*, P_i optimally offers w_i^*, and (2) given $(w_1^*, \ldots, w_N^*)$, A optimally chooses a^* (abusing notation slightly). We will say that $w^* = w_1^* + \cdots + w_N^*$ is an **equilibrium aggregate contract**, $b^* = b_1^* + \cdots + b_N^*$ is an **equilibrium aggregate bonus**, and a^* is an **equilibrium action** if they are part of an equilibrium. Denote $\mathcal{A}^* \subset \mathcal{A}$ to be the **set of equilibrium actions**. Our objective is to characterize this set and to describe how it depends on properties of the function $c(\cdot)$.

3.1 Preliminaries

In this section, we will establish an equivalence between the set of equilibrium actions and the solutions to a self-generating maximization program. To do so, we will introduce a few pieces of terminology and notation. For a function $f : [0,1] \to \mathbb{R}$, the **left- and right-derivatives** of f at x are denoted by $\partial^- f(x)$ and $\partial^+ f(x)$. The **convexification** of a function $f : \mathcal{A} \to \mathbb{R}_+$ is the largest convex function $\tilde{f} : [0,1] \to \mathbb{R}_+$ such that $\tilde{f}(x) \leq f(x)$ for all $x \in [0,1]$. The set of **incentive-feasible** actions is the set of actions for which the agent's cost function coincides with its convexification $\mathcal{A}^{feas} = \{a \in \mathcal{A} : \tilde{c}(a) = c(a)\}$. The set $\mathcal{A}^{feas} \subset \mathcal{A}$ is compact, because c is lower semicontinuous and $\mathcal{A}$ is compact. We describe below the sense in which $\mathcal{A}^{feas}$ corresponds to the set of implementable actions.

The **single-principal cost function** $C : \mathcal{A} \to \mathbb{R}_+$ is the solution to the cost-minimization problem

$$C(a) = \min_{w \in \mathcal{W}} s + ba$$

subject to the agent's incentive-compatibility constraint

$$a \in \operatorname*{argmax}_{a'} s + ba' - c(a').$$

Because the Agent's preferences are additively separable in money and costs, the solution to this problem can always be written as $C(a) = c(a) + R(a)$ for all $a \in \mathcal{A}^{feas}$, where $R(a)$ are the **incentive rents** required to get the agent to choose action a. Note that for all $a \notin \mathcal{A}^{feas}$, $C(a) = +\infty$, since if $\tilde{c}(a) \neq c(a)$, then there does not exist a contract under which a is incentive-compatible. We will say that a solution to this problem is a **cost-minimizing contract implementing action** a and denote the resulting contract by w_a^*, the resulting salary component by s_a^*, and the resulting bonus component by b_a^*. We show in Lemma 5 in the Appendix that cost-minimizing contracts satisfy $s_a^* = 0$ and $b_a^* = \partial^- \tilde{c}(a)$. We will say that P_i **supports** action a if she offers the Agent the contract $w_i = w_a^*/N$.

We introduce the following conditions that will be used in some of the results. We indicate in the statement of each result where each condition is imposed. Our main results require both conditions to hold, but several of our intermediate results apply more generally.

CONDITION S. *Principals are symmetric, so that* $B_1 = \cdots = B_N = B/N$.

Because we are focusing on a restricted contracting space, we are only able to develop necessary and sufficient conditions for $a \in \mathcal{A}^*$ when Condition S holds. When Condition S does not hold, the conditions we identify are necessary but not sufficient.

To state the next condition, define the quantity

$$Z(a,a') = \frac{\partial^- \tilde{c}(a) - \partial^- \tilde{c}(a')}{a-a'}.$$

CONDITION Z. *For each $a, a' \in \mathcal{A}^{feas}$ with $a \geq a'$, $Z(a,a')$ is increasing in a and a'.*

This condition implies that the rents that must be provided to the agent to choose action a are convex-extensible on $[0,1]$,[9] which allows us to characterize equilibrium outcomes using marginal conditions. When c is everywhere thrice differentiable, Condition Z is implied by $c'''(a) \geq 0$. If c is strictly convex on $\mathcal{A} = [0,1]$, then Condition Z implies c is differentiable.

Finally, we will say that a cost function is **well-behaved** if it satisfies the following condition.

CONDITION W. $\mathcal{A} = [0,1]$, c is twice-differentiable with $c', c'' > 0$, and $c''' \geq 0$.

Before stating the main results, it will be useful to establish benchmarks to which to compare equilibrium actions. A **first-best action** a social planner would choose is any action satisfying

$$a^{FB} \in \underset{a \in \mathcal{A}^{feas}}{\operatorname{argmax}} \, Ba - c(a).$$

It will always be the case that $a^{FB} \in \mathcal{A}^{feas}$, since it is an action A would be willing to choose if the aggregate bonus were $b = B$. Because actions are not directly contractible, the first-best action will not typically be an equilibrium action.

A **second-best action** is any action a single principal would implement, or

$$a^{SB} \in \underset{a \in \mathcal{A}^{feas}, w \in \mathcal{W}}{\operatorname{argmax}} \quad Ba - s - ba$$

subject to the agent's incentive-compatibility constraint

$$a \in \underset{a' \in \mathcal{A}}{\operatorname{argmax}} \, s + ba' - c(a').$$

We can equivalently define the second-best action using marginal conditions. Define $MC^-(a) \equiv \partial^- \tilde{C}(a)$ and $MC^+(a) \equiv \partial^+ \tilde{C}(a)$, and define the single-principal marginal-cost correspondence

$$MC(a) = \left\{ x \middle| \, MC^-(a) \leq x \leq MC^+(a) \right\}.$$

[9] $R : \mathcal{A}^{feas} \to \mathbb{R}$ is **convex-extensible on** $[0,1]$ if R is the restriction of a convex function $\tilde{R} : [0,1] \to \mathbb{R}$ to the domain $\mathcal{A}^{feas}$. (See Kiselman and Samieinia (Forthcoming))

The second-best action a^{SB} satisfies $B \in MC\left(a^{SB}\right)$. In general, we will have that $a^{SB} \leq A^{FB}$, because $R(a)$ is increasing. As we establish below, a^{SB} in general represents an upper bound on equilibrium actions in the common-agency game.

3.2 Necessary and Sufficient Conditions for Equilibrium Actions

In this subsection, we will show that an action a^* is an equilibrium action when there are $N \geq 2$ Principals if and only if

$$a^* \in \underset{a \in \mathcal{A}^{feas}}{\operatorname{argmax}} \frac{1}{N} Ba - C_i\left(a, a^*\right),$$

for some function $C_i\left(a, a^*\right)$, which is described in Theorem 1 below, and which we will now derive. Our analysis parallels Bernheim and Whinston's (1986b), but it differs in two respects, because all contractual payments are required to be nonnegative. First, since the Agent's outside option yields a payoff of zero to the Agent, we can without loss of generality assume that the Agent will accept all contracts he is offered. In our setting, therefore, there are no equilibria "without participation" as there are in Bernheim and Whinston's setting. Second, we cannot make use of Bernheim and Whinston's two-step argument that each Principal "undoes" all other Principals' contracts and puts in place her own contract with the Agent, because the resulting contracts may be infeasible. It turns out that, nevertheless, all equilibrium aggregate contracts are cost-minimizing contracts for some action. This result, however, does not immediately imply that we can restrict attention to cost-minimizing contracts when solving each Principal's problem.

