Sticking Points: Common-Agency Problems and Contracting in the U.S. Healthcare System

Brigham Frandsen   Michael Powell   James B. Rebitzer

May 11, 2018

Abstract

We propose a "common agency" model for explaining inefficient contracting in the U.S. healthcare system. We study the common-agency problems that arise when multiple payers seek to motivate a provider to invest in improved care coordination. We highlight the possibility of "sticking points," that is, Pareto-dominated equilibria in which payers coordinate around contracts that provide weak incentives to the provider. Sticking points rationalize three hard-to-explain features of the U.S. healthcare system: widespread fee-for-service arrangements outside of Medicare; problematic care coordination; and the historical reliance on single-specialty practices to deliver care. The model also analyzes the effects of policies aiming to promote more efficient contracting between payers and providers.

Keywords: Accountable Care Organizations, Common Agency, Moral Hazard

JEL classifications: D8, I10, I18

The authors would like to thank Dan Barron, Robert Gibbons, Jin Li, Thomas McGuire, Lars Stole, Jeroen Swinkels and Mike Whinston and for helpful comments along with participants at the MIT Organizational Economics Seminar; the NBER Personnel Economics Summer Institute meeting; the BU, Harvard, MIT Health Economics Seminar; The Maryland Workshop on Health Information Technology, The Society for Institutional and Organizational Economics Conference, and the European Health Economics Workshop in Lausanne Switzerland. We are, of course, responsible for any remaining errors. The views expressed herein are those of the authors and do not necessarily reflect the views of the National Bureau of Economic Research.

1 Introduction

The U.S. healthcare system is famously inefficient, but the causes are poorly understood (Baicker and Chandra, 2011). One candidate explanation that has received considerable attention from analysts and policy makers is inefficient contracting between payers and providers, in particular, the heavy reliance on fee-for-service pay arrangements with physicians. This paper proposes a "common agency" model for explaining the puzzling prevalence of inefficient contracting in the U.S. healthcare system.

Our analysis focuses on the common-agency problems that arise when multiple payers seek to motivate a provider to take actions that benefit the payers. The provider in this case acts as a "common agent" to the various payers. Strategic interactions among payers introduce two distortions that shape the equilibrium contracts offered to the provider. The first distortion, which has been extensively analyzed in other contexts but only rarely in healthcare, is the free-rider problem in which payers offer weak incentives to the provider because any given payer reaps only a fraction of the marginal benefit of stronger incentives.[1] The second distortion is a coordination failure between payers. Unlike the free-rider problem, coordination failures lead to the emergence of what we refer to as "sticking points," that is, Pareto-dominated equilibria in which all payers offer contracts that may entirely omit incentives for making efficient investments.

Previous analyses of common-agency problems have not emphasized the role of coordination failures among payers and the resulting sticking-point equilibria. A central contribution of this paper is to trace out the implications these equilibria have for the U.S. healthcare system and healthcare policy. In brief, we find that sticking-point equilibria offer a straightforward explanation for three long-observed but difficult-to to-explain features of the U.S. healthcare system: the ubiquity of fee-for-service contracting arrangements for physician services among private payers;[2] poor care coordination across providers;[3] and the

[1]For an important exception see Glazer and McGuire (2002), which is, to the best of our knowledge, the first application of common-agency models to the study of the U.S. healthcare system.

[2]Fee-for-service contracts reward physicians who are able to generate volume without consideration of the cost to payers. For this reason, the National Commission on Physician Payment Reform recommended a rapid transition away from fee-for-service payment because of its "inherent inefficiencies and problematic financial incentives." (The National Commission on Physician Payment Reform 2013, p. 3) Zuvekas and Cohen (2016) reports that 94.7% of physician office visits in 2013 were covered under fee-for-service arrangements (p. 412). Burns and Pauly (2018), in a review of physician compensation, similarly concludes that adoption of alternative payment mechanisms that link physician pay to quality or other performance metrics was low and had relatively little impact on physician incomes or how they deliver patient care.

[3]The Institute of Medicine's assessment of care quality in the U.S. healthcare system found that care

historic reliance on small, single-specialty practices rather than larger multi-specialty group practices.[4] The common-agency model we propose also provides insights on the effects of policies such as Accountable Care Organizations that aim to promote more efficient forms of contracting between payers and providers.

To illustrate the logic of our model, imagine two private payers who would like to encourage a common provider, for example, an independent physician practice, to implement an electronic medical record system. Implementing such a system requires the provider to exert effort, but many of the benefits of the effort spent introducing the electronic medical record system accrue to the payers, not the provider.[5] A natural way to motivate the provider to implement the new system would be for payers to move away from traditional fee-for-service contracts and offer the provider a share of the savings that the electronic medical records system generates for the payer. In a common-agency setting, however, the shared-savings incentives offered by one payer will also result in benefits to the patients covered by the other payer. Not surprisingly, this externality results in an equilibrium in which both payers offer

delivery is often complex and poorly coordinated, leading to wasted resources, gaps in coverage, loss of information, and reductions in the speed and safety with which care is delivered (IOM, 2001). Many of the problems of poor coordination result either from mishandled referrals to specialists or from fragmented care delivery. Both problems are exacerbated by relatively weak investments in technology and process improvements that strengthen integration across providers and organizations. For a review of the problems resulting from poorly handled referrals, see Mehrotra, Forrest, and Lin (2011). For a review and discussion of the problem of care fragmentation, see Cebul, Rebitzer, Taylor, and Votruba (2008) and Rebitzer and Votruba (2011). For estimates of the costs resulting from fragmented care delivery, see Frandsen et al. (2015), Hussey et al. (2014), Agha et al. (2016), and Romano, Segal, and Pollack (2015). For instances of investments in effective care coordination that deliver cost and quality benefits, see Milstein and Gilbertson (2009) and Simon et al. (2017)

[4]Historically, medical care in the U.S. was delivered by practitioners operating out of their own offices or as attendings in hospitals. This arrangement granted physicians a great deal of professional autonomy which, as a group, they were loathe to surrender to larger organizations (Starr, 1984; Robinson, 1999). Burns et al. (2013) documents that a large percentage of physician groups practice in small, physician-owned practices, although there is also a small but rapidly growing percentage of physicians operating in large groups that have been organized by non-physician owners. Baker, Bundorf, and Royalty (2014) estimates that as late as 1998, 29 percent of physicians worked in solo practices, and 55 percent in practices of 9 or fewer physicians. In contrast, only 19 percent of physicians were employed in practices having 50 or more physicians. Since the beginning of this century, physicians have slowly migrated towards larger practices so that, by 2010, 18 percent were solo practitioners, and only 40 percent worked in practices of 9 or fewer physicians. (Baker, Bundorf, and Royalty, 2014, Table 1). Despite this migration, 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. Burns and Pauly (2018), in a review of physician models of practice organization, concludes that overall the majority of physicians remain in solo or small group practices.

[5]For example, the electronic medical record system could allow the payer to track and discourage duplicative testing, treatments that are not cost-effective, excessive referrals to specialists or unwarranted emergency room visits.

weak incentives. As a result, the provider devotes low levels of effort and attention toward implementing the electronic medical records system.

If free riding were the only common-agency induced market failure at work, we would expect to see an equilibrium in which private payers offered weak incentives for improved care coordination. We would *not*, however, expect to find insurers and other private payers relying almost entirely on fee-for-service contracts that do not share any of the payer's gains from improved care coordination with providers. Yet this is the pattern historically observed in most of the U.S. healthcare system.

To explain the anomalous reliance on fee-for-service contracts within a common-agency framework, we consider coordination failures among payers. We introduce coordination failures into our previous example by adding the reasonable assumption that the transition to electronic medical records also requires providers to purchase a new health information technology (HIT) system. The fixed, up-front costs of purchasing such a system subtly alter the common-agency problem and make it more severe.[6] This is because the two payers will only deviate from traditional fee-for-service arrangements if they each believe that incentives jointly offered by *both* payers fully compensate the provider for these fixed up-front costs. If, in contrast, each payer believes that the other payer's contract includes weak incentives, then neither payer will find it optimal to shoulder the entire burden of motivating the provider to incur the fixed cost of the HIT system. In this case, the payers will stay with the traditional fee-for-service contracts—even if such contracts are Pareto-dominated higher-powered contracts each payer would offer if they believed the other would as well. To the extent that many organizational innovations that improve care coordination involve sizeable fixed costs, sticking-point equilibria within a common-agency model offer a plausible account for the persistence of both fee-for-service payments and poor care coordination.[7]

In a sticking-point equilibrium, providers also face weakened incentives to form integrated, multi-specialty group practices and so elect to deliver care through small, single-specialty practices. To see why, return again to the decision to invest in an HIT system. These systems enable superior coordination and information handoffs in referrals, but they typically do not

[6]See Simon et al. (2007) for evidence showing the most cited barrier to adopting HIT is up-front costs

[7]A partial list of organizational innovations that improve the integration of care across providers includes: investments in 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 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. Each of these innovations plausibly involves a combination of up front fixed investments and ongoing expenditures of effort.

allow for interoperability, that is, the easy exchange of information across organizations. In this setting, the gains from investing in HIT systems are greater when providers operate within a multi-specialty group practice, and the gains from forming multi-specialty group practices are similarly enhanced by investments in HIT systems. Put differently, because HIT investments and multi-specialty group practices are complementary, the failure to write incentive contracts that encourage efficient investments in HIT systems also depresses the returns to forming integrated multi-specialty group practices.[8]

Many others have observed that fee-for-service pay structures suppress investments in the technology and processes required for care coordination and integrated care delivery.[9] Our key contribution is to provide an explanation for the persistence of these pay practices. By rooting the persistence of inefficient fee-for-service contracts in a specific market failure, our common-agency model also provides insights on the effects of public policies that aim to promote more efficient forms of contracting between payers and providers.

The common-agency approach we develop in this paper has three policy implications that are not apparent in more familiar principal-agent models. First, in traditional principal-agent models, one might observe fee-for-service contracts, but such contracts would persist only when they are efficient. Thus in traditional principal-agent models—but not in our common-agency model when there are sticking-point equilibria—policies aimed at promoting shared saving incentives would be counterproductive in the sense that they would move the market away from efficient contracts. Secondly, common-agency problems become more severe as the number of payers increases: simply increasing the competitiveness of insurance markets may therefore not lead to more efficient contracting between insurers and providers. A subtle policy corollary, however, is that the efficiency of contracting under common agency improves when a larger share of a particular physician's patients are concentrated in a single insurer. These gains can be realized without an overall decrease in the competitiveness of insurance markets.

The third policy implication concerns the effect of government interventions on private

[8] Burns et al. (2013) and Burns and Pauly (2018) argue that the long persistence of small group practices is due to the limited scale and scope economies in physician practices. Our model highlights a different explanation: the complementarities between investments in care coordination and organizational form. If common-agency problems inhibit investments in care coordination, they will also inhibit the formation of multi-specialty group practices. Other sources of complementarity are also likely to be important because multi-specialty groups typically confine their referrals to specialists and hospitals who are also in the group. Simon et al. (2007) shows that electronic health record adoption is significantly more common in larger, more integrated practices.