We are ultimately interested in describing when a particular action a^* is an equilibrium action, but a^* is an equilibrium action if and only if there is an equilibrium aggregate contract w^* under which the Agent finds it optimal to choose a^*. Since the Principals' choices pertain to contracts rather than directly to the Agent's action choice, we have to first describe the conditions under which an aggregate contract w^* is an equilibrium aggregate contract. An aggregate contract $w^* \in \mathcal{W}$ is an equilibrium aggregate contract if and only if there exists $w_1^*, \ldots, w_N^*$, with $w_i^* \in \mathcal{W}$ and $\sum_i w_i^* = w^*$ such that for each i, Principal P_i prefers to offer contract w_i^* if she believes the other principals will offer contracts $\left\{w_j^*\right\}_{j \neq i}$. Denote the sum of the other principals' contracts by $w_{-i}^* = \sum_{j \neq i} w_j^*$. Define b_{-i}^* similarly. Let $a(b)$ denote the Agent's choice of action when he faces an aggregate contract with aggregate bonus b. We therefore have that w^* is an equilibrium aggregate contract if and only if there exists feasible contracts $w_1^*, \ldots, w_N^*$ such that $\sum_i w_i^* = w^*$ and for each i,

$$w_i^* \in \underset{w_i \in \mathcal{W}}{\operatorname{argmax}} \frac{1}{N} Ba\left(b_i + b_{-i}^*\right) - b_i a\left(b_i + b_{-i}^*\right) - s_i.$$

Each Principal takes as given the contracts other Principals offer and chooses a contract that maximizes her expected profits. Offering a contract with a higher bonus increases the amount that P_i will pay the Agent when success occurs, but it also may increase the action the Agent will choose.

As is standard in the common-agency literature, we can perform a change of variables and think of P_i as directly choosing the aggregate contract that the Agent will face: that is, we can change P_i's choice variable from w_i to $w = w_{-i}^* + w_i$. In order for w to be feasible for P_i given w_{-i}^*, it has to be the case that $w \in \mathcal{W} + w_{-i}^* = \left\{\hat{w} + w_{-i}^*, \hat{w} \in \mathcal{W}\right\}$. Under this change of variables, P_i's optimal choice solves

$$w^* \in \underset{w \in \mathcal{W} + w_{-i}^*}{\operatorname{argmax}} \frac{1}{N} Ba(b) - \left(b - b_{-i}^*\right) a(b) - \left(s - s_{-i}^*\right).$$

Equilibrium aggregate contracts have to satisfy what Martimort and Stole (2012) refer to as the "aggregate concurrence principle": in equilibrium, all Principals must agree on the aggregate contract the Agent will face. We show in Corollary 2 in the Appendix that w^* is an equilibrium aggregate contract if and only if there is a symmetric equilibrium in which w^* is the resulting aggregate contract, so in characterizing the set of equilibrium aggregate contracts, we can without loss of generality assume that $w_{-i}^* = (1 - 1/N)\, w^*$.

Instead of fully characterizing the set of equilibrium aggregate contracts, we are able to focus on the simpler problem of characterizing the set of equilibrium actions. To do so, we first recognize (see Proposition 2A in the appendix) that all equilibrium aggregate contracts are cost-minimizing contracts for some action a. This result parallels Bernheim and Whinston's (1986b) Theorem 1, but their results do not imply ours. In particular, their result requires that each Principal is able to undo all other Principals' contracts and put in place her own—and if she is going to do so, she will clearly put in place a cost-minimizing contract for whichever effort level she wants the Agent to choose. Such contracts are not in general feasible in our setting.

The fact that equilibrium aggregate contracts are cost-minimizing contracts in our setting follows from risk-neutrality. Since the Agent is risk-neutral, the set of contracts that get the agent to choose a particular action a is a convex set. Suppose the equilibrium aggregate contract is not cost-minimizing. Then the convex hull of the equilibrium aggregate contract and the cost-minimizing contract implementing that action must intersect the set of feasible aggregate contracts. All aggregate contracts in the intersection are feasible and those in the interior of this intersection necessarily must involve lower expected costs to P_i, since she is also risk-neutral. Combined, these results show that if other Principals support an equilibrium action a^*, then it is a best response for P_i to support a^*.

More generally, we show in Lemma 4 in the Appendix that, given other Principals put in place a set of contracts that add up to w_{-i}, it is a best response for P_i to put in place an aggregate contract that is **cost-minimizing relative to** w_{-i}. That is, we can think of P_i as choosing an action a and the minimum-cost feasible aggregate contract, given that others' offers add up to w_{-i}, under which the Agent chooses action a. It is therefore without loss of generality to assume that each Principal chooses an action at minimum cost to herself, even though the resulting contract need not be a cost-minimizing contract for that action. In particular, if all other Principals support an action a, there is a minimum feasible action, which we will denote $a_{\min}(a)$ that P_i can implement by choosing $w_i = 0$. The resulting aggregate contract will involve aggregate bonus $b = (1 - 1/N)\, b_a^*$, which may not be part of a cost-minimizing contract for action $a_{\min}(a)$. Denote the cost-minimizing contract for action a relative to $(1 - 1/N)\, w_{a^*}^*$ by $\hat{b}_{a,a^*}$.

Combining these results, an action a^* is an equilibrium action if and only if

$$a^* \in \underset{a \geq a_{\min}(a^*)}{\operatorname{argmax}} \frac{1}{N} Ba - \left(\hat{b}_{a,a^*} - \left(1 - \frac{1}{N}\right) b_{a^*}^*\right) a.$$

We can define Principal i's **effective cost function given** $\bar{a}$ as

$$C_i(a, \bar{a}) = \max\left\{C(a), (1 - 1/N)\, b_{\bar{a}}^* a_{\min}(\bar{a})\right\} - (1 - 1/N)\, b_{\bar{a}}^* a.$$

Finally, in order to state our first theorem, define the set of optimal action choices for P_i given others support action $\bar{a}$

$$\hat{a}(\bar{a}) = \underset{a \in \mathcal{A}^{feas}}{\operatorname{argmax}} \frac{1}{N} Ba - C_i(a, \bar{a}).$$

The following theorem provides necessary and sufficient conditions for an action a^* to be an equilibrium action.

THEOREM 1. Suppose Condition S holds. Then $a^* \in \mathcal{A}^*$ if and only if $a^* \in \hat{a}(a^*)$.

PROOF OF THEOREM 1. See Appendix.

Theorem 1 provides a complete characterization of the set of equilibrium actions. In particular, it shows that we can characterize the set of equilibrium actions by looking for the solutions to a problem of a single Principal choosing an action given a modified cost function, which in turn takes as a parameter a "proposed" equilibrium action. If the solution to the problem coincides with the proposed equilibrium action a^*, then a^* is indeed an equilibrium action. In other words, the set of equilibrium actions are the set of fixed points of the operator defined by $\hat{a}(\cdot)$. This operator indeed has a fixed point, so there is at least one equilibrium action.

THEOREM 2. Suppose Condition S holds. The set of equilibrium actions $\mathcal{A}^*$ is nonempty.

PROOF OF THEOREM 2. See Appendix.

Theorem 2 follows by an application of Tarski's fixed-point theorem given the observation that $C_i(a, \bar{a})$ satisfies decreasing differences in a and $\bar{a}$, which we show in Lemma 7 in the Appendix. When it comes to characterizing the set of equilibrium contracts, there is a sense in which there are strategic substitutes and a sense in which there are strategic complements across Principals. Given an action, there are strategic substitutes in contract offerings—if others offer contracts that provide a higher bonus payment to the Agent, then Principal i can implement the given action by offering a lower bonus payment. However, when it comes to characterizing the set of equilibrium actions, the formulation we derive in Theorem 1 suggests that there are, in some sense, strategic complements across actions: if other principals support a higher action, then the marginal cost to Principal i of supporting a higher action fall.

In the next section, we explore the conditions on the Agent's cost function under which complementarities across Principals is sufficiently strong to guarantee there will be multiple equilibrium actions. In order to do so, we will now describe marginal conditions for an action a^* to be an equilibrium action. If Condition Z is satisfied, the function C is convex-extensible on $[0, 1]$, which implies that $C(a) = \tilde{C}(a)$ for all $a \in \mathcal{A}^{feas}$. Corollary 3 in the appendix shows that this property is inherited by the function $C_i(\cdot; \bar{a})$ given any $\bar{a} \in \mathcal{A}^{feas}$. Define $MC_i^-(a; \bar{a}) \equiv \partial^- \tilde{C}_i(a; \bar{a})$ and $MC_i^+(a; \bar{a}) = \partial^+ \tilde{C}_i(a; \bar{a})$ and the marginal-cost correspondence at a given $\bar{a}$ by

$$MC_i(a; \bar{a}) = \left\{ x \middle| MC_i^-(a; \bar{a}) \leq x \leq MC_i^+(a; \bar{a}) \right\}.$$

An action a^* is an equilibrium action if and only if $\frac{1}{N}B \in MC_i(a^*; a^*)$. We can therefore figure out which actions are equilibrium actions by computing the marginal-cost correspondence

$MC_i(\bar{a};\bar{a})$ for all $\bar{a}$. Each intersection of this correspondence with the value $\frac{1}{N}B$ correspond to an equilibrium action. We will make use of this characterization of equilibrium actions in what follows.