[9] See for example Crosson (2009), Burns and Pauly (2002), and Blumenthal (2011).

contracting initiatives. Consider that under the Affordable Care Act, Medicare was empowered to write shared savings contracts with newly constituted Accountable Care Organizations (ACOs). Our common-agency model suggests two distinct mechanisms through which this policy can influence the contracts that private payers write with providers. If there were no coordination failure, so that contracts were shaped only by common-agency-induced free-riding among payers, Medicare's new incentive contracts will partially or fully crowd out already existing shared-savings contracts. Things are different if common agency also leads to a coordination failure among payers. In this case, introducing Medicare ACOs can "crowd-in" new, more efficient, private-sector contracts. Our model also suggests, however, that this jump-start effect may only manifest if Medicare's intervention is sufficiently aggressive.[10]

Taken individually, it is not hard to find other explanations for each of the stylized facts we highlight in this paper. Our common-agency analysis, however, can also account for features of healthcare contracting and delivery in the U.S. that are missing from these alternative stories. For example, the simplest explanation for the persistence of fee-for-service is that the benefits of the system (such as transparency and also the assurance that providers are fairly compensated for treating even the sickest patients) exceeds the cost of foregone incentives. This account is, however, hard to square with the history of Medicare's successful move away from fee-for-service payments for hospitals. Medicare adopted a prospective payment system late in 1983. Under this system, hospitals received a fixed payment for each Medicare patient in a given diagnosis related group (DRG), regardless of the actual expenses the hospital incurred in caring for the patient. Other public and private insurers subsequently implemented similar payment schemes, and this change had a substantial effect on the number of hospital admissions and length of stay.[11] As we have already noted and

[10]Weak ACO incentives, in contrast, may not have such a transformative effect. In this case, ACOs would not succeed in stimulating new investments in care coordination. Early evidence on the cost impacts of ACOs is consistent with this "weak incentives" view. The incentives in ACOs tend to be small and, for the vast majority of ACOs, involve no downside risk. On this basis, it is not surprising that early results suggest quite modest savings from ACOs (McWilliams, 2016). In addition, these savings do not appear to be generated from improvements in care coordination, although there is some evidence of a reduction in the use of low-value services (McWilliams et al., 2017). Less consistent with weak incentives is the early evidence of a "jump-start" effect—the majority of Medicare ACOs also contract with commercial payers (Peiris et al., 2016). It is not yet clear, however, if these contracts will persist if the effects of Medicare's intervention continue to be so modest.

[11]That private payers responded to Medicare's introduction of prospective pay systems is common knowledge (Dafny, 2005; Carter 1994). For a review of the literature on the effects of prospective payment on hospital admissions and length of stay see Cutler and Zeckhauser, 2000). See also the discussion in Baker, Bundorf, Devlin, and Kessler (2016).

will demonstrate more formally below, Medicare's apparent ability to jumpstart the introduction effective, high-powered incentive contracts is a natural consequence of the existence of sticking-point equilibria.

Similarly, the simplest alternative explanation for the long absence of large multi-specialty groups is that economies of scale and scope in the delivery of physician services are modest (Goldsmith et al., 2015; Burns and Pauly, 2018). But this explanation is hard to square with the observation that such groups have managed to thrive in some local markets but not others.[12] Market-level variation in the success of large medical groups, however, is a natural outcome of our common-agency approach in which such groups are inhibited in markets where sticking-point equilibria prevail.[13]

Our common-agency model also helps explain another market-level feature of the U.S. healthcare system: the curious failure of integrated insurance plans, like Kaiser Permanente, to spread into new markets. After all, these organizations would seem to have the advantage of internalizing the benefits of provider investments in more integrated care and, because the plan either owns or has exclusive contracts with physicians, there is no coordination failure among payers. The most straightforward explanation for the limited spread of vertically integrated payers is that they do not offer enough value to enough consumers to support their provider networks (Ho, 2009; Goldsmith et al. 2015), but this explanation is hard to reconcile with the success that insurers like Kaiser have had in their home regions.[14] Kaiser-like entities will create the most value when they enter a market that is in a sticking-point equilibrium. Entry requires contracting with a local set of providers. If the entity enters at small scale and covers only a small share of providers' patients, it will not be able to disrupt the sticking-point equilibrium and so will not be able to encourage investments in integrated

[12]For a qualitative discussion of the varying regional fortunes of multi-specialty groups, see Goldsmith and Burns (2016) and Burns and Pauly (2018). For a map of regional variation in the presence of large medical groups, see Welch et al. (2013).

[13]There are, of course, other potential explanations for market-level variations. It may be, for example, that locally dominant star hospitals such as Boston's Mass. General or Cleveland's Cleveland Clinic form large groups so as to lock in patient referrals and improve their bargaining power with private insurers. This is a reasonable story but if it were generally the driver of market variation, we would expect to see large groups emerge as part of vertically integrated delivery systems associated with locally dominant hospitals, and the actual appearance of large groups is much more diverse than this. The point of the model is not to eliminate alternative market stories but rather to highlight the potential importance of a previously overlooked possibility: coordination failures among payers.

[14]Kaiser enrollment exceeds 10 million, and more than 80% of that is in two states—Oregon and California—where Kaiser originated. The percentage of Kaiser enrollment outside of these states has been stable for a decade (Goldsmith and Burns, 2016). The national expansion program Kaiser undertook in the 1990s was an expensive failure (Goldsmith and Burns, 2016; Gitterman et al., 2003)

care, negating its competitive advantage. Things are different if the new entity is able to enter at large scale. In this case the sticking-point equilibrium could be disrupted, and the integrated plan could realize its competitive advantage, but convincing existing providers to give up all but their Kaiser patients is a formidable hurdle to clear.

We conclude this introduction by briefly situating our approach within the larger literature. Common-agency models were first introduced by Bernheim and Whinston (1986b). Much of the subsequent literature on common-agency models focuses on problems of lobbying and influence in political settings.[15] Ours is a complete-information moral-hazard model with risk-neutral parties and no negative transfers, so that agency costs arise from a trade-off between incentive provision and rent extraction (Sappington, 1983; Innes, 1990). Our set-up relies on public contracting variables, so that all payers' contracts depend on the same publicly observed variables.

Our model differs from the models commonly used in the literature in several ways. Imposing limits on transfers requires us to depart from the standard tools used to analyze the set of equilibria in common-agency models.[16] Our objective is also different. We seek to characterize the entire set of equilibrium action choices by the provider within a specific class of games. This differs from the more common approach of describing the distributional properties of a subset of equilibria in a general class of games. 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) In our model, whether there are multiple equilibrium actions depends on features of the provider's cost function, and we provide necessary and sufficient conditions for there to be multiple equilibrium actions.

The remainder of the paper proceeds as follows. In Section 2, we set up the model and describe necessary and sufficient conditions for equilibrium. In Section 3, we describe necessary and sufficient conditions under which our conclusions about sticking points hold. Section 4 considers the effects of policies (such as ACOs) aimed at promoting more efficient contracts. Section 5 considers the effects of common-agency problems on the formation of integrated, multi-specialty groups. Section 6 examines how alternative assumptions affect

[15]See for example Dixit, Grossman, and Helpman (1997), Besley and Coate (2001), and Kirchsteiger and Prat (2001). To our knowledge, the only other heathcare application is Glazer and McGuire (2002).

[16]In the terminology of Martimort and Stole (2012), our contracting space is not bijective, meaning that the payers cannot effectively "undo" the contracts put in place by others.

the model's results. Section 7 presents a set of testable hypotheses that derive from our model and relates them to stylized facts emerging from the empirical literature. Section 8 concludes by discussing the broader implications of our approach for other healthcare policy and management concerns.

2 Theoretical Analysis

In this section, we develop a simple model that highlights both the free-rider problem and coordination failures that emerge under common agency. The principals in our model are the payers who wish to motivate a common agent (the provider) to invest in improved care coordination. Because the costs of these investments accrue to the provider and many of the benefits accrue to payers, some sort of incentive contract is required. These contracts will typically not provide first-best incentives because incentive contracts entail some agency costs.

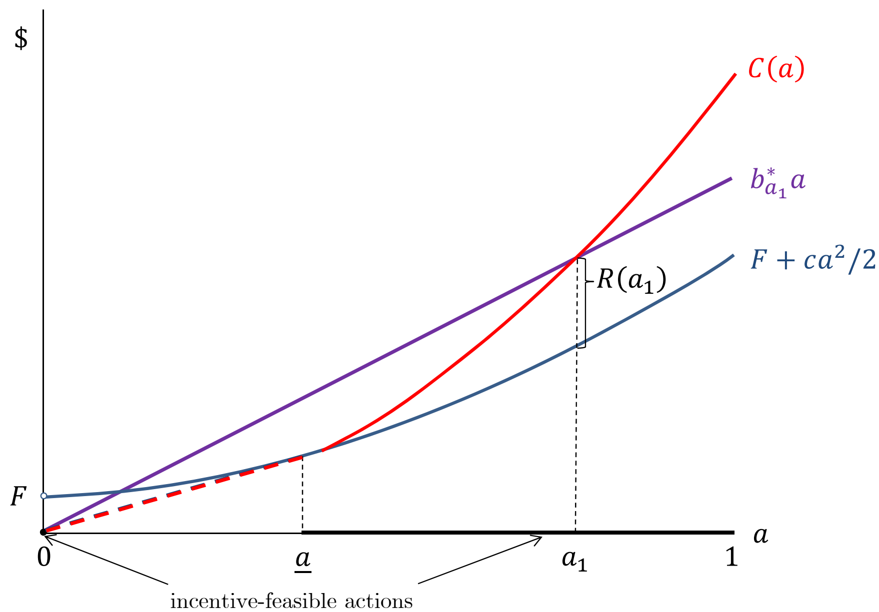
In our set-up, introduced formally in Section 2.1, payers cannot reduce up-front payments to capture all the rents that the contract generates. Higher powered incentives create more surplus, but a limited-liability restriction causes some of the surplus to flow to the provider as rent. Agency costs therefore emerge and result in a trade-off between incentive provision and rent extraction. In Section 2.2, we analyze equilibrium actions under our common-agency model. In comparison to the simpler and better known case of a single principal and agent, we find that the presence of multiple payers effectively amplifies agency costs. The magnitude of this amplification depends on what contracts payers expect other payers to offer.

To build intuition for our analysis, we use a running example in which payers seek to motivate a shared provider, for example, an independent physician practice, to invest in an electronic medical records system. This system reduces the cost of care delivery to payers by improving care coordination, but it imposes two costs on the provider. The first of these is the fixed cost of the HIT system. The second is the variable cost of the effort required to ensure successful implementation. Providers lack the knowledge to directly purchase or monitor implementation efforts. They therefore offer the provider a non-negative bonus if certain cost targets are met. This bonus is a vehicle for sharing the cost savings generated by the HIT system with the provider.

Figure 1 depicts the familiar incentive design problem that a unitary payer would face. The provider chooses the probability, a, that a binary contractible outcome is successful.

The provider's choice of investment level is on the horizontal axis. In this context, since outcomes are binary, any contract can be represented as a straight line from the origin. Given a contract, the provider chooses an investment level to maximize the difference between his expected rewards and his action costs. Consistent with our HIT system example, Figure 1 depicts both a fixed cost (the purchase of the system) and a variable cost representing the effort and attention required for successful implementation.

Figure 1: Unitary Payer Incentive Problem



This figure illustrates the provider's cost function (blue), the cost-minimizing contract for a particular investment level (purple) and its associated incentive rents, the set of incentive-feasible levels of invesment—levels for which there is a contract that could get the provider to choose that level of investment—and the unitary payer cost function (red) for what we will refer to as the health-IT (HIT) example.

Given the depicted contract, the provider would choose investment level a_1 at cost $c\,(a_1) = F + ca_1^2/2$ if $a_1 > 0$ and $c(0) = 0$, and he would receive expected rents $R\,(a_1)$ equal to the difference between his expected benefits and costs. The effective cost to the payer of getting the provider to choose investment level a_1, denoted $C\,(a_1)$, is therefore equal to the provider's costs, $c\,(a_1)$, plus the rents required to get him to choose a_1, $R\,(a_1)$. The need to give the provider rents is the source of agency costs that leads the payer to offer a contract that does not maximize total surplus.

When there are multiple payers, these agency costs are amplified by two additional sources of contracting frictions. The first additional source of friction is a free-rider problem among payers: at the margin, to get the provider to choose a higher level of investment, each payer has to effectively top off all the contracts she believes all others are putting in place, which means that she is effectively facing the entire marginal agency costs while receiving only a fraction of the marginal benefits. The second friction results from a coordination failure. If a payer believes no other payer will offer incentive contracts, then she has to shoulder the entire burden of getting the provider to choose any positive investment level, which can be substantial when the costs of the provider's actions are lumpy. However, if she believes the other payers are offering high-powered incentive contracts, she only has to shoulder a small part of the burden of getting the provider to choose a positive investment level. In this example, and as we demonstrate much more generally below, these coordination problems only emerge when there are fixed costs in addition to the standard variable costs.

2.1 The Model

There are N risk-neutral payers (denoted $P_1, \ldots, P_N$) and a single risk-neutral provider. There is a binary outcome, $y \in \{0, 1\}$, which can either be success at hitting a cost target, or failure, and the probability of success is determined by the provider's action choice a: $\Pr[y = 1 | a] = a \in \mathcal{A} \subseteq [0, 1]$, where $\mathcal{A}$ is a compact set. We refer to a as the provider's **care-coordination investment level**. A successful outcome yields a total benefit B to all the payers, and we assume this benefit is equally distributed among them, so that payer i receives benefit $B_i = B/N$ if $y = 1$ and 0 if $y = 0$. The action a is costly to the provider: choosing a costs $c(a)$, where c is lower semicontinuous and nonnegative.