4 Equilibrium, Efficiency, and Coordination Failures

Our first result in this section is that the market failures in common-agency games are generally more severe than in conventional principal-agent models. Specifically, the highest equilibrium in the common agency problem, a_H^*, is inefficient in that it results in actions that are no greater than second-best actions.

PROPOSITION 3. Suppose Conditions S and Z hold. The highest equilibrium action a_H^* is bounded from above by a^{SB}.

Proof of Proposition 3. See Appendix.

The inefficiency in equilibrium actions stems from potentially three sources. The first is the inability to contract on the provider's action, combined with some contracting costs.[10] This source accounts for the difference between a^{FB} and a^{SB}. The second potential source of inefficiency is free riding among the principals. This source arises because the agents' investments benefit *all* the principals—whether or not the principals offered contracts with strong or weak incentives. The resulting free riding produces a lower than optimal set of incentive contracts and so the highest equilibrium action will generally be less than the second-best action. The third potential source of inefficiency is due to possible coordination failures. Coordination failures occur when there are multiple equilibrium actions.

As we demonstrated in our illustrative example, not all common-agency games result in coordination failures. An important implication of our characterization of equilibrium actions is that if c is well-behaved, then there is a unique equilibrium action. When c is well-behaved, $MC_i(\bar{a};\bar{a})$ is a singleton and is equal to $c'(\bar{a}) + R'(\bar{a})$, both of which are increasing in $\bar{a}$. We will say that there is a **sticking point** at a if $\partial^-\tilde{c}(a) \neq \partial^+\tilde{c}(a)$. Proposition 4 provides necessary conditions for there to be multiple equilibrium actions.

PROPOSITION 4. *Suppose Conditions S and Z hold. If there are multiple equilibrium actions, a_L^* and $a_H^* > a_L^*$, then there is a sticking point at a_L^*. If, in addition, Condition W holds, then there is a unique equilibrium action a^*.*

[10] In our set-up the contracting costs emerge from our stipulation that all elements of the incentive contract are non-negative.

Proof of Proposition 4. See Appendix.

This result means that the possibility of principals coordinating on an inefficient action when a more efficient equilibrium exists can only arise when there is a nondifferentiability in the (convexication of the) function mapping the desired action to the amount a single principal would have to compensate the agent to induce that action. This nondifferentiability condition appears to be a narrow and technical one, but it has broad and important economic implications. For example, it is satisfied in the case of discrete investments or (as in our model in Section 2) when investments have a discrete component such as a fixed cost. In the healthcare context, innovations involving new organizational structures, care processes or information technologies appear likely to meet the nondifferentiability criteria. The criteria will manifestly not be satisfied, however, in the case most studied by prior common-agency models—when the cost of the agent's action is continuous. Thus coordination failures should not arise when principals are trying to elicit greater effort and attention from a common-agent, unless effort/attention involves a substantial fixed cost.

5 Policy Interventions

The source of inefficiency—free riding or coordination failure—matters for policy. In this section we analyze how an ACO-style market intervention affects equilibrium incentives and investment. Through ACO policies, Medicare influences providers' incentives to invest in efficient care practices by sharing realized savings. We show that in the absence of coordination failure—that is, when the inefficiency is driven by free riding—the ACO intervention has the unintended consequence of reducing the incentives for investments in efficient care offered by other payers in the market. In some instances the ACO intervention completely crowds out private incentives and has no impact on provider investment levels at all. In the presence of coordination failure, however, ACO interventions have the potential to eliminate inefficient equilibria completely and trigger a shift towards stronger incentives and larger investments in efficient care. The particular market failure driving the inefficiency therefore has dramatic implications for the likely consequences of policy interventions like ACOs.

We model the ACO intervention as an additional payment τ to the provider in the event of success. Including incentives offered by the payers, the provider's payoff is therefore $w+\tau$ if $y = 1$ and zero otherwise. To make clear the equilibrium actions depend on the ACO intervention, denote the least equilibrium action $a_L^*(\tau)$ and the highest equilibrium action

$a_H^*(\tau)$. The equilibrium aggregate contract will be $w^*(\tau)$.

The consequences of the ACO incentive τ for equilibrium investment levels depends on what they would be in the absence of an intervention. First, in settings in which there are no sticking points—that is, where free-riding is the only source of inefficiency introduced by common agency—the ACO intervention has the perverse effect of reducing the incentives for investment offered by the payers, as the following proposition establishes.

PROPOSITION 5. *Suppose Conditions S, Z, and W hold. Then for each τ, there is a unique aggregate equilibrium contract $w^*(\tau)$, which is decreasing in τ.*

Proof of Proposition 5. See Appendix.

Condition W implies that c is well-behaved and that there are no sticking points such as those resulting from fixed costs in the agent's cost function, and by Proposition 4, there is a unique equilibrium action. In this setting the market failure responsible for inefficiently low incentives for investment is free riding among the multiple principals. The result says that in such settings market interventions such as ACO policies will partially crowd out private incentives. When free-riding drives the inefficiency, policies may be less effective than hoped.

When coordination failures drive the inefficiency, however, ACO-style policies have the potential to trigger large changes in equilibrium investment in efficient care, as the following proposition shows.

PROPOSITION 6. *Suppose Conditions S and Z hold and suppose $a_L^*(0) < a_H^*(0)$. Then there exists some τ^* such that $w_L^*(\tau) \geq w_H^*(0)$ for all $\tau > \tau^*$.*

Proof of Proposition 6. See Appendix.

Proposition 6 implies that when there are multiple equilibrium actions in the absence of an ACO intervention—which by Proposition 4 requires the existence of sticking points or fixed costs in the cost function—then there exists an ACO payout capable of ensuring that the least equilibrium exceeds the original highest equilibrium action. Thus in a setting where prior to an ACO intervention payers had been coordinating on an inefficient equilibrium, the introduction of ACOs has the potential to jump-start incentive provision by private payers.

5.1 Discussion

The model's key implication for healthcare fragmentation is Proposition 4, which provides a formal economic mechanism by which healthcare may remain persistently and inefficiently fragmented. The result suggests that payers may provide little incentive for, and providers may in turn undertake little investment in organizational innovations required for integrated care, even though doing so could be Pareto-improving. The result says that a necessary condition for this to be the case is that the cost function a single payer would face to induce a given level of effort from the provider be nondifferentiable at the lower level of effort.

Put differently, Proposition 4 implies that in a common-agency setting, the efficiency properties of outcomes depend critically on the nature of the task principals seek to elicit from agents. The numerous fixed, indivisible investments in the form of health IT infrastructure, and measurement and management systems involved in adopting efficient integrated care would lead to very different outcomes than incentives aimed towards eliciting more effort or attention from agents. Intuitively, half of an electronic health record system delivers no benefit so these sort of investments are inherently "lumpy." In this case incremental incentives that are insufficient to cover the jump in costs will not increase investment, and thus payers will not offer them. The result may be an equilibrium with virtually no incentives to undertake the organizational innovation. In the healthcare setting, this equilibrium corresponds to a status quo of fee-for-service payment arrangements with little coordination or integration. Only if the aggregate incentives offered by the other payers are high enough will a payer be willing to offer the higher incentives needed to induce the higher investment level from the provider. On the other hand, Proposition 4 also shows that if the investments required to adopt efficient, integrated care are, in fact, incremental or continuous in nature, then coordination failures will not arise. Investment may still be inefficiently low (i.e., less than second-best), but because of free riding among the payers, not because of coordination failures.

The model shows that the nature of the required investments, and thus the source of inefficient fragmentation—coordination failure or moral hazard—also matters for policy. Proposition 6 says that in the presence of coordination failures, policy interventions such as ACOs, which provide additional incentives to providers to invest in efficient care, have the potential to eliminate inefficient equilibria and shift the market towards an equilibrium with substantial investments and for efficient care. In this light, health reform initiatives such as ACOs or health information technology subsidies can be seen as effective policy responses to coordination failures. On the other hand, in the absence of coordination failures, such policies are more likely to crowd out private incentives for investments in efficient care, as Proposition 5 shows.

6 Conclusion

The core of our argument is that integration in a fragmented healthcare delivery system requires substantial up-front investments and that these may not be forthcoming when the health insurance system has weak incentives to control costs. The weak incentives persist because of common-agency problems among insurers.

Specifically we find that common agency leads to weaker incentives than found in conventional principal-agent settings. This heightened inefficiency has two sources. The first is free-riding among payers because provider investments in integrated care benefit all the payers whether or not they offer providers cost-saving incentives. The second is coordination failures in which payers become stuck in a highly inefficient equilibrium and offer contracts having no incentives at all even though Pareto dominating alternative equilibria are feasible.

The source of inefficiency—free-riding or coordination failure—matters for policy. In a setting with only free-riding, Medicare's ACO intervention has the perverse effect of reducing incentives for investment offered by private payers. Things are different when coordination failures drive the inefficiencies. Here ACO-style policies have the potential to trigger jump-start changes in incentives throughout the private sector—provided that the intervention is sufficiently high-powered and aggressive. If not, the policies may have no effect at all.

Our results also have implications for the applied theory literature on common-agency. We find that the possibility of multiple equilibria and coordination failure depend critically on the sort of actions incentive contracts seek to elicit. When principals wish to encourage more effort, attention or similarly continuous actions, coordination failures will not appear. Coordination failures may appear when contracts are aimed at inducing innovations involving fixed costs or lumpy investments. Efforts to move from fragmented to integrated care delivery clearly require providers to undertake these sorts of investments.

Investigating other common-agency applications involving these sorts of investments is an important area for future research.

Appendix

This Appendix develops the arguments to prove the results in Sections 3 and 4, including equilibrium existence and our characterization of the set of equilibria. The first subsection develops Theorem 1, which characterizes the set of equilibrium actions $\mathcal{A}^*$ as the solution to a self-generating maximization program. In particular, we show that $a^* \in \mathcal{A}^*$ if and only if

$$a^* \in \hat{a}\left(a^*\right) \equiv \underset{a \in \mathcal{A}^{feas}}{\operatorname{argmax}} \tilde{\Lambda}\left(a, a^*\right).$$

The second subsection shows that the operator $\hat{a}\left(\cdot\right)$ is monotone, so it always has at least one fixed point, proving Theorem 2.

Aggregate Representation

In this subsection, we develop necessary and sufficient conditions for an action a^* to be an equilibrium action. The results of this subsection hold for more general output spaces and more general contracting spaces than we assume in our main model. Further, while our sufficient conditions require Condition S to be satisfied, our necessary conditions do not. The results in the following subsections make use of our assumptions that the output space is binary and contracts are nonnegative and nondecreasing.

Before we outline the argument, we define some notation and terms that will be convenient in the arguments. First, denote the **distribution over output induced by action** a by $\phi\left(a\right)$, and the Agent's **optimal action given aggregate contract** w by $a\left(w\right)$. Recall our tie-breaking assumption on the Agent's choice: if the Agent is indifferent among two or more actions, he chooses the highest action he is indifferent among. The **set of feasible contracts that support action** a is the subdifferential of c at a:

$$\partial c\left(a\right) = \left\{w \in \mathcal{W} : w \cdot \phi\left(a\right) - c\left(a\right) \geq w \cdot \phi\left(a'\right) - c\left(a'\right) \text{ for all } a' \in \mathcal{A}\right\},$$

where $w \cdot \phi\left(a\right)$ is the inner product of w and $\phi\left(a\right)$. A **cost-minimizing contract for** a is denoted by w_a^*, and it solves

$$w_a^* \in \underset{w \in \mathcal{W}}{\operatorname{argmin}} \left\{w \cdot \phi\left(a\right) : w \in \partial c\left(a\right)\right\}.$$

The set of **feasible actions relative to** $\bar{w}$ is denoted by

$$\mathcal{A}_{\bar{w}}^{feas} = \left\{a \in \mathcal{A}^{feas} : w \in \partial c\left(a\right) \text{ for some } w \in \mathcal{W} + \left(1 - 1/N\right)\bar{w}\right\}.$$

A **cost-minimizing contract for** $a \in \mathcal{A}_{\bar{w}}^{feas}$ **relative to** $\bar{w}$, denoted by $w_{a,\bar{w}}^*$, solves

$$w_{a,\bar{w}}^* \in \underset{w \in \mathcal{W} + \left(1-1/N\right)\bar{w}}{\operatorname{argmin}} \left\{w \cdot \phi\left(a\right) : w \in \partial c\left(a\right)\right\}$$

Our analysis in this subsection proceeds in four steps:

1. We first extend the analysis of Martimort and Stole (2012) to a setting in which the aggregate is not bijective, we will show that when Condition S is satisfied, w^* is an equilibrium aggregate contract if and only if

$$w^* \in \hat{w}\left(w^*\right) = \underset{w \in \mathcal{W}+(1-1/N)w^*}{\operatorname{argmax}} \left(\frac{1}{N}B - w + \left(1 - \frac{1}{N}\right)w^*\right) \cdot \phi\left(a\left(w\right)\right).$$

2. We then extend Theorem 1 of Bernheim and Whinston (1986b) to this setting, showing that if $w^* \in \hat{w}\left(w^*\right)$, then w^* is a cost-minimizing contract for some action $a \in \mathcal{A}^{feas}$.

3. Next, we show that given any aggregate contract $\bar{w}$, any $w \in \hat{w}\left(\bar{w}\right)$ will be a cost-minimizing contract for some action relative to $\bar{w}$.

4. Finally, we show that $w^* \in \hat{w}\left(w^*\right)$ if and only if $w^* = w^*_{a^*}$ where

$$a^* \in \hat{a}\left(a^*\right) = \underset{a \in \mathcal{A}^{feas}}{\operatorname{argmax}} \frac{1}{N}Ba - C_i\left(a, a^*\right).$$

In proceeding from the self-generating maximization program derived in Step 1 to the simpler self-generating maximization program derived in Step 4, Step 2 is restricts the domain of the contracting space that needs to be searched over, and Step 3 restricts the range. In particular, Steps 2 and 3 show that both the domain and the range can, without loss of generality, be restricted to a set that is isomorphic to the set of incentive-feasible actions, which is a compact subset of $[0, 1]$.

Step 1 Given w_{-i}, P_i chooses w_i to solve

$$\max_{w_i \in \mathcal{W}} \left(B_i - w_i\right) \cdot \phi\left(a\left(w\right)\right) = \max_{w_i \in \mathcal{W}} u_i\left(w_i, w\right).$$

We can instead think of P_i as choosing $w = w_i + \sum_{j \neq i} w_j$. Then $w_i \in \mathcal{W}$ if and only if $w \in \mathcal{W} + \sum_{j \neq i} w_j$. P_i's problem is therefore

$$\max_{w \in \mathcal{W}+\sum_{j \neq i} w_j} u_i\left(w - \sum_{j \neq i} w_j, w\right).$$

If w^* is an equilibrium aggregate contract, then for each i,

$$w^* \in \underset{w \in \mathcal{W}+w^*-w_i^*}{\operatorname{argmax}} u_i\left(w - \sum_{j \neq i} w_j^*, w\right).$$

Since w^* solves this program for each i, it also solves these programs on average:

$$w^* \in \underset{w \in \cap_{i=1}^N \mathcal{W}+w^*-w_i^*}{\operatorname{argmax}} \sum_{i=1}^N u_i \left(w - \sum_{j \neq i} w_j^*, w \right).$$

Define the quantity

$$\Lambda(w, \bar{w}) = \frac{1}{N} \sum_{i=1}^N u_i \left(w - \sum_{j \neq i} \bar{w}_j, w \right) = \frac{1}{N} \sum_{i=1}^N \left(B_i - \left(w - \sum_{j \neq i} \bar{w}_j \right) \right) \cdot \phi(a(w)),$$

or

$$\Lambda(w, \bar{w}) = \left(\frac{1}{N} B - w + \left(1 - \frac{1}{N} \right) \bar{w} \right) \cdot \phi(a(w)).$$

Therefore, if $w^* \in \mathcal{W}$ is an aggregate equilibrium contract, then

$$w^* \in \underset{w \in \cap_{i=1}^N \mathcal{W}+w^*-w_i^*}{\operatorname{argmax}} \Lambda(w, w^*)$$

for some $w_1^*, \ldots, w_N^*$ such that $\sum_{i=1}^N w_i^* = w^*$. This leads to the following Lemma.