Payers simultaneously and noncooperatively offer bonus contracts $b_i \geq 0$, which specify a nonnegative payment to be made from P_i to the provider if $y = 1$ and zero if $y = 0$. If $b_i = 0$, we will say that P_i offers a **fee-for-service contract**, and if $b_i > 0$, we will say that P_i offers a **shared savings contract**. The provider 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 provider, we can without loss of generality assume he accepts all contracts. As a result, the provider cares about, and is motivated by, the **aggregate contract** $b = b_1 + \cdots + b_N$. If the provider is indifferent among several action choices, we assume he chooses the highest action he is indifferent among.

The timing of the game is as follows. First, $P_1, \ldots, P_N$ simultaneously offer $b_i \geq 0$ to the provider. The provider then chooses an action $a \in \mathcal{A}$ at cost $c(a)$. The outcome $y \in \{0, 1\}$ is realized, and $b_i y$ is paid from P_i to the provider.

A **subgame-perfect equilibrium** of this game is a set of nonnegative contracts $b_1^*, \ldots, b_N^*$ and an action-choice function a^* such that: (1) given b_{-i}^* and the provider's action-choice function, P_i optimally offers b_i^*, and (2) given $b_1^*, \ldots, b_N^*$, the provider optimally chooses a^*. We will say that $b^* = b_1^* + \cdots + b_N^*$ is an **equilibrium aggregate contract**, 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 $c(\cdot)$. In particular, our general specification of the cost function will allow us to identify the precise properties of the cost function that are necessary for there to be multiple equilibrium actions.

2.2 Computing Equilibrium Actions

In order to compute the set of equilibrium actions, it will first be useful to solve the problem that a unitary payer would face if there were no other payers. The payer wants to choose an action a she wants the provider to undertake, and the cheapest contract b that gets him to take action a. Define the **unitary payer cost function** $C : \mathcal{A} \to \mathbb{R}_+$ to be the solution to the unitary payer's cost-minimization problem

$$C(a) = \min_{b \geq 0} ba$$

subject to the provider's incentive-compatibility constraint

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

Note that there may be some actions a for which no contract b could get the provider to choose a. We will refer to the actions that the payer could in principle get the provider to choose as **incentive-feasible actions**, and we will denote the set of such actions by $\mathcal{A}^{feas}$. Because the provider's preferences are additively separable in money and costs, we can always write the unitary payer's cost function as the sum of the provider's action costs and the agency costs, $C(a) = c(a) + R(a)$ for all $a \in \mathcal{A}^{feas}$, where $R(a)$ are the **incentive rents** required to get the provider to choose action a. For any action that is not incentive-feasible, the payer's objective function is $C(a) = +\infty$. We will say that a solution to this problem is a **cost-minimizing contract implementing action** a and denote the resulting contract by b_a^*. We show in Lemma 6 in the appendix that cost-minimizing contracts satisfy $b_a^* = \partial^- c(a)$, where $\partial^- c(a)$ is the smallest subgradient of c at a, and therefore $C(a) = b_a^* a$.

We begin by laying out two conditions that will be used in some of the results. The first condition allows us to generalize the first-order conditions to cases where the provider's cost function is not well behaved. We are specifically interested in cases where the provider's cost function entails fixed or "lumpy" costs, as this is precisely when common-agency problems generate coordination failures. Preliminary to stating this condition, we define the quantity

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

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

In our setup, to motivate the provider to take an action a, the payer has to give the provider incentive rents $R(a)$. In models of this sort with limited-liability constraints, inducing a higher action requires the principal to provide the agent with higher rents, implying that $R(a)$ is increasing. Condition CR further implies that the incentive rents schedule is not only increasing but is essentially a convex function.[17] Condition CR is implied by $c'''(a) \geq 0$, which is a standard condition invoked in moral-hazard models with limited liability and binary outcomes. We will say that c is **well-behaved** if it satisfies the following condition.

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

When the provider's cost function is well-behaved, there will be a unique equilibrium action. When the provider's cost function is not well-behaved, there may be multiple equilibrium actions.

Figure 1 illustrates, for the HIT example, the provider's cost function, the set of incentive-feasible investment levels (actions), the cost-minimizing contract for investment level a_1 and its associated incentive rents $R(a_1)$, and the unitary payer's cost function $C(a)$. Because the provider must cover his fixed costs in order to be willing to choose any positive level of investment, there will be some set of investment levels $(0, \underline{a})$ that he would not be willing to choose for any contract he faces. For higher values of a, the gap between $C(a)$ and $c(a)$ is increasing and convex, since the cost function in the HIT example satisfies Condition CR.

[17]More precisely, it ensures that $R(a)$ is convex-extensible on $[0, 1]$, where $R : \mathcal{A}^{feas} \to \mathbb{R}$ 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 (2017))

The following efficiency benchmarks will help us to interpret the common-agency equilibrium actions derived below. A **first-best action** a social planner would choose is any action satisfying

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

The first-best action is always incentive-feasible since it is an action the provider would be willing to choose if the aggregate bonus were $b = B$. Because actions are not directly contractible, the first-best action will typically not be an equilibrium action.

A **second-best action** is any action a unitary payer would implement or

$$a^{SB} \in \underset{a \in \mathcal{A}}{\operatorname{argmax}}\, Ba - C(a),$$

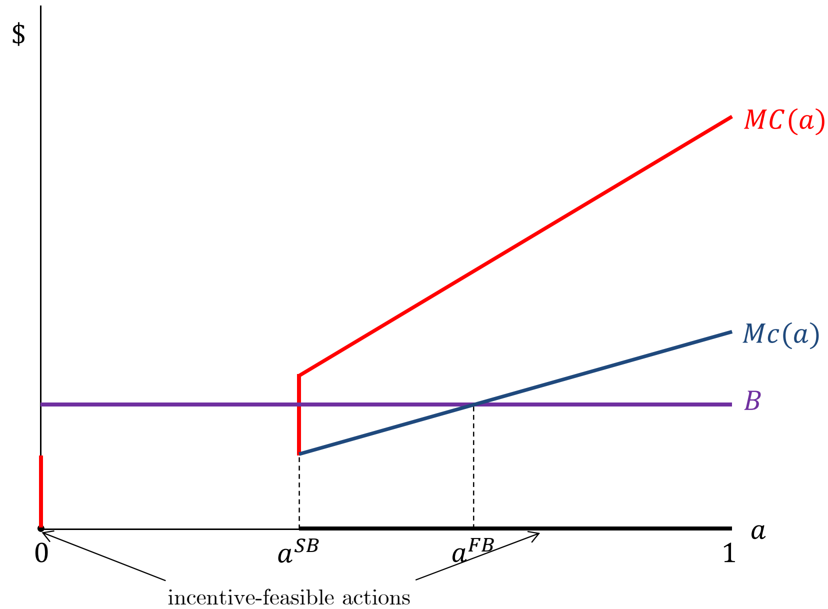
where $C(a)$ is the unitary payer's cost function defined above. The second-best action differs from the first-best because of agency costs. Function $c(a)$ includes only the provider's action costs, while $C(a)$ additionally includes the cost to the payer of rents earned by the provider. As we discuss below, actions under common agency (which may be termed "third-best") differ from the conventional principal-agent second-best action defined here because agency costs are typically strictly higher in common-agency settings.

We can characterize the second-best action using marginal conditions. Define $MC(a)$ to be the set of subgradients of C at a. If C is everywhere differentiable, then for all a, this set will be a singleton equal to $C'(a)$. We refer to $MC(a)$ as the **unitary payer's marginal-cost correspondence**. If Condition CR holds, the second-best action a^{SB} satisfies $B \in MC\left(a^{SB}\right)$. In general, the second-best action will be below the first-best action, because the incentive-rents schedule, $R(\cdot)$, is an increasing function. Moreover, we will show below that the second-best action in general represents an upper bound on equilibrium actions in the common-agency game.

Figure 2 below illustrates, for the HIT example, the unitary payer's marginal-cost correspondence $MC(a)$ as well as the provider's marginal-cost correspondence, which we denote by $Mc(a)$, and the payer's marginal benefit B. The marginal benefit B intersects $MC(a)$ at the second-best investment level, and it intersects $Mc(a)$ at the first-best investment level.

We are now in a position to describe the equilibrium conditions under common agency, that is, when there are $N \geq 2$ payers. We show in Corollary 1 in the appendix that when it comes to characterizing the set of equilibrium actions, $\mathcal{A}^*$, it is without loss of generality to focus on symmetric equilibria in which all payers offer the same contract to the provider. We will say that payer i **supports action** $\bar{a}$ if she offers the provider the contract $b_i = b^*_{\bar{a}}/N$. An

Figure 2: Marginal Conditions for Unitary Payer's Problem



This figure illustrates, for the HIT example, the provider's marginal-cost correspondence (blue), the unitary payer's marginal-cost correspondence (red), and the unitary payer's marginal benefit (purple). It also illustrates the first-best and second-best investment levels, a^{FB} and a^{SB}, respectively.

action a^* will therefore be an equilibrium action if and only if, whenever all payers other than payer i support action a^*, payer i also wants to support action a^*. We refer to the function $C_N(a, \bar{a}, N)$ as a payer's **effective cost function** given others support $\bar{a}$. Theorem 1 below provides necessary and sufficient conditions for an action a^* to be an equilibrium action.

THEOREM 1. *There is at least one equilibrium action, and there exists a function $C_N(a, \bar{a}, N)$ such that action a^* is an equilibrium action if and only if*

$$a^* \in \underset{a \in \mathcal{A}}{\operatorname{argmax}}\, Ba - C_N(a, a^*, N).$$

The function $C_N(a, \bar{a}, N)$ exhibits increasing differences in (a, N) and decreasing differences in $(a, \bar{a})$, and it is given by

$$C_N(a, \bar{a}, N) = \max\{NC(a), (N-1)\, b^*_{\bar{a}} a_{\min}(\bar{a})\} - (N-1)\, b^*_{\bar{a}} a,$$

where $a_{\min}(\bar{a})$ is the smallest action the payer can get the provider to choose if all other payers support action $\bar{a}$.

Proof of Theorem 1. See appendix.

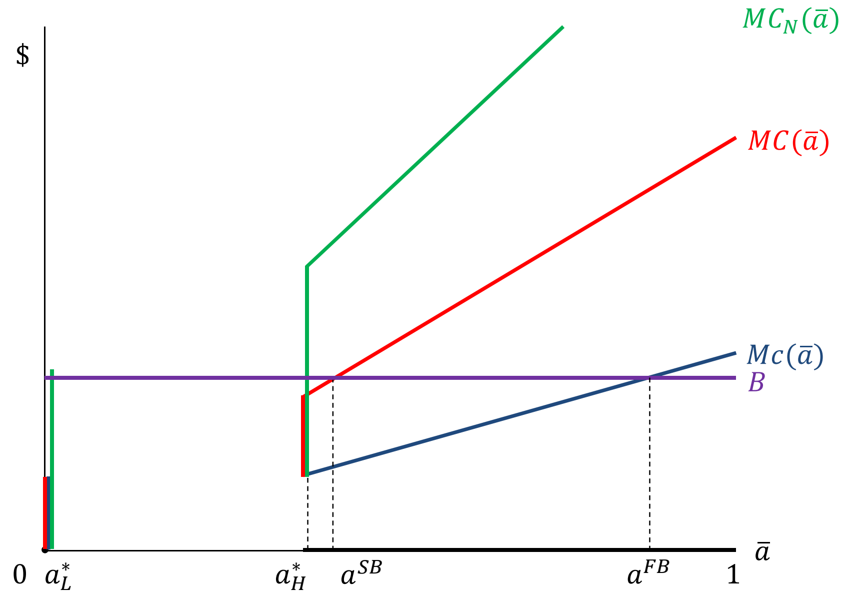
Following the general approach of Martimort and Stole (2012), Theorem 1 characterizes the common-agency game's set of equilibrium actions—which is a potentially complicated object involving the strategies of multiple principals—as the solution to a self-generating maximization problem. The intuition allowing this simplification is that a single payer can be thought of as choosing not only her own contract, b_i, but—since she takes the other payers' contracts as given—the aggregate contract b and therefore the action, a. In particular, the theorem shows that we can characterize the set of equilibrium actions by looking for the solutions to a problem of a unitary payer 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.

The particular form of $C_N(a, \bar{a}, N)$ has several elements and is worth commenting briefly on. First, if payer i believes all other payers will support action $\bar{a}$ and therefore each offer a contract $b^*_{\bar{a}}/N$, then she has effectively two options. She can offer a contract $b_i = 0$, in which case her costs will be zero, and the provider will choose action $a_{\min}(\bar{a})$. Alternatively, she can offer a contract that gets the provider to choose a higher action, in which case she will do so at lowest cost to herself. The lowest-cost contract she can offer to get the provider to choose action a is $b_i = b^*_a - (1 - 1/N)\, b^*_{\bar{a}}$, in which case her costs are $C(a) - (1 - 1/N)\, b^*_{\bar{a}} a$. Finally, to more easily compare the individual payer's maximization problem with that of a unitary payoff, we have multiplied her benefits and costs by N.

Characterizing the equilibria of the common-agency game in terms of solutions to an individual payer's decision allows us to draw an analogy with the unitary-payer setting and highlight the additional sources of inefficiency that arise from the common-agency problem. Analogous to the case with a unitary payer, we can define $MC_N(\bar{a})$ to be the set of subgradients of $C_N(a, \bar{a}, N)$ (with respect to a) evaluated at $\bar{a}$. We refer to $MC_N(\bar{a})$ as the **multiple-payer's marginal cost correspondence**. An action a^* is an equilibrium action if and only if the marginal cost correspondence $MC_N(a^*)$ contains the payers' aggregate marginal benefit, B.

Figure 3 illustrates, for the HIT example, the multiple-payer's marginal cost correspondence $MC_N(\bar{a})$ as well as the unitary payer's marginal cost correspondence and the provider's marginal cost correspondence. The payers' aggregate marginal benefit, B, intersects the $MC_N(\bar{a})$ curve twice, which implies that there are two equilibrium levels of investment: the low equilibrium level, a^*_L, and the high equilibrium level, a^*_H.

Figure 3: Marginal Conditions for the Common-Agency Problem



This figure illustrates, for the HIT example, the equilibria for three different games, which are each defined by the intersection of the marginal benefit curve (horizontal purple line B) and the respective marginal cost correspondences (blue, red, and green). The blue curve corresponds to the marginal costs faced by a social planner (i.e., the provider's marginal cost), and the intersection with B corresponds to the first-best action, a^{FB}. The red curve includes marginal costs and agency costs and corresponds to the marginal cost faced by a unitary payer. The intersection with B corresponds to the equilibrium of a conventional principal-agent game. The green curve shows the marginal cost faced by a payer in a multiple-payer setting, and the intersections with B correspond to equilibrium actions of the common-agency game. Action a_H^* is the third-best action under common agency, and a^{SB} is the second-best action under traditional agency with a single payer. Action a_L^* is the action without any shared savings incentives.

As Figure 3 illustrates, both equilibrium levels of investment are below the second-best, which in turn is below the first-best level of investment. The difference between the low and the high equilibrium levels of investment results from a coordination failure. If payer i believes all other payers are supporting zero investment, then if payer i wants the provider to choose a higher investment level, she has to shoulder the entire burden of getting him to do so. However, if she believes all other payers are supporting action a_H^*, then she only has to shoulder a small part of the burden of getting him to choose a_H^*. As a result, payers can get stuck in a vicious cycle in which none of them offers high-powered shared savings contracts because they think the others will not offer high-powered shared savings contracts. We now explore the implications of this result.

3 Equilibrium, Efficiency, and Coordination Failures

Common-agency problems amplify the agency costs that would arise in the interaction between a unitary payer and a provider. 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 action in the common-agency problem, a_H^*, is inefficient in that it results in actions that are no greater than second-best actions.

PROPOSITION 1. *The highest equilibrium action a_H^* is bounded from above by a^{SB}.*

Proof of Proposition 1. See appendix A.

In general, the inefficiency in equilibrium actions arises from three sources. The difference between the first-best action and the second-best action is the standard distortion that arises because of agency costs resulting from the trade-off between incentive provision and rent extraction in the unitary-payer problem. The second potential source of inefficiency is a free-rider problem among payers: at the margin, to get the provider to choose a higher action, each payer has to effectively top off all the contracts she believes all others are putting in place, which means that she is effectively facing the entire marginal cost while receiving only a fraction of the marginal benefit.

The third potential source of inefficiency is due to possible coordination failures, which arise when there are multiple equilibrium actions. Specifically, we will say $a_L^* = 0$ is a **sticking-point equilibrium** if $a_L^* = 0$ is an equilibrium action, and $a_H^* > 0$ is also an equilibrium action. In a sticking-point equilibrium, all payers offer fee-for-service contracts (i.e., $b_i^* = 0$ for all payers), and the provider does not undertake any care-coordination investment.

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, there is a unique equilibrium action. When Condition W is satisfied, $MC_N(\bar{a})$ 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 **nondifferentiability at** a if $\partial^- c(a) < \partial^+ c(a)$, where $\partial^+ c(a)$ is the largest subgradient of c at a. The next proposition provides necessary and sufficient conditions for there to be multiple equilibrium actions.

PROPOSITION 2. *Suppose Condition CR holds. If there are multiple equilibrium actions a_L^* and $a_H^* > a_L^*$, then there is a nondifferentiability. If there is a nondifferentiability at $\hat{a}$,* *then there exists a B for which $a_L^* = \hat{a}$ and $a_H^* > \hat{a}$. If Condition W holds, then there is a unique equilibrium action a^*.*

Proof of Proposition 2. See appendix A.

Proposition 2 implies that payers coordinating on an inefficient action when a more efficient equilibrium action exists can only occur when there is a nondifferentiability in the provider's cost function. In particular, this result implies that a sticking-point equilibrium can only arise if the provider's cost function is nondifferentiable at 0.

This nondifferentiability at 0 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 HIT example, when investments have a discrete component such as a fixed cost. In the healthcare context, innovations involving new care processes or information technologies appear likely to meet the nondifferentiability criterion. The criterion will manifestly not be satisfied, however, in the case most studied by prior common agency models—when the provider's cost function is well-behaved. Thus, under a conventional setup with well-behaved cost functions, we should not observe private payers relying strictly on fee-for-service compensation systems.

The next proposition shows that equilibrium actions are, in some sense, Pareto ranked.

PROPOSITION 3. *Suppose Condition CR holds. If there are multiple equilibrium actions, a_L^* and $a_H^* > a_L^*$, then (i) there exists an equilibrium with $a^* = a_H^*$ that Pareto dominates an equilibrium with $a^* = a_L^*$, and (ii) there does not exist an equilibrium with $a^* = a_L^*$ that Pareto dominates any equilibrium with $a^* = a_H^*$.*

Proof of Proposition 3. See appendix A.

For the first part of the proposition, note that symmetric equilibria are Pareto ranked, because each payer receives a share $(1/N)$ of the profits the unitary payer would receive, and the unitary-payer's profits are increasing in the provider's action (among his incentive-feasible actions) for all actions below a^{SB}. Further, since incentive rents $R(a)$ are increasing in a for $a \in \mathcal{A}^{feas}$, the provider's profits are also increasing in the action he is induced to take. For the second part of the claim, since $R(a)$ is increasing in a for $a \in \mathcal{A}^{feas}$, the provider is worse off for lower actions, so it cannot be that any equilibrium with $a^* = a_L^*$ Pareto dominates any equilibrium with $a^* = a_H^*$.

The possibility of an inefficient equilibrium with zero investment depends crucially on the common-agency concerns arising with multiple payers $(N > 1)$. In a setting with a unitary payer, non-differentiabilities in the provider's cost function at zero can certainly lead to an equilibrium with zero investment and zero incentives, but only when zero investment is efficient. To see this, note that if positive investment is efficient, there must be a level a' where average benefit Ba' exceeds the cost $c(a')$, which means there is some $b' < B$ for which $b'a'$ also exceeds $c(a')$. The provider will strictly prefer investing at the level a' to zero if given incentive b', and such a b' also makes the payer strictly better off than setting $b = 0$.

4 Common Agency and Public Policy to Improve Contracting

We have demonstrated that common-agency problems lead to third-best incentive contracts or, in the case of sticking-point equilibria, to outcomes that are Pareto dominated by the third-best outcomes. In this section we consider whether and how public policy interventions might be used to improve contracting. We will focus our discussion on a particular policy intervention that has gained a great deal of recent attention: Accountable Care Organizations (ACOs). Our findings, however, emerge from the fundamental logic of common-agency market failures and are not limited to this particular policy.

ACOs are entities composed of hospitals and/or other providers that contract with The Centers for Medicare and Medicaid 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.[18] As a public policy intervention, ACOs are essentially a commitment by Medicare to reduce reliance on fee-for-service and to engage in a new form of contracting with providers. By introducing shared savings contracts, ACOs seek to directly stimulate provider investments in more efficient, integrated care delivery. As the common-agency problem makes clear, the efficacy of these incentives also depends on the contracts offered by private sector payers. An important goal of ACOs, therefore, is to use Medicare's contracts to jump-start the introduction of similarly efficient shared savings contracts by

[18] ACOs can be formed by groups of tremendously varied size and integration, from integrated delivery systems such as Kaiser Permanente and Geisinger Health Systems, to loosely affiliated networks of providers. These latter ACOs typically lack a large, salaried multi-specialty group of physicians; they frequently lack a hospital as part of the entity; and may often have little experience in managing contracts that deviate from the fee-for-service norm. Allowing loosely affiliated networks to form ACOs greatly expands the potential reach of the policy. ACOs are also expected to develop similar contracts with private insurers, thereby spreading cost-effective, integrated care throughout the health care system.

private payers. In this way, ACOs offer the prospect of transforming incentives throughout the healthcare system. Whether ACOs can, in fact, play such a transformative role depends on the specifics of the market failures that inhibit efficient contracting. In the context of our common-agency model, we find that ACOs can either crowd-out or crowd-in efficient private sector contracting. As we detail below, the specific outcomes depend on the nature of the provider's cost function and the magnitude of the ACO intervention.

To model the effect of ACOs, we return to our main model with a single provider and multiple private payers. In this setting, ACO contracts with Medicare act as an additional **shared-savings payment**, S, that is chosen exogenously via public policy. This payment is common knowledge to all other payers. In this setting, the provider's payoff is $b+S$ if $y=1$ and 0 if $y=0$. To make clear that equilibrium actions depend on the ACO shared-savings contract, denote the least equilibrium action by $a_L^*(S)$ and the greatest equilibrium action by $a_H^*(S)$, and denote equilibrium aggregate contracts by $b^*(S)$.

As in the main model, there exists a function $C_N(a, \bar{a}, N, S)$ such that an action a^* is an equilibrium action if and only if

$$a^* \in \operatorname*{argmax}_{a \in \mathcal{A}} Ba - C_N(a, a^*, N, S),$$

and the function described in Theorem 1 is a special case of this function with $S=0$. Importantly, C_N satisfies decreasing differences in a and S: a higher-powered ACO contract decreases the payers' costs of getting the provider to undertake a higher action. The consequences of the ACO shared-savings payment S for equilibrium actions depends on what the equilibrium actions would be in the absence of an intervention. In settings in which there are no nondifferentiabilities in the provider's cost function, the ACO shared-savings contract has the perverse effect of reducing the incentives for payers to offer high-powered contracts, as the following proposition establishes.

PROPOSITION 4. *Suppose Condition W holds. Then for each S, there is a unique aggregate equilibrium contract $b^*(S)$, which is decreasing in S.*

Proof of Proposition 4. See appendix A.

When the provider's cost function is well-behaved, there is a unique equilibrium action. In this setting, the market failure responsible for inefficiently low-powered incentives for care-coordination investments is free riding among the multiple payers. Proposition 4 shows that ACO shared-savings payments will partially crowd out private shared savings incentive contracts in this case.