LEMMA 1. *If w^* is an equilibrium aggregate contract, then*

$$w^* \in \underset{w \in \mathcal{W}(w^*)}{\operatorname{argmax}} \Lambda(w, w^*),$$

where

$$\mathcal{W}(w^*) = \bigcup_{\substack{w_1^*+\cdots+w_N^*=w^* \\ w_i^* \in \mathcal{W}}} \bigcap_{i=1}^N \mathcal{W} + w^* - w_i^*.$$

The set $\mathcal{W}(w^*)$ seems unwieldy, but it is simplified by the fact that $\mathcal{W}$ is a convex cone, as the following Lemmas show.

LEMMA 2. *Suppose $X_1, \ldots, X_N$ are convex cones. Then*

$$\bigcap_{i=1}^N X_i \subset \sum_{i=1}^N \frac{1}{N} X_i.$$

PROOF OF LEMMA 2. Take $x \in \cap_{i=1}^N X_i$. Then $x \in X_i$. Since X_i is a convex cone, $\frac{1}{N} x \in X_i$ for each i. Therefore, $x = \sum_{i=1}^N \frac{1}{N} x \in \sum_{i=1}^N \frac{1}{N} X_i$.■

LEMMA 3. *For any $\bar{x} \in X$,*

$$\bigcup_{\substack{\bar{x}_1+\cdots+\bar{x}_N=\bar{x} \\ \bar{x}_i \in X}} \bigcap_{i=1}^{N} X + x_i = X + \frac{1}{N}\bar{x}.$$

PROOF OF LEMMA 3. To prove ($\subset$), take $x_1, \ldots, x_N$ such that $\sum_{i=1}^{N} x_i = \bar{x}$. From Lemma 2,

$$\bigcap_{i=1}^{N} X + x_i \subset \sum_{i=1}^{N} \frac{1}{N}\left(X + x_i\right).$$

Since X is a convex cone, $\frac{1}{N}X = X$ and $\sum_{i=1}^{N} X = X$, so

$$\sum_{i=1}^{N} \frac{1}{N}\left(X + x_i\right) = X + \sum_{i=1}^{N} \frac{1}{N}x_i = X + \frac{1}{N}\bar{x},$$

since x_i are scalars. Since $x_1, \ldots, x_N$ were arbitrary, this holds for all $x_1, \ldots, x_N$ for which $\sum_{i=1}^{N} x_i = \bar{x}$. To prove ($\supset$), suppose $x \in X + \frac{1}{N}\bar{x}$. Let $x_i = \frac{1}{N}x$. Then $\sum_{i=1}^{N} x_i = \bar{x}$, and $\frac{1}{N}x \in X + \frac{1}{N}x$.■

The following Corollary results from Lemmas 2 and 3.

COROLLARY 1. *For any $\bar{w} \in W$,*

$$\mathcal{W}\left(\bar{w}\right) = \mathcal{W} + \left(1 - 1/N\right)\bar{w}.$$

Putting these results together, if $w^* \in \mathcal{W}$ is an equilibrium aggregate contract, then

$$w^* \in \underset{w \in \mathcal{W}+(1-1/N)\bar{w}}{\operatorname{argmax}} \Lambda\left(w, \bar{w}\right).$$

The converse holds if Condition S is satisfied, as the following Proposition shows.

PROPOSITION 1A. *Suppose Condition S is satisfied. Then $w^* \in W$ is an equilibrium aggregate contract if and only if*

$$w^* \in \underset{w \in \mathcal{W}+(1-1/N)w^*}{\operatorname{argmax}} \Lambda\left(w, w^*\right).$$

PROOF OF PROPOSITION 1A. Necessity follows from Lemma 1 and Corollary 1. Now, suppose w^* solves this program. Let $w_i^* = \frac{1}{N} w^*$ for $i = 1, \ldots, N$. P_i's program is therefore

$$\begin{aligned} & \max_{w \in \mathcal{W} + (1-1/N) w^*} \left(B_i - \left(w - \sum_{j \neq i} w_j^* \right) \right) \cdot \phi\left(a\left(w\right)\right) \\ = & \max_{w \in \mathcal{W} + (1-1/N) w^*} \left(B_i - \left(w - \left(1 - \frac{1}{N}\right) w^* \right) \right) \cdot \phi\left(a\left(w\right)\right). \end{aligned}$$

Since $B_i = \frac{1}{N} B$, this program is equivalent to

$$\max_{w \in \mathcal{W} + (1-1/N) w^*} \left(\frac{1}{N} B - w + \left(1 - \frac{1}{N}\right) w^* \right) \cdot \phi\left(a\left(w\right)\right) = \max_{w \in \mathcal{W} + (1-1/N)} \Lambda\left(w, w^*\right).$$

Since w^* solves the aggregate problem, it also solves each Principal's problem.■

Proposition 1A completes the first step of the analysis. One immediate Corollary of Corollary of Proposition 1A is that when Condition S is satisfied, if w^* is an equilibrium aggregate contract, there is a symmetric equilibrium in which w^* is the associated equilibrium aggregate contract.

COROLLARY 2. *Suppose Condition S is satisfied. If w^* is an equilibrium aggregate contract, there is a symmetric equilibrium in which each Principal chooses $w_i^* = \frac{1}{N} w^*$.*

Step 2 We now turn to the second step, showing that any equilibrium aggregate contract must be a cost-minimizing contract for some action, as long as the contracting space $\mathcal{W}$ is a convex cone. This result is captured in Proposition 2A. Note in particular that Proposition 2A does not require that Condition S is satisfied. It therefore shows that Theorem 1 of Bernheim and Whinston (1986b) can be extended to environments with restricted contracting spaces, as long as the contract space is a convex cone, and players are risk-neutral.

PROPOSITION 2A. *Suppose Condition S is satisfied and*

$$w^* \in \underset{w \in \mathcal{W} + (1-1/N) w^*}{\operatorname{argmax}} \Lambda\left(w, w^*\right).$$

Then $w^* = w_a^*$ for some $a \in A^{feas}$.

PROOF OF PROPOSITION 2A. Suppose

$$w^* \in \underset{w \in \mathcal{W} + (1-1/N) w^*}{\operatorname{argmax}} \left(\frac{1}{N} B - w + \left(1 - \frac{1}{N}\right) w^* \right) \cdot \phi\left(a\left(w\right)\right).$$

Then w^* implements some action $a^* = a\left(w^*\right)$. In order to get a contradiction, suppose there is some contract $\hat{w} \in \mathcal{W}$ that also implements a^* but $\hat{w} \cdot \phi\left(a^*\right) < w^* \cdot \phi\left(a^*\right)$. First, note that if $a\left(w^*\right) = a\left(\hat{w}\right) = a$, then for any $\lambda \in [0, 1]$, $a\left(\left(1 - \lambda\right) \hat{w} + \lambda w^*\right) = a$. That is, if two contracts implement the same action, then so does any convex combination. This is because the Agent is risk-neutral.