When coordination failures drive the inefficiency, however, even small ACO shared-savings contracts can *increase* private shared savings incentive contracts and improve social welfare. To make this point precise, let $W_L(S) = Ba_L^*(S) - c(a_L^*(S))$ denote aggregate surplus in the least equilibrium under ACO shared-savings payment S. We show that even small increases in S can substantially increase $W_L(S)$. Of course, this improvement comes at the cost of the expected subsidy expenditure $K(S) = S \cdot a_L^*(S)$. Since public funds may be costly to raise due to distortionary taxes or other considerations, whether or not the ACO shared-savings contract improves social welfare depends on whether its social return, defined as $(W_L(S) - W_L(0))/K(S)$, clears some hurdle rate κ, which corresponds to the cost of raising public funds. Proposition 5 shows that when there are coordination failures, ACO shared-savings contracts can increase the strength of private-payer contracts and increase social welfare at a return greater than any given hurdle rate κ.

PROPOSITION 5. *Suppose Condition CR holds, and there is a sticking-point equilibrium. Then there exists a B and a shared-savings payment $S > 0$ such that $b_L^*(S) > b_L^*(0)$. Additionally, for any value $\kappa > 0$, there exists a B for which the returns to an ACO intervention are greater than κ in the least equilibrium.*

Proof of Proposition 5. See appendix A.

In contrast to Proposition 4, Proposition 5 shows that when there are multiple equilibrium actions in the absence of an ACO shared-savings contract, an ACO shared-savings contract can in fact lead to an increase in the strength of the incentives in private-payer contracts. Thus, in a setting where prior to the introduction of ACOs, payers had been in a sticking-point equilibrium, the introduction of ACOs has the potential to jump-start incentive provision by private payers. Moreover, the social returns of such a jump-start can potentially be very large relative to the cost of the intervention itself. Note that Proposition 5 provides a *lower* bound on the potential social returns to an ACO policy when the initial (pre-ACO policy) equilibrium is at a sticking point, since the result bounds the social return corresponding to the lowest equilibrium; the return would be even higher if payers happened to coordinate on a different equilibrium following the introduction of the ACO contract.

The welfare-improving function of ACO contracts depends on two factors: nondifferentiabilities in the provider's cost function and common agency. As discussed above, nondifferentiabilities arise from fixed costs or lumpy investments of the sort often found in investments in organizational processes and technologies that aim to promote improved health care coordination and integrated care. The presence of fixed costs or other sources of nondifferentiability alone, however, is not sufficient for ACO-like subsidies to increase welfare. If there was only one payer, then ACO subsidies that jump-start investment above zero reduce welfare. This occurs because in the one-payer case with a nondifferentiability in the provider's cost function at zero, equilibrium investment is only zero if that is efficient, as noted above—a policy which increases investment above zero would in this case reduce welfare. Common-agency concerns are therefore an essential part of the theoretical rationale for ACO and other policy interventions aimed at improving contracting.

These two propositions highlight the different policy implications of the two different types of distortions introduced by the common-agency market failure. When free riding prevails, Proposition 4 establishes that policy interventions aimed at subsidizing improved contracting crowd out private-sector contracts. When coordination failures prevail as they do in sticking-point equilibria, Proposition 5 establishes that policies aimed at promoting improved contracting can "crowd in" new investments and generate a positive social return.

5 Common Agency and the Formation of Integrated, Multi-specialty Groups

This section of the paper considers the implications of the common-agency market failure for health care delivery organizations. Specifically, we argue that investments in care coordination are complements to the formation of multi-specialty integrated care delivery organizations. By discouraging the former, common-agency problems also discourage the latter. In this way, common agency helps support a fragmented system of care delivery. The analysis in this section is based on the firm-boundaries model of Hart and Holmström (2010) and Legros and Newman (2013) and emphasizes the trade off between coordination under integration and professional autonomy under non-integration.

We introduce the problem of integrated care delivery through our now-familiar example of an HIT system investment. Specifically, we imagine a setting in which there are two doctors in different specialties who operate independent practices. By virtue of their different specialties, each has a preference for a different type of HIT system, but the benefits of care coordination are greatest when both physicians invest in the same system. One way to ensure that the doctors each choose the same system is for each doctor to join a multi-specialty integrated practice and to give the decision about which IT system to purchase to the integrated practice. By agreeing to unified control of the investment decision, the doctors are trading off professional autonomy for enhanced integration and coordination. To the extent that professional autonomy is valuable to physicians while many of the gains from enhanced coordination accrue to payers, the payers will wish to promote integration by offering such practices shared savings incentives contracts. The payoffs from these incentives can then be used to compensate physicians for the loss of professional autonomy integration entails.

To make this argument more formally, suppose there are two doctors, A and B, who must make a pair of **horizontal coordination decisions** $d_1, d_2 \in \{0, 1\}$ and choose an action $a \in \mathcal{A} \subseteq [0, 1]$ at financial cost $c(a)$. The action and the horizontal decisions determine the probability that a public outcome $y \in \{0, 1\}$ is equal to 1, with $\Pr[y = 1 | a, d_1, d_2] = a(1 - |d_1 - d_2|)$. The public outcome, along with the aggregate bonus b offered by the payers determines the **monetary payoffs** the doctors receive, $\pi = by - c(a)$. Further, the doctors receive **private benefits** associated with the horizontal coordination decisions that are made. Doctor A receives $u_A = -d_1$ and therefore prefers $d_1 = 0$, and doctor B receives $u_B = -\alpha(1 - d_2)$ and therefore prefers $d_2 = 1$. The parameter α scales the relative value the two doctors place on professional autonomy, and we assume without loss of generality that $\alpha \leq 1$, so that doctor A incurs a larger cost if her preferred horizontal decision is not made. The horizontal coordination decisions and the private benefits are non-contractible, while the rights to make the horizontal coordination decisions, the right to choose the action, and the monetary payoffs are alienable and ex ante contractible. Neither the decisions nor the action is ex post contractible.

We will consider two governance structures, which we denote by $g \in \{I, NI\}$, where $g = I$ denotes provider integration into a multi-specialty group practice, and $g = NI$ denotes non-integration (or two single-specialty practices). Under integration, doctor A receives the monetary payoffs, makes both horizontal coordination decisions, and chooses the action. Under non-integration, doctor A receives the monetary payoffs, makes horizontal decision d_1, and chooses the action. Doctor B makes horizontal decision d_2.[19]

The timing of the game with provider organizational choice is as follows. First, $P_1, \ldots, P_N$

[19]Since there are four alienable items, there are sixteen possible governance structures (i.e., allocations of control, decisions, and monetary payoffs). We show in Lemma 10 in the appendix that if any governance structure is optimal, either integration or non-integration is optimal, so it is without loss of generality to focus on these two governance structures. We further show that if we allow for continuous revenue sharing in which doctor A receives monetary payoffs $\lambda\pi$, and doctor B receives monetary payoffs $(1 - \lambda)\pi$, then if any governance structure is optimal, there is an optimal governance structure in which doctor A receives all the monetary payoffs.

simultaneously offer $b_i \geq 0$ to the doctors. The doctors then bargain over a governance structure $g \in \{I, NI\}$. Whoever possesses control under g makes decisions and chooses the action $a \in \mathcal{A}$, and whoever possesses the monetary payoffs incurs the associated cost, $c(a)$. The outcome $y \in \{0, 1\}$ is realized, $b_i y$ is paid from P_i to whomever possesses the monetary payoffs, and private costs are realized. A subgame-perfect equilibrium of this game is a set of contracts, a governance structure choice, horizontal decisions, and an action such that each player is choosing optimally given others' choices. Define $V(b)$ to be the maximized *monetary* payoffs attainable by the two doctors given aggregate bonus b:

$$V(b) = \max_a ba - c(a).$$

Note that by the envelope theorem $V(b)$ is non-decreasing.

Under non-integration, doctor B will choose $d_2 = 1$, so doctor A's problem is:

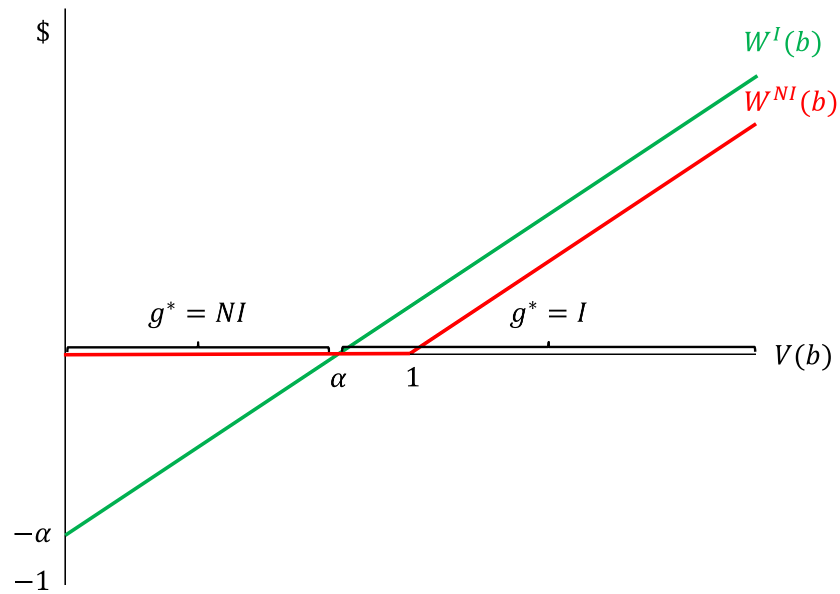
$$\max_{a, d_1} ba(1 - |d_1 - 1|) - d_1 - c(a).$$

Her problem is therefore to choose whether to minimize her private costs by choosing $d_1 = 0$, in which case she will also prefer to choose $a = 0$, or to coordinate with doctor B by choosing $d_1 = 1$, in which case she will choose a to maximize the monetary payoffs. She will opt for the former if b is small and for the latter if b is large. We will denote total surplus for the doctors under non-integration by $W^{NI}(b)$, and by Lemma 11 in the appendix, we have that $W^{NI}(b) = \max\{V(b) - 1, 0\}$.

Under integration, doctor A will choose $d_1 = d_2 = 1$, and she will choose a to maximize the monetary payoffs. Total surplus for the doctors under integration is denoted by $W^I(b)$, and by Lemma 11 in the appendix, we have that $W^I(b) = V(b) - \alpha$. Given aggregate bonus b, the optimal governance structure therefore solves $\max_{g \in \{I, NI\}} W^g(b)$. The solution to this problem is depicted in Figure 4 below. Provider incentives to form multi-specialty group practices arise from the incentive contracts providers have with their payers: as Figure 4 illustrates, when incentive contracts with payers are low-powered (i.e., b, and therefore $V(b)$ is small), providers will not find it optimal to forego the private benefits of professional autonomy. Moreover, integration increases the returns to coordinating horizontal decisions, and coordination of horizontal decisions complement care-coordination investments.

The complementarity between integrated organizations and the strength of shared-savings incentives also exacerbates the distortions resulting from common-agency problems. The reason for this is that the rents from selecting an integrated organizational form increases

Figure 4: Provider Integration Decision



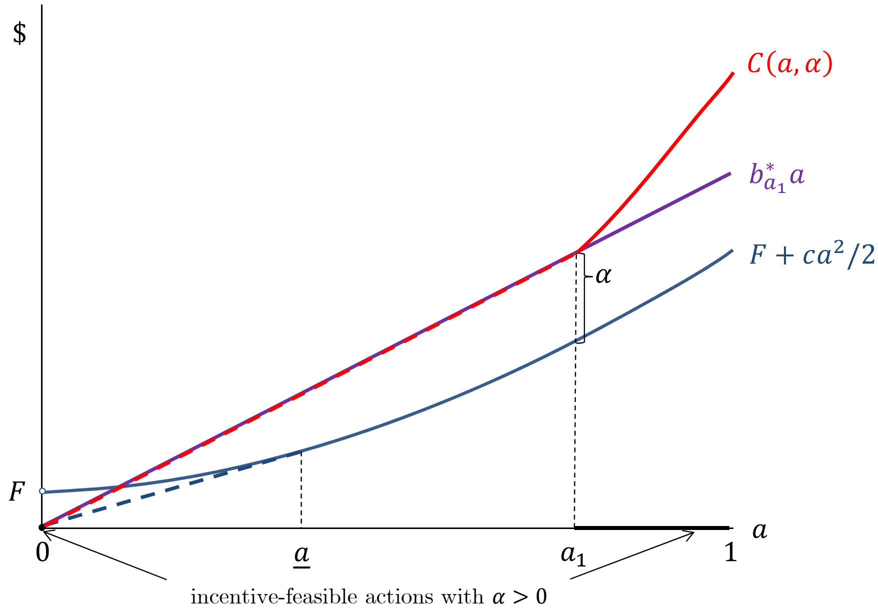
This figure illustrates the total benefits for the doctors under physician integration (green line) and under non-integration (red line) as a function of the maximized monetary payoffs attainable given an aggregate bonus level. To the left of α, these total benefits are higher under non-integration, and to the right of α, they are higher under integration.

the magnitude of the non-differentiability in the provider's cost function. We illustrate this discontinuity in Figure 5 in the simple case of a unitary payer, but the same logic holds in the common-agency setting. For any $\alpha > 0$, the rents providers earn under an integrated care delivery organization must be at least α. This result implies that the set of incentive-feasible actions is smaller than it is in the main model (which in fact corresponds $\alpha = 0$). This lacuna increases the scope for sticking-point equilibria.

We introduced our analysis of integrated multi-specialty groups through the example of investments in HIT systems, but the formal model indicates that complementarity between integrated care and investments in care coordination is far more general. If, for example, a PCP and a specialist develop procedures for jointly tracking and treating their shared patients, the returns to these investments will be higher within multi-specialty groups because the number of within-firm referrals and shared patients will be higher than if the PCP and specialist were not working in the same organization.[20]

[20]Stark laws and anti-kickback laws prohibit contractual arrangements ensuring that PCPs refer repeatedly to a particular set of specialists. Within-firm referrals, however, can be supported by profit-sharing arrangements that are allowed under the law: for example, a simple per capita division of profits is allowed

Figure 5: Unitary Payer's Incentive Problem with Provider Integration Decision



This figure illustrates the unitary payer's cost function when $\alpha > 0$. Private costs of integration reduce the set of incentive-feasible actions, since if the incentive rents from a contract are smaller than α, the providers will opt for non-integration and will not invest.

6 Model Extensions

Our discussion so far rests on a number of simplifying assumptions that streamline both the exposition and the formal analysis. Specifically, in the core common-agency model we give payers all the bargaining power, and we assume that the payers are symmetric. In addition, our analysis of multi-specialty groups assumes that payers have the initiative and that providers decide whether or not to integrate in response to payer incentive contracts. In this section we explore what happens in less restrictive environments. In each case our conclusions regarding the importance of sticking-point equilibria hold, but we also present some new results with additional economic implications.

Provider bargaining power: The baseline model in Section 2.1 assumes that the payers make contract offers to the provider, who can accept any subset of them. The ability

within multi-specialty groups. Since specialist visits are typically more profitable than primary care visits, profit sharing would give PCPs incentives to refer patients to specialists within their firm. These financial incentives may become diluted in large groups, but this dilution is offset by mutual monitoring and peer pressure that is reinforced by the colocation of specialists and PCPs (Kandel and Lazear, 1992).

to make offers gives payers something like "bargaining power," even though they do not receive all the social surplus from the contract. Indeed in models like ours, the split of the surplus between payers and providers is not a good indicator of relative bargaining power. This is because the limited liability constraint makes the division of the surplus between providers and payers an endogenous result of equilibrium contracts.

A natural way to introduce provider bargaining power is to assume that the provider has an outside option yielding profits $\bar{u} \geq 0$ that he receives if he rejects all the contracts the payers offer him. Intuitively, the more favorable the outside option, the more powerful the provider becomes because he can be "choosier" about accepting offers. Appendix B extends the model in this way. The results imply that increasing the provider's bargaining power actually *increases* the relevance of sticking-point equilibria. Specifically, we find that even in settings where there is no sticking-point equilibrium when $\bar{u} = 0$, increasing the provider's outside option can create a sticking-point equilibrium. Increasing the provider's outside option can also increase the level of investment that takes place in non-sticking-point equilibria. In short, as provider outside options improve, payers will either "go big or go home." "Going big" reflects a non-sticking point equilibrium with higher powered incentives and higher levels of provider investment. "Going home" means being in a sticking-point equilibrium in which incentive contracts to do not emerge in equilibrium.

Payer asymmetries: Throughout the analysis, we maintained the assumption that payers are symmetric, each responsible for the same number of the provider's patients and each receiving B/N if $y = 1$. In reality, payers will differ in the share of a provider's patients they cover. Often in a given market there will be a dominant payer covering the majority of a provider's patients and several smaller payers. In appendix B we extend the model to allow for such asymmetries. Three key implications follow, all of which suggest that asymmetries among payers can lead to more efficient equilibrium actions. First, asymmetries do not create sticking-point equilibria: if there is no sticking-point equilibrium when payers are symmetric, there is no sticking-point equilibrium when payers are asymmetric. Second, asymmetries among payers can eliminate a sticking-point equilibrium: even if there is a sticking-point equilibrium when payers are symmetric, this is not the case when they are sufficiently asymmetric. Finally, introducing asymmetries can increase the highest equilibrium action. These results have an appealing intuition if one recognizes that in the most extreme form of asymmetry with only one payer, the coordination and free-riding distortions resulting from common agency vanish.

Provider-driven change: Our analysis of the formation of multi-specialty groups in Section 5 gave the initiative to payers. Specifically our set-up presumed that payers first offered contracts and then providers decided whether to integrate or not. A natural question is whether similar results hold when the timing and initiative is reversed, that is, when providers make the integration decision, and payers respond by offering incentive contracts. Similar to our original results in Section 5, we find that under this alternative timing, common-agency problems among payers can interact with the integration decision among payers to depress both investment in care coordination and provider integration. Specifically, as we show formally in appendix B, the looming possibility of a sticking-point equilibrium can make non-integration strictly optimal for providers when it is the providers who move first.

7 Empirical Extensions

We have argued that our model of common agency and shared-savings contracts can account for a number of otherwise hard to explain features of the U.S. healthcare system. In this section we discuss a number of additional empirical implications that can, in principle, be used to assess the model's empirical relevance. As we discuss below, there is suggestive evidence consistent with these implications, but implementing rigorous empirical tests may be challenging.

The first empirical implication we discuss is that heightened competition between insurers may lead to less efficient contracting between insurers and payers. This results from the fact that both "free-riding" and "coordination failures" are exacerbated as the number of payers increases. More formally, this emerges from the result in Theorem 1 that $C_N(a, a^*, N)$ exhibits increasing differences in a, the level of provider investment, and N, the number of insurers. Thus as the number of insurers in a market increases, the marginal cost of increasing a increases. This implies that the lowest and highest equilibrium level of incentives are decreasing in N.

Observed correlations between payer market structure and efficient contracting are consistent with this prediction. Rosenthal et al.'s (2006) study of the adoption of pay-for-performance contracts among HMOs documents that payers (in their case, HMOs) with a higher market share are more likely to enter into pay-for-performance contracts with providers. This observed correlation however, is not sufficient to verify the prediction, since even conditional on their controls, it is likely that unobserved characteristics of HMOs—for example, quality of management—could be driving both market share and efficient contracting. The model's prediction is about a causal effect, not an observed correlation, and therefore testing it requires identifying an exogenous change in the concentration of payers in a market and examining effects on subsequent adoption of pay-for-performance contracts. A possible source of exogenous variation in payer concentration could be a merger of national-level payers whose impacts on local healthcare markets may be taken as exogenous.

The second empirical implication is that Medicare's introduction of shared-savings contracts will increase the overall incentives for physicians to invest in care coordination, and these heightened incentives will lead to more investment. If cost functions are well-behaved, we showed in Proposition 4 that Medicare's ACO incentives will only be partially offset by decreased incentives from private payers, so that the total incentive a provider faces is still increasing in the Medicare subsidy. If cost functions are not well-behaved, and the common agency game is at a sticking-point equilibrium, the appearance of ACOs can jump-start additional shared-savings contracts in the private sector.

As we discuss in the introduction, it is generally acknowledged that Medicare's introduction of its prospective payment system for hospitals had a jump-start effect on private payers' adoption of prospective payment. The evidence for ACOs having a similar effect is much weaker. For example, Hsiao and Hing (2014) documents dramatic increases in the adoption of electronic health records in the 2011-2013 period when ACOs were introduced. It is, of course, hard to distinguish the effect of ACOs from direct subsidies to these systems resulting from the HITECH Act. Consistent with the presence of a jump-start, some survey data suggest that many ACOs are forming contracts with private payers as well (Peiris et al., 2016). Identifying the causal effect of ACOs on private sector contracting, however, requires more than surveys. At a minimum analysts would have to find some exogenous cause of variation in ACO adoption. The rich empirical predictions of the model itself further complicate empirical tests—especially the result that if the market is at a sticking point, the introduction of ACOs with weak incentives should not be expected to disrupt the equilibrium or to transform payer incentives.

Finally, the complementarity between investments in care coordination and multi-specialty group practices suggests that markets with higher ACO penetration should experience an increase in these sorts of practices. This prediction follows from the results in Section 5. Suggestive evidence supporting the complementarity between investments in care coordination and integrated practices comes from Simon et al. (2007). The paper shows that measures of integration, namely practice size and hospital affiliation, significantly predict adoption of electronic health records. Incidentally they also find that among practices that have not adopted electronic health records, the most-cited barrier was prohibitive start-up costs. This finding is consistent with our assumption that up-front fixed costs play an important role in provider cost functions. While the correlation between integration and adoption of electronic health records is suggestive, it does not prove complementarity, as unobserved characteristics affecting the returns to both care coordination and integrated practices could be driving the correlation. Athey and Stern (1998) offers a method to demonstrate complementarity, in which one jointly estimates equations describing productivity and the adoption care coordination and practice integration using instruments that predict care coordination and integration but do not otherwise affect productivity, a challenge in a healthcare setting where productivity is difficult to measure.

8 Conclusion

In this paper we have proposed a common-agency model for explaining inefficient contracting in the U.S. healthcare system. Common-agency problems arise when multiple payers seek to motivate a shared provider to undertake investments that reduce payers' costs. Our approach differs from other common-agency models in that we analyze sticking points, that is, equilibria in which payers coordinate around Pareto-dominated contracts that do not offer providers incentives to implement efficient investments. These sticking points can account for three long-observed but hard to explain features of the U.S. healthcare system: the ubiquity of fee-for-service contracting arrangements outside of Medicare; problematic care coordination; and the historic reliance on small single-specialty practices rather than larger multi-specialty group practices to deliver care.

Market failures induced by common agency also have broader implications for healthcare policy and management. One implication concerns Medicare's promotion of Accountable Care Organizations. A move by Medicare to introduce more efficient contracting in the form of ACOs may eliminate inefficient sticking-point equilibria, but only if the ACO incentives are sufficiently strong. Weak incentives may fail to eliminate the inefficient equilibrium and so fail to materially change incentives to make investments in care coordination or complementary organizational or technological innovations.

A second set of implications involves anti-trust concerns. As we have discussed, common-agency problems become more severe as the number of payers increases so that increased competition need not lead to more efficient contracting. Contracting efficiency, however, may increase when a provider's patients become increasingly concentrated in a single payer. In this way, narrow networks among insurers can potentially realize efficiency gains without necessarily decreasing market competition between insurers.

A third set of broader implications for our common-agency model concerns the persistently high administrative costs in the U.S. healthcare system. Cutler and Ly (2011 p. 6), for example, estimates that 39 percent of the $1,589 per capita difference in healthcare spending between the U.S. and Canada is due to higher expenditures on administration. Very little is known about the drivers of high administrative expenditures in the U.S., but a likely contributor is the near universal reliance on a claims-based system of reimbursing providers. Under a claims system, every billable activity is submitted to a payer as a claim, not a bill. This means that rather than simply paying the bill at the previously negotiated price, the insurer evaluates each claim and rejects many of them. The providers can then resubmit a modified claim or challenge the rejection. The payer can then accept the modified claim or respond to the challenge. Eventually the claim is adjudicated through a specified process. The claims-based payment system is costly and slow, involving tens of billions in challenged revenues, slow payments, and a small army of administrators at insurance companies and providers to manage the process (Gottlieb et al., 2018). One can imagine moving to a more streamlined billing system in which providers simply submit bills and are reimbursed, as in conventional businesses. To ensure that providers submit bills with the appropriate and accurate information, payers could test a sample of submitted claims and pay a bonus to providers based on the rate of acceptable bills. If the collective bonuses from all the payers were large enough, such bonuses could induce the common agent (the provider) to absorb the fixed and variable costs entailed by introducing systems that ensured appropriate billing. If the savings from improved billing processes by a provider spill over among the various payers, however, these incentive payments could by stymied by the common-agency induced coordination problems we have analyzed here. Coordination failures among payers could thus prevent the move to a streamlined billing system just as we have argued they do for fee-for-service payments. Indeed, the coordination problem involved in the transition from a claims system to a billing system is likely more severe because the various providers would also have to jointly coordinate around standards for billing requirements and also for credentialing providers.