There are then two cases. First if $\hat{w} \in \mathcal{W}+(1-1/N)\, w^*$, then $\hat{w}$ is feasible and does better than w^* in the aggregate program, so $w^* \notin \hat{w}(w^*)$. Next, suppose $\hat{w} \notin \mathcal{W}+(1-1/N)\, w^*$. Then, we can always draw a line segment connecting $\hat{w}$ and w^*. This line segment will intersect $\mathcal{W}+(1-1/N)\, w^*$. Take $\hat{\lambda}$ such that $\hat{\lambda}\hat{w}+(1-1/N)\, w^*$ is in the boundary of $\mathcal{W}+(1-1/N)\, w^*$. Then $\hat{\lambda}\hat{w}+\left(1-\hat{\lambda}\right) w^*$ is feasible and cheaper than w^*, so $w^* \notin \hat{w}(w^*)$.■

Proposition 2A effectively restricts the domain over which we have to search when looking for fixed points of the $\hat{w}(\cdot)$ operator. In particular, we only have to look for w_a^* such that $w_a^* \in \hat{w}(w_a^*)$.

Step 3 We will now proceed to the third step, which shows that any contract in $\hat{w}(\bar{w})$ is cost-minimizing relative to $\bar{w}$. This result is described in the following Lemma.

LEMMA 4. *Suppose Condition S is satisfied and $w \in \hat{w}(\bar{w})$. Then w is cost-minimizing for some a relative to $\bar{w}$.*

PROOF OF LEMMA 4. In order to get a contradiction, suppose w is not cost-minimizing for any action relative to $\bar{w}$. Let $a = a(w)$. Since $w_{a,\bar{w}}^*$ is a cost-minimizing contract for a relative to $\bar{w}$, it is feasible, and we have that $w_{a,\bar{w}}^* \cdot \phi(a) < w \cdot \phi(a)$, which implies that

$$\left(\frac{1}{N}B - w_{a,\bar{w}}^* + \left(1-\frac{1}{N}\right)\bar{w}\right)\cdot\phi(a) > \left(\frac{1}{N}B - w + \left(1-\frac{1}{N}\right)\bar{w}\right)\cdot\phi(a),$$

which contradicts the claim that $w \in \hat{w}(\bar{w})$.■

One implication of Lemma 4 is that in solving for $\hat{w}(\bar{w})$, it is without loss of generality to consider cost-minimizing contracts relative to $\bar{w}$. That is,

$$\underset{w\in\mathcal{W}+(1-1/N)\bar{w}}{\operatorname{argmax}}\ \Lambda(w,\bar{w}) = \underset{w_{a,\bar{w}}^*\in\mathcal{W}+(1-1/N)\bar{w}}{\operatorname{argmax}}\ \Lambda(w,\bar{w}).$$

Proposition 2A restricts the domain over which we have to search when looking for fixed points of the $\hat{w}(\cdot)$ operator. Lemma 4 shows that, given a cost-minimizing contract $w_{\bar{a}}^*$, we can restrict attention to looking for cost-minimizing contracts relative to $w_{\bar{a}}^*$. Denote a cost-minimizing contract for action a relative to $w_{\bar{a}}^*$ by $w_{a,\bar{a}}^*$, and denote the set of feasible actions relative to $w_{\bar{a}}^*$ by $\mathcal{A}_{\bar{a}}^{feas}$. Without loss of generality, we can therefore restrict attention to a domain and a range that are each isomorphic to $\mathcal{A}^{feas}$.

Step 4 Before we can state and prove Theorem 1A, define the function

$$\bar{C}_i(a,\bar{a}) = w_{a,\bar{a}}^*\cdot\phi(a) - \left(1-\frac{1}{N}\right) w_{\bar{a}}^*\cdot\phi(a).$$

Our main characterization theorem follows.

THEOREM 1A. *Suppose Condition S is satisfied.* a^* *is an equilibrium action if and only if*

$$a^* \in \hat{a}(a^*) = \underset{a \in \mathcal{A}^{feas}_{a^*}}{\operatorname{argmax}} \frac{1}{N} Ba - \bar{C}_i(a, a^*). \tag{9}$$

PROOF OF THEOREM 1A. Suppose a^* is an equilibrium action. Then $w^*_{a^*}$ is an equilibrium aggregate contract (Lemma 3), which in turn implies that $w^*_{a^*} \in \hat{w}(w^*_{a^*})$ (Lemma 1). Since all $w \in \hat{w}(w^*_{a^*})$ are cost-minimizing relative to $w^*_{a^*}$ (Lemma 4), $w^*_{a^*} \in \hat{w}(w^*_{a^*})$ implies that $a^* \in \hat{a}(a^*)$. Conversely, suppose $a^* \in \hat{a}(a^*)$. Then $w^*_{a^*}$ is the best cost-minimizing contract relative to $w^*_{a^*}$, which implies that $w^*_{a^*} \in \hat{w}(w^*_{a^*})$ (Lemma 4).■

Theorem 1A shows that instead of solving for fixed points of $\hat{w}(\cdot)$, an equivalent problem is the simpler problem of solving for fixed points of $\hat{a}(\cdot)$. This problem is simpler, because the action space is simpler than the contracting space.

Monotonicity

In this subsection, we show that the operator $\hat{a}(\cdot)$ is increasing, which in turn allows us to make use of monotonicity-based fixed-point theorems to establish the existence of an equilibrium action and to derive some properties of the set of equilibrium effort levels. The analysis in this subsection proceeds in four steps. Recall that we have denoted $\partial^- \tilde{c}(a)$ and $\partial^+ \tilde{c}(a)$ to be the left- and right-derivative of $\tilde{c}$ at a, where $\tilde{c}(a)$ is the convexification of the Agent's cost function, $c(a)$. $\partial^- \tilde{c}(a)$ and $\partial^+ \tilde{c}(a)$ are also the inf and sup of the subdifferential of c at a. By convention, we will denote $\partial^- \tilde{c}(0) = 0$.

1. For all $a \in \mathcal{A}^{feas}$, $w^*_a(0) = 0$ and $w^*_a(1) = \partial^- \tilde{c}(a)$, so that we can write $s^*_a = 0$ and $b^*_a = \partial^- \tilde{c}(a)$.

2. Next, we will show that

$$\hat{a}(\bar{a}) = \underset{a \in \mathcal{A}^{feas}}{\operatorname{argmax}} \frac{1}{N} Ba - C_i(a, \bar{a}),$$

   where

$$C_i(a, \bar{a}) = \max\{C(a), (1 - 1/N) b^*_{\bar{a}} a_{\min}(\bar{a})\} - (1 - 1/N) b^*_{\bar{a}} a.$$

3. We will then show that $C_i(a, \bar{a})$ satisfies decreasing differences in $(a, \bar{a})$ on $\mathcal{A}^{feas}$, which, by Topkis's (1998) theorem, implies that $\hat{a}(\cdot)$ is increasing.

4. Finally, by Zhou's (1994) extension of Tarski's (1955) fixed-point theorem, the set of fixed points of $\hat{a}(\cdot)$ is nonempty and compact.

Step 1 Lemma 5 establishes the first result, solving for the set of cost-minimizing contracts of the single-principal problem in our setting.

LEMMA 5. *If $w \in \partial c(a)$, then there is a $\tilde{w} \in W$ with $\tilde{w}(0) = 0$ such that $\tilde{w} \in \partial c(a)$ and $\tilde{w} \cdot \phi(a) \leq w \cdot \phi(a)$, so for each $a \in A^{feas}$, $w_a^*(0) = 0$. Further, $w_a^*(1) = \partial^- \tilde{c}(a)$.*

PROOF OF LEMMA 5. Suppose $w \in \partial c(a)$. Then

$$\partial^- \tilde{c}(a) \leq w(1) - w(0) \leq \partial^- \tilde{c}(a).$$

Define $\tilde{w} = (0, w(1), w(0))$. By construction, $\tilde{w} \in \partial c(a)$, and

$$[w(1) - w(0)]\, a \leq w(0)(1-a) + w(1)\, a = w(0) + [w(1) - w(0)]\, a,$$

since $w(0) \geq 0$. This inequality is strict, unless $w(0) = 0$, so $w_a^*(0) = 0$.