The possibility that high administrative costs can result from common agency illustrates that the market failure we identify has broad implications for the healthcare system.

However, there are also important features of healthcare markets that we do not take into account. For example, providers in our model are risk-neutral and so our contracting problem abstracts from risk aversion. We also do not consider the multi-task problems entailed by shared-savings incentives when quality is hard to measure. Finally, our model ignores the complications introduced by multiple providers. With multiple providers, for example, payers could gain leverage over providers by threatening to leave some of them out of the network. This enhanced bargaining power may act to reduce the severity of common-agency problems for contracting. A full-fledged multi-agent/multi-principal model is beyond the scope of this paper. We hope, however, that the results we find in this stylized context will stimulate more research into related market failures in a richer environment.

Our results also have implications for the applied theory literature on common agency. Most noteworthy is our finding that the outcome of common-agency market failures depend critically on the sort of actions incentive contracts seek to elicit. When principals wish to encourage more effort, attention, or similarly continuous actions, equilibria involve third-best incentive contracts. When agent actions involve fixed costs or lumpy investments, as is often the case when agents are asked to implement new technology and management systems, there may also be coordination failures. In this case, outcomes can be much worse than third-best. The implications of these coordination failures for management and for public policy have not been fully worked out, and this may be an important direction for future theoretical research.

Appendix A: Proofs of Theoretical Results

The first two subsections of this appendix develop the arguments to prove the results in Section 2.2, including equilibrium existence and our characterization of the set of equilibrium actions. The first subsection develops Theorem 1A, which characterizes the set of equilibrium actions $\mathcal{A}^*$ as the solution to a self-generating maximization program, as in Martimort and Stole (2012). In particular, we show that $a^* \in \mathcal{A}^*$ if and only if

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

for some function $\tilde{\Lambda}(a, a^*)$. The second subsection shows that the operator $\hat{a}(\cdot)$ is monotone, so it always has at least one fixed point—and therefore an equilibrium action exists. In the process of proving these two results, we establish Theorem 1. The remaining subsections establish the results in Sections 3, 4, and 5.

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. 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 provider's **optimal action given aggregate contract** b by $a(b)$. Recall our tie-breaking assumption on the provider's choice: if the provider 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:

$$\partial c(a) = \{b \geq 0 : ba - c(a) \geq ba' - c(a') \text{ for all } a' \in \mathcal{A}\}.$$

A **cost-minimizing contract for** a is denoted by b_a^*, and it solves

$$b_a^* \in \underset{b \geq 0}{\operatorname{argmin}} \{ba : b \in \partial c(a)\}.$$

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

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

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

$$b_{a,\bar{b}}^* \in \underset{b \geq (1-1/N)\bar{b}}{\operatorname{argmin}} \{ba : b \in \partial c(a)\}.$$

Finally, if P_j, $j \neq i$ each choose $b^*_{\bar{a}}/N$, then we define the **minimum action relative to** $\bar{a}$, denoted $a_{\min}(\bar{a})$, to be the action that the provider will choose if P_i chooses $b_i = 0$.

Our analysis in this subsection proceeds in four steps. We first show that b^* is an equilibrium aggregate contract if and only if

$$b^* \in \hat{b}(b^*) = \underset{b \geq (1-1/N)b^*}{\operatorname{argmax}} \frac{1}{N} Ba(b) - \left(b - \left(1 - \frac{1}{N}\right) b^*\right) a(b).$$

An implication of this step is that b^* is an equilibrium aggregate contract if and only if there is a symmetric equilibrium in which b^* is the resulting aggregate contract. We then show that if $b^* \in \hat{b}(b^*)$, then b^* is a cost-minimizing contract for some action $a \in \mathcal{A}^{feas}$, and given any aggregate contract $\bar{b}$, any $b \in \hat{b}(\bar{b})$ will be a cost-minimizing contract for some action relative to $\bar{b}$. Finally, we show that $b^* \in \hat{b}(b^*)$ if and only if $b^* = b^*_{a^*}$, where

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

for some function $C_N(a, a^*, N)$.

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 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 b_{-i}, P_i chooses b_i to solve

$$\max_{b_i \geq 0} \left(\frac{1}{N} B - b_i\right) a(b) = \max_{b_i \geq 0} u_i(b_i, b)$$

We can instead think of P_i as choosing $b = b_i + b_{-i}$. Then $b_i \geq 0$ if and only if $b \geq b_{-i}$. P_i's problem is therefore

$$\max_{b \geq b_{-i}} u_i(b - b_{-i}, b).$$

If b^* is an equilibrium aggregate contract, then there exists $b^*_1, \ldots, b^*_N$ such that $b^*_j \geq 0$, $\sum_{j=1}^N b^*_j = b^*$, and, for each i,

$$b^* \in \underset{b \geq b^* - b^*_i}{\operatorname{argmax}} u_i\left(b - \sum_{j \neq i} b^*_j, b\right).$$

Since b^* solves this program for each i, it is therefore feasible for each i, and it also solves these programs on average:

$$b^* \in \underset{b \geq b^* - \min_j b_j^*}{\operatorname{argmax}} \frac{1}{N} \sum_{i=1}^{N} u_i \left(b - \sum_{j \neq i} b_j^*, b \right).$$

Define the quantity

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

or

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

Therefore, if $b^* \geq 0$ is an aggregate equilibrium contract, then

$$b^* \in \underset{b \geq b^* - \min_j b_j^*}{\operatorname{argmax}} \Lambda \left(b, b^* \right)$$

for some $b_1^*, \ldots, b_N^*$ such that $\sum_{i=1}^{N} b_i^* = b^*$. This argument establishes the following Lemma.

LEMMA 1. *If b^* is an equilibrium aggregate contract, then for some $b_1^*, \ldots, b_N^* \geq 0$ such that $\sum_{i=1}^{N} b_i^* = b^*$,*

$$b^* \in \underset{b \geq b^* - \min_j b_j^*}{\operatorname{argmax}} \Lambda \left(b, b^* \right).$$

The next lemma shows that we can replace the set of feasible contracts in this maximization problem by $b \geq (1 - 1/N)\, b^*$.

LEMMA 2. *If* b^* is an equilibrium aggregate contract, then

$$b^* \in \underset{b \geq (1-1/N) b^*}{\operatorname{argmax}} \Lambda \left(b, b^* \right).$$

Proof of Lemma 2. Lemma 2 is not directly implied by Lemma 1, because the objective in Lemma 2 involves a larger domain. Nevertheless, in order to get a contradiction, suppose b^* is an equilibrium aggregate contract, and suppose there is some b' such that $(1 - 1/N)\, b^* \leq b' \leq b^* - \min_j b_j^*$ and $\Lambda \left(b', b^* \right) > \Lambda \left(b^*, b^* \right)$. Then there must be some P_k such that

$$\left(\frac{1}{N} B - b' \right) a \left(b' \right) + b_{-j}^* a \left(b' \right) > \left(\frac{1}{N} B - b^* \right) a \left(b^* \right) + b_{-j}^* a \left(b^* \right),$$

but since P_k was optimizing, for P_k not to have chosen b', it must be the case that $b_j^* = 0$, and therefore,

$$b^* \in \underset{b \geq b^*}{\operatorname{argmax}} \left(\frac{1}{N}B - b\right) a(b) + b^* a(b). \tag{1}$$

Since $b_j^* = 0$, there must be some ℓ for which $b_{-\ell}^* \leq (1 - 1/N) b^*$ and for which

$$b^* \in \underset{b \geq b_{-\ell}^*}{\operatorname{argmax}} \left(\frac{1}{N}B - b\right) a(b) + b_{-\ell}^* a(b). \tag{2}$$

Since $a(b)$ is weakly increasing in b, (1) and (2) imply that

$$b^* \in \underset{b \geq b_{-\ell}^*}{\operatorname{argmax}} \Lambda(b, b^*),$$

and if b^* maximizes $\Lambda(b, b^*)$ over all $b \geq b_{-\ell}^*$, it also maximizes $\Lambda(b, b^*)$ over all $b \geq (1 - 1/N) b^*$, which contradicts the assumption that $\Lambda(b', b^*) > \Lambda(b^*, b^*)$.■

The following Lemma establishes the converse of Lemma 2.

LEMMA 3. $b^* \geq 0$ *is an equilibrium aggregate contract if and only if*

$$b^* \in \underset{b \geq (1-1/N)b^*}{\operatorname{argmax}} \Lambda(b, b^*).$$

Proof of Lemma 3. Necessity follows from Lemma 2. Now, suppose b^* solves this program. Let $b_i^* = \frac{1}{N} b^*$ for $i = 1, \ldots, N$. P_i's program is therefore

$$\max_{b \geq (1-1/N)b^*} \left(\frac{1}{N}B - \left(b - b_{-i}^*\right)\right) a(b) = \max_{b \geq (1-1/N)b^*} \left(\frac{1}{N}B - \left(b - \left(1 - \frac{1}{N}\right) b^*\right)\right) a(b),$$

which is the aggregate problem described in Lemma 2. Since b^* solves the aggregate problem, it therefore also solves each payer's problem.■

Lemma 3 completes the first step of the analysis. One immediate Corollary of Lemma 3 is that if b^* is an equilibrium aggregate contract, there is a symmetric equilibrium in which b^* is the associated equilibrium aggregate contract.

COROLLARY 1. *If* b^* *is an equilibrium aggregate contract, there is a symmetric equilibrium in which each Principal chooses* $b_i^* = \frac{1}{N} b^*$.

Step 2 We now turn to the second step, showing that any equilibrium aggregate contract must be a cost-minimizing contract for some action. This result is captured in Lemma 4.

LEMMA 4. *Suppose*

$$b^* \in \underset{b \geq (1-1/N)b^*}{\operatorname{argmax}} \Lambda(b, b^*).$$

Then $b^* = b_a^*$ *for some* $a \in A^{feas}$.

Proof of Lemma 4. Suppose

$$b^* \in \underset{b \geq (1-1/N)b^*}{\operatorname{argmax}} \left(\frac{1}{N} B - b + \left(1 - \frac{1}{N}\right) b^* \right) a\left(b^*\right).$$

Then b^* implements some action $a^* = a(b^*)$. In order to get a contradiction, suppose there is some contract $\tilde{b} \geq 0$ that also implements a^* but $\tilde{b}a^* < b^* a^*$. First, note that if $a(b^*) = a\left(\tilde{b}\right) = a^*$, then for any $\lambda \in [0,1]$, $a\left((1-\lambda)\tilde{b} + \lambda b^*\right) = a$. 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 $\hat{b} \geq (1 - 1/N) b^*$, then $\hat{b}$ is feasible and does better than b^* in the aggregate program, so $b^* \notin \hat{b}(b^*)$. Next, suppose $\hat{b} < (1 - 1/N) b^*$. Then, the contract $(1 - 1/N) b^*$ implements the same action and does better in the aggregate program, so $b^* \notin \hat{b}(b^*)$.■

Lemma 4 effectively restricts the domain over which we have to search when looking for fixed points of the $\hat{b}(\cdot)$ operator. In particular, we only have to look for b_a^* such that $b_a^* \in \hat{b}(b_a^*)$.