Next, in order for $(0, w(1)) \in \partial c(a)$, it has to be the case that $\partial^- \tilde{c}(a) \leq w(1) \leq \partial^+ \tilde{c}(a)$. Let $\tilde{w} = (0, \partial^- \tilde{c}(a))$. Then $\tilde{w} \in \partial c(a)$, and

$$\partial^- \tilde{c}(a)\, a \leq w(1)\, a,$$

and therefore $w_a^*(1) = \partial^- \tilde{c}(a)$.

We will subsequently denote $s_a^* = w_a^*(0) = 0$ and $b_a^* = w_a^*(1) - w_a^*(0) = \partial^- \tilde{c}(a)$.

Step 2 We now show that the maximization program (9) defined in Theorem 1A is solution-equivalent to an unconstrained maximization program obtained by replacing $\bar{C}_i(a, \bar{a})$ with

$$C_i(a, \bar{a}) = \max\{C(a), (1 - 1/N)\, b_{\bar{a}}^* a_{\min}(\bar{a})\} - (1 - 1/N)\, b_{\bar{a}}^* a.$$

LEMMA 6. *For all $\bar{a} \in A^{feas}$, the solutions to maximization program defined in (9), $\hat{a}(\bar{a})$, coincide with*

$$\underset{a \in \mathcal{A}^{feas}}{\operatorname{argmax}} \frac{1}{N} B a - C_i(a, \bar{a}).$$

PROOF OF LEMMA 6. In this setting, we have $\mathcal{A}_{\bar{a}}^{feas} = \mathcal{A}^{feas} \cap [a_{\min}(\bar{a}), 1]$, $\bar{C}_i(a, \bar{a}) = b_{a,\bar{a}}^* a - (1 - 1/N)\, b_{\bar{a}}^* a$, and

$$b_{a,\bar{a}}^* a = \begin{cases} (1 - 1/N)\, b_{\bar{a}}^* a & a = a_{\min}(\bar{a}) \\ C(a) & a > a_{\min}(\bar{a}). \end{cases}$$

By definition of $a_{\min}(\bar{a})$, for all $a \leq a_{\min}(\bar{a})$, $b_a^* \leq (1 - 1/N)\, b_{\bar{a}}^*$. We therefore have that for all $a \in \mathcal{A}_{\bar{a}}^{feas}$, $\bar{C}_i(a, \bar{a}) = C_i(a, \bar{a})$. Finally, for all $a < a_{\min}(\bar{a})$,

$$C_i(a, \bar{a}) \geq C_i(a_{\min}(\bar{a}), \bar{a}),$$

so that

$$\underset{a\in\mathcal{A}_{\bar{a}}^{feas}}{\operatorname{argmax}}\frac{1}{N}Ba-C_i\left(a,\bar{a}\right)=\underset{a\in\mathcal{A}^{feas}}{\operatorname{argmax}}\frac{1}{N}Ba-C_i\left(a,\bar{a}\right),$$

which completes the proof.■

Step 3 If $C_i\left(a,\bar{a}\right)$ satisfies decreasing differences in $\left(a,\bar{a}\right)$ on $\mathcal{A}^{feas}$, then $\tilde{\Lambda}\left(a,\bar{a}\right)$ satisfies increasing differences in $\left(a,\bar{a}\right)$ on $\mathcal{A}^{feas}$. This is the case, as the following Lemma shows.

LEMMA 7. $C_i\left(a,\bar{a}\right)$ satisfies decreasing differences in $\left(a,\bar{a}\right)$ on $\mathcal{A}^{feas}$. Consequently, $\tilde{\Lambda}\left(a,\bar{a}\right)$ satisfies increasing differences in $\left(a,\bar{a}\right)$ on $\mathcal{A}^{feas}$.

PROOF OF LEMMA 7. Let $a\geq a'$ and $\bar{a}\geq\bar{a}'$ with $a,a',\bar{a},\bar{a}'\in\mathcal{A}^{feas}$. Define the difference $\Delta\left(\bar{a}\right)\equiv C_i\left(a,\bar{a}\right)-C_i\left(a',\bar{a}\right)$ and the value $\delta=\left(b_{\bar{a}}^*-b_{\bar{a}'}^*\right)\left(a-a'\right)\geq 0$. There are six cases we need to consider. They are tedious but straightforward.
Case 1. If $C\left(a\right)\geq C\left(a'\right)\geq\left(1-1/N\right)b_{\bar{a}}^*a_{\min}\left(\bar{a}\right)\geq\left(1-1/N\right)b_{\bar{a}'}^*a_{\min}\left(\bar{a}'\right)$, then

$$\Delta\left(\bar{a}\right)-\Delta\left(\bar{a}'\right)=-\delta\leq 0$$

Case 2. If $C\left(a\right)\geq\left(1-1/N\right)b_{\bar{a}}^*a_{\min}\left(\bar{a}\right)\geq C\left(a'\right)\geq\left(1-1/N\right)b_{\bar{a}'}^*a_{\min}\left(\bar{a}'\right)$, then

$$\Delta\left(\bar{a}\right)-\Delta\left(\bar{a}'\right)=C\left(a'\right)-\left(1-1/N\right)b_{\bar{a}}^*a_{\min}\left(\bar{a}\right)-\delta\leq 0$$

Case 3. If $C\left(a\right)\geq\left(1-1/N\right)b_{\bar{a}}^*a_{\min}\left(\bar{a}\right)\geq\left(1-1/N\right)b_{\bar{a}'}^*a_{\min}\left(\bar{a}'\right)\geq C\left(a'\right)$, then

$$\Delta\left(\bar{a}\right)-\Delta\left(\bar{a}'\right)=\left(1-1/N\right)b_{\bar{a}'}^*a_{\min}\left(\bar{a}'\right)-b_{\bar{a}}^*a_{\min}\left(\bar{a}\right)-\delta\leq 0$$

Case 4. If $\left(1-1/N\right)b_{\bar{a}}^*a_{\min}\left(\bar{a}\right)\geq C\left(a\right)\geq C\left(a'\right)\geq\left(1-1/N\right)b_{\bar{a}'}^*a_{\min}\left(\bar{a}'\right)$, then

$$\Delta\left(\bar{a}\right)-\Delta\left(\bar{a}'\right)=C\left(a'\right)-C\left(a\right)-\delta\leq 0$$

Case 5. If $\left(1-1/N\right)b_{\bar{a}}^*a_{\min}\left(\bar{a}\right)\geq C\left(a\right)\geq\left(1-1/N\right)b_{\bar{a}'}^*a_{\min}\left(\bar{a}'\right)\geq C\left(a'\right)$, then

$$\Delta\left(\bar{a}\right)-\Delta\left(\bar{a}'\right)=\left(1-1/N\right)b_{\bar{a}'}^*a_{\min}\left(\bar{a}'\right)-C\left(a\right)-\delta\leq 0$$

Case 6. If $\left(1-1/N\right)b_{\bar{a}}^*a_{\min}\left(\bar{a}\right)\geq\left(1-1/N\right)b_{\bar{a}'}^*a_{\min}\left(\bar{a}'\right)\geq C\left(a\right)\geq C\left(a'\right)$, then

$$\Delta\left(\bar{a}\right)-\Delta\left(\bar{a}'\right)=-\delta\leq 0.$$

Since $\tilde{\Lambda}\left(a,\bar{a}\right)=\frac{1}{N}Ba-C_i\left(a,\bar{a}\right)$, $\tilde{\Lambda}\left(a,\bar{a}\right)$ satisfies increasing differences in $\left(a,\bar{a}\right)$ on $\mathcal{A}^{feas}$.■

We can therefore apply Topkis's theorem to show that $\hat{a}\left(\cdot\right)$ is increasing.

LEMMA 8. $\hat{a}\left(\cdot\right)$ is increasing on $\mathcal{A}^{feas}$.

PROOF OF LEMMA 8. Follows directly from Topkis's theorem.

The intuition behind Lemma 8 is that, given any cost-minimizing target contract, $w_{\bar{a}}^*$, each Principal P_i either wants to leave $\left(1-1/N\right)w_{\bar{a}}^*$ in place by contributing $w_i=0$, or they want to tup up $(1-1/N)\, w_{\bar{a}}^*$. If they choose to top it up, they will top it up to a cost-minimizing contract, which is feasible, because w_a^* is increasing in a.

Step 4 Our second theorem follows from Lemma 8.

THEOREM 2. Suppose Condition S holds. The set of equilibrium actions $\mathcal{A}^*$ is nonempty and compact.

PROOF OF THEOREM 2. By Lemma 8 and the fact that $\mathcal{A}^{feas}$ is a compact subset of $[0,1]$, $\hat{a}(\cdot)$ is a monotone operator on a complete lattice. By Zhou's (1994) extension of Tarki's fixed-point theorem to correspondences, the set of fixed points of $\hat{a}(\cdot)$ is a nonempty complete lattice, which in turn implies that $\mathcal{A}^*$ is a compact subset of $[0,1]$.$\blacksquare$

6.1 Efficiency, Coordination Failures, and Policy

LEMMA 9. *Let $f : R \to R$ be convex and nondecreasing, and let $g : A^{feas} \to R$ be convex-extensible on $[0,1]$. Then $f \circ g : A^{feas} \to R$ is convex-extensible on $[0,1]$.*

Proof of Lemma 9. Let $\tilde{g}$ be the convexification of g on $[0,1]$. Then $f \circ \tilde{g}$ is convex, and $f \circ g(a) = f \circ \tilde{g}(a)$ for all $a \in \mathcal{A}^{feas}$. Therefore, $f \circ g$ is convex-extensible on $[0,1]$.$\blacksquare$

COROLLARY 3. *For any $\bar{a} \in A^{feas}$, the function $C_i(\cdot;\bar{a})$ is convex-extensible on $[0,1]$.*

Proof of Corollary 3. Since $f(x) = \max\{x,k\}$ is convex and nondecreasing, Lemma 9 implies that $f \circ C(a)$ is convex-extensible on $[0,1]$. $C_i(\cdot;\bar{a})$ is the sum of an affine function and $f \circ C(a)$ for $k = (1-1/N)\, b_{\bar{a}}^* a_{\min}(\bar{a})$ and is therefore convex-extensible on $[0,1]$.$\blacksquare$

PROPOSITION 3. Suppose Conditions S and Z hold. The highest equilibrium action a_H^* is bounded from above by a^{SB}.

Proof of Proposition 3. TBA, but follows from the fact that $MC(a) \leq N \cdot MC_i(a)$ in the strong set order.

PROPOSITION 4. *Suppose Conditions S and Z hold. If c is well-behaved, then there is a unique equilibrium action a^*. If there are multiple equilibrium actions, a_L^* and $a_H^* > a_L^*$, then there is a sticking point at a_L^*.*

Proof of Proposition 4. TBA.

PROPOSITION 5. *Suppose Conditions S, Z, and W hold. Then for each τ, there is a unique aggregate equilibrium contract $w^*(\tau)$, which is decreasing in τ.*

Proof of Proposition 5. TBA.

PROPOSITION 6. *Suppose Conditions S and Z hold and suppose $a_L^*(0) < a_H^*(0)$. Then there exists some τ^* such that $w_L^*(\tau) \geq w_H^*(0)$ for all $\tau > \tau^*$.*

Proof of Proposition 6. TBA.

References

[1] Baicker, Katherine and Amitabh Chandra. 2011. Aspirin, Angioplasty, and Proton Beam Therapy: The Economics of Smarter Health Care Spending. In *Jackson Hole Economic Policy Symposium.*

[2] Baker, Laurence, Kate Bundorf, and Anne Royalty. 2014. Measuring Physician Practice Competition Using Medicare Data. In *Measuring and Modeling Health Care Costs*, eds. Ana Aizcorbe, Colin Baker, Ernst Berndt, and David Cutler. University of Chicago Press.

[3] Bernheim, B. Douglas and Michael Whinston. 1986a. Menu Auctions, Resource Allocation, and Economic Influence. *Quarterly Journal of Economics*, 101(1): 1-32.

[4] Bernheim, B. Douglas and Michael Whinston. 1986b. Common Agency. *Econometrica*, 54(4): 923-942.

[5] Besley, Timothy and Stephen Coate. 2001. Lobbying and Welfare in a Representative Democracy. *Review of Economic Studies*, 68: 67-82.

[6] Blumenthal, David. 2011. Wiring the Health System Origins and Provisions of a New Federal Program. *New England Journal of Medicine*, 365: 2323-2329.

[7] Burns, Lawton and Mark Pauly. 2002. Integrated Delivery Networks: A Detour on the Road to Integrated Health Care? *Health Affairs*, 21(4): 128-143.

[8] Cebul, Randall, James Rebitzer, Lowell Taylor, and Mark Votruba. 2008. Organizational Fragmentation and Care Quality in the U.S. Healthcare System. *Journal of Economic Perspectives*, 22(4): 93-114.

[9] Centers for Disease Control. 2010. National Ambulatory Care Survey: 2010 Summary Tables.

[10] Crosson, Francis. 2009. Medicare: The Place to Start Delivery System Reform. *Health Affairs*, 28(2): w232-w234.

[11] Dixit, Avinash, Gene Grossman, and Elhanan Helpman. 1997. Common Agency and Coordination: General Theory and Application to Government Policy Making. *Journal of Political Economy*, 105(4): 752-769.

[12] Glazer, Jacob and Thomas McGuire. 2002. Multiple Payers, Commonality and Free-Riding in Health Care: Medicare and Private Payers. *Journal of Health Economics*, 21(6): 1049-1069.

[13] Hussey, Peter, Eric Schneider, Robert Rudin, Steven Fox, Julie Lai, and Craig Pollack. 2014. Continuity and the Costs of Care for Chronic Disease. *JAMA Internal Medicine*, 174(5): 742-748.

[14] Innes, Robert. 1990. Limited Liability and Incentive Contracting with Ex-ante Action Choices. *Journal of Economic Theory*, 52: 45-67.

[15] Kirchsteiger, Georg and Andrea Prat. 2001. Inefficient Equilibria in Lobbying. *Journal of Public Economics*, 82(3): 349-375.

[16] Kiselman, Christer and Shiva Samieinia. Forthcoming. *Journal of Mathematical Sciences*.

[17] Martimort, David and Lars Stole. 2009. Selecting Equilibria in Common Agency Games. *Journal of Economic Theory*, 144(2): 604-634.

[18] Martimort, David and Lars Stole. 2012. Representing Equilibrium Aggregates in Aggregate Games with Applications to Common Agency. *Games and Economic Behavior*, 76: 753-772.

[19] Milstein, Arnold and Elizabeth Gilbertson. 2009. American Medical Home Runs. *Health Affairs*, 28(5): 1317-1326.

[20] Moses, Hamilton, David Matheson, Ray Dorsey, Benjamin George, David Sadoff, and Satoshi Yoshimura. 2013. *JAMA*, 310(18): 1947-1964.

[21] Rebitzer, James and Mark Votruba. 2011. Organizational Economics and Physician Practices. *Working Paper*.

[22] Robinson, James. 1999. *The Corporate Practice of Medicine: Competition and Innovation in Health Care*. University of California Press.

[23] Romano, Max, Jodi Segal, and Craig Pollack. 2015. The Association Between Continuity of Care and the Overuse of Medical Procedures. *JAMA Internal Medicine*, 175(7): 1148-1154.

[24] Ross, Joseph. 2002. The Committee on the Costs of Medical Care and the History of Health Insurance in the United States. *Einstein Journal of Biology and Medicine*, 19: 129-134.

[25] Sappington, David. 1983. Limited Liability Contracts between Principal and Agent. *Journal of Economic Theory*, 29: 1-21.

[26] Stigler, George. 1958. The Economies of Scale. *Journal of Law and Economics*, 1: 54-71.

[27] Tarski, Alfred. 1955. A Lattice-Theoretical Fixpoint Theorem and its Applications. *Pacific Journal of Mathematics*, 5: 285-309.

[28] Topkis, Donald. 1998. *Supermodularity and Complementarity*. Princeton University Press.