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

LEMMA 5. *Suppose $b \in \hat{b}\left(\bar{b}\right)$. Then b is cost-minimizing for some a relative to $\bar{b}$.*

Proof of Lemma 5. To get a contradiction, suppose b is not cost-minimizing for any action relative to $\bar{b}$. Let $a = a(b)$. Since $b^*_{a,\bar{b}}$ is a cost-minimizing contract for a relative to $\bar{b}$, it is feasible, and we have $b^*_{a,\bar{b}} a < ba$, which implies that

$$\left(\frac{1}{N}B - b^*_{a,\bar{b}} + \left(1 - \frac{1}{N}\right)\bar{b}\right) a > \left(\frac{1}{N}B - b + \left(1 - \frac{1}{N}\right)\bar{b}\right) a,$$

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

The main implication of Lemma 5 is that in solving for $\hat{b}\left(\bar{b}\right)$, it is without loss of generality to consider cost-minimizing contracts relative to $\bar{b}$. That is,

$$\underset{b \geq (1-1/N)\bar{b}}{\operatorname{argmax}} \Lambda\left(b, \bar{b}\right) = \underset{b^*_{a,\bar{b}} \geq (1-1/N)\bar{b}}{\operatorname{argmax}} \Lambda\left(b, \bar{b}\right).$$

Lemma 4 restricts the domain over which we have to search when looking for fixed points of the $\hat{b}(\cdot)$ operator. Lemma 5 shows that, given a cost-minimizing contract $b^*_{\bar{a}}$, we can restrict attention to looking for cost-minimizing contracts relative to $b^*_{\bar{a}}$. Denote a cost-minimizing contract for action a relative to $b^*_{\bar{a}}$ by $b^*_{a,\bar{a}}$, and denote the set of feasible actions relative to $b^*_{\bar{a}}$ by $\mathcal{A}^{feas}_{\bar{a}}$. Without loss of generality, we can therefore restrict attention to a domain and a range that are each isomorophic to $\mathcal{A}^{feas}$.

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

$$\bar{C}_N(a, \bar{a}, N) = N b^*_{a,\bar{a}} a - (N-1) b^*_{\bar{a}} a.$$

Our main characterization theorem follows.

THEOREM 1A. a^* *is an equilibrium action if and only if*

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

Proof of Theorem 1A. Suppose a^* is an equilibrium action. Then $b^*_{a^*}$ is an equilibrium aggregate contract (Lemma 4), which in turn implies that $b^*_{a^*} \in \hat{b}(b^*_{a^*})$ (Lemma 1). Since all $b \in \hat{b}(b^*_{a^*})$ are cost-minimizing relative to $b^*_{a^*}$ (Lemma 5), $b^*_{a^*} \in \hat{b}(b^*_{a^*})$ implies that $a^* \in \hat{a}(a^*)$. Conversely, suppose $a^* \in \hat{a}(a^*)$. Then $b^*_{a^*}$ is the best cost-minimizing contract relative to $b^*_{a^*}$, which implies that $b^*_{a^*} \in \hat{b}(b^*_{a^*})$ (Lemma 5).■

Theorem 1A shows that instead of solving for fixed points of $\hat{b}(\cdot)$, 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 of this subsection proceeds in four steps. Recall that we have denoted $\partial^- c(a)$ and $\partial^+ c(a)$ to be the smallest and largest subgradients of c at a. By convention, we will denote $\partial^- c(0) = 0$.

First, we will show that for all $a \in \mathcal{A}^{feas}$, $b^*_a = \partial^- c(a)$. We will then show that

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

where

$$C_N(a, \bar{a}, N) = \max\{NC(a), (N-1) b^*_{\bar{a}} a_{\min}(\bar{a})\} - (N-1) b^*_{\bar{a}} a,$$

and we will establish that $C_N(a, \bar{a}, N)$ satisfies decreasing differences in $(a, \bar{a})$ on $\mathcal{A}^{feas}$. By Topkis's (1998) theorem, this result implies that $\hat{a}(\cdot)$ is increasing, so 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 6 establishes the first result, solving for the set of cost-minimizing contracts of the unitary-payer problem in our setting.

LEMMA 6. $b^*_a = \partial^- c(a)$.

Proof of Lemma 6. In order for $b \in \partial c(a)$, it has to be the case that $\partial^{-} c(a) \leq b \leq \partial^{+} c(a)$. Given any $\tilde{b}$ such that $a\left(\tilde{b}\right) = a$, then setting $\tilde{b} = \partial^{-} c(a)$ implements the same action at weakly lower cost, and therefore $b_a^* = \partial^{-} c(a)$.■

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

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

LEMMA 7. *For all $\bar{a} \in A^{feas}$, the solutions to maximization program defined in* (1), $\hat{a}(\bar{a})$, *coincide with*

$$\underset{a \in \mathcal{A}^{feas}}{\operatorname{argmax}}\, Ba - C_N(a, \bar{a}, N).$$

Proof of Lemma 7. In this setting, we have $\mathcal{A}_{\bar{a}}^{feas} = \mathcal{A}^{feas} \cap [a_{\min}(\bar{a}), 1]$, $\bar{C}_N(a, \bar{a}, N) = N b_{a,\bar{a}}^* a - (N-1)\, b_{\bar{a}}^* a$, 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 $a \in \mathcal{A}_{\bar{a}}^{feas}$, $\bar{C}_N(a, \bar{a}, N) = C_N(a, \bar{a}, N)$. Finally, for all $a < a_{\min}(\bar{a})$,

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

so that

$$\underset{a \in \mathcal{A}_{\bar{a}}^{feas}}{\operatorname{argmax}}\, Ba - C_N(a, \bar{a}, N) = \underset{a \in \mathcal{A}^{feas}}{\operatorname{argmax}}\, Ba - C_N(a, \bar{a}, N),$$

which completes the proof.■

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

LEMMA 8. $C_N(a, \bar{a}, N)$ *satisfies decreasing differences in* $(a, \bar{a})$ *and increasing differences in* (a, N) *on* A^{feas}. *Consequently,* $\tilde{\Lambda}(a, \bar{a})$ *satisfies increasing differences in* $(a, \bar{a})$ *and decreasing differences in* (a, N) *on* A^{feas}.

Proof of Lemma 8. Let $a \geq a'$ and $\bar{a} \geq \bar{a}'$ with $a, a', \bar{a}, \bar{a}' \in \mathcal{A}^{feas}$. Define the difference $\Delta(\bar{a}) \equiv C_N(a, \bar{a}, N) - C_N(a', \bar{a}, N)$ and the value $\delta = (b_{\bar{a}}^* - b_{\bar{a}'}^*)(a - a') \geq 0$. There are six cases we need to consider. They are tedious but straightforward.

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

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

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

$$\Delta(\bar{a}) - \Delta(\bar{a}') = C(a') - (1-1/N)\, b^*_{\bar{a}} a_{\min}(\bar{a}) - \delta \leq 0$$

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

$$\Delta(\bar{a}) - \Delta(\bar{a}') = (1-1/N)\, b^*_{\bar{a}'} a_{\min}(\bar{a}') - b^*_{\bar{a}} a_{\min}(\bar{a}) - \delta \leq 0$$

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

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

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

$$\Delta(\bar{a}) - \Delta(\bar{a}') = (1-1/N)\, b^*_{\bar{a}'} a_{\min}(\bar{a}') - C(a) - \delta \leq 0$$

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

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

Since $\tilde{\Lambda}(a, \bar{a}) = \frac{1}{N} Ba - C_N(a, \bar{a}, N)$, $\tilde{\Lambda}(a, \bar{a})$ satisfies increasing differences in $(a, \bar{a})$ on $\mathcal{A}^{feas}$. The argument that $C_N(a, \bar{a}, N)$ satisfies increasing differences in (a, N) is similar.■

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

LEMMA 9. $\hat{a}(\cdot)$ is increasing on $\mathcal{A}^{feas}$.

Proof of Lemma 9. Follows directly from Topkis's theorem.■

The intuition behind Lemma 9 is that, given any cost-minimizing target contract, $b^*_{\bar{a}}$, each payer P_i either wants to leave $(1-1/N)\, b^*_{\bar{a}}$ in place by contributing $b_i = 0$, or they want to top up $(1-1/N)\, b^*_{\bar{a}}$. If they choose to top it up, they will top it up to a cost-minimizing contract, which is feasible, because b^*_a is increasing in a.

Step 4 Our second theorem follows from Lemma 9.

THEOREM 2A. *The set of equilibrium actions A^* is nonempty and compact.*

Proof of Theorem 2A. By Lemma 9 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 $[0,1]$.■

Putting all these results together, we get Theorem 1.

THEOREM 1. *There is at least one equilibrium action, and there exists a function* $C_N(a, \bar{a}, N)$ *such that action* a^* *is an equilibrium action if and only if*

$$a^* \in \underset{a \in \mathcal{A}}{\operatorname{argmax}}\, Ba - C_N(a, a^*, N).$$

The function $C_N(a, \bar{a}, N)$ *exhibits increasing differences in* (a, N) *and decreasing differences in* $(a, \bar{a})$*, and it is given by*

$$C_N(a, \bar{a}, N) = \max\{NC(a), (N-1)\, b^*_{\bar{a}} a_{\min}(\bar{a})\} - (N-1)\, b^*_{\bar{a}} a,$$

where $a_{\min}(\bar{a})$ *is the smallest action the payer can get the provider to choose if all other payers support action* $\bar{a}$.

Equilibrium, Efficiency, and Coordination Failures

PROPOSITION 1. *The highest equilibrium action* a^*_H *is bounded from above by* a^{SB}.

Proof of Proposition 1. By Lemma 8, $C_N(a, \bar{a}, N)$ satisfies increasing differences in (a, N), which implies that for any $\bar{a}$, $\hat{a}(\bar{a}, N)$ is smaller for larger values of N. By Topkis's theorem, this in turn implies that the set of fixed points to $a^* \in \hat{a}(a^*, N)$ is decreasing in strong set order in N. Since a^{SB} is the unique solution to $a^* \in \hat{a}(a^*, 1)$, the result follows.■

PROPOSITION 2. *Suppose Condition CR holds. If there are multiple equilibrium actions* a^*_L *and* $a^*_H > a^*_L$*, then there is a nondifferentiability. If there is a nondifferentiability at* $\hat{a}$*, then there exists a* B *for which* $a^*_L = \hat{a}$ *and* $a^*_H > \hat{a}$*. If Condition W holds, then there is a unique equilibrium action* a^*.

Proof of Proposition 2. For the first part of the proposition, define the quantity $H(a) = b^*_a + a\partial^- b^*_a$ on $a \in \mathcal{A}^{feas}$, where $\partial^- b^*_a = \lim_{a' \uparrow a} (\partial^- c(a) - \partial^- c(a')) / (a - a')$. $H(a)$ is strictly increasing in a, because b^*_a is strictly increasing in a on $\mathcal{A}^{feas}$ and $\partial^- b^*_a$ is weakly increasing in a by Condition CR. For a^*_L and $a^*_H > a^*_L$ to be equilibrium actions, it has to be the case that

$$\begin{aligned} \partial^- C_N(a^*_L) &\leq B \leq \partial^+ C_N(a^*_L) \\ \partial^- C_N(a^*_H) &\leq B \leq \partial^+ C_N(a^*_H), \end{aligned}$$

which implies that $\partial^+ C_N(a^*_L) \geq \partial^- C_N(a^*_H)$. Define $\Delta(a^*_L) = \partial^+ C_N(a^*_L) - \partial^- C_N(a^*_L)$. Then this last inequality implies that

$$N\Delta(a^*_L) \geq H(a^*_H) - H(a^*_L) > 0,$$

where the strict inequality follows from the argument above that $H(a)$ is strictly increasing in a. The result that $\Delta(a^*_L) > 0$ means that C is not differentiable at a^*_L.

For the second part of the proposition, suppose $\Delta(\bar{a}) > 0$ for some $\bar{a}$. Set $B = \partial^+ C_N(\bar{a})$ and define $\bar{a}^+ = \lim_{a \downarrow \bar{a}, a \in \mathcal{A}^{feas}} a$ to be the smallest incentive-feasible action larger than $\bar{a}$.

Since

$$\begin{aligned}\partial^{+}C_{N}\left(\bar{a}\right) &= N\partial^{+}C\left(\bar{a}\right)-\left(N-1\right)b_{\bar{a}}^{*}=N\partial^{-}C\left(\bar{a}^{+}\right)-\left(N-1\right)b_{\bar{a}}^{*}\\ &> N\partial^{-}C\left(\bar{a}^{+}\right)-\left(N-1\right)b_{\bar{a}^{+}}^{*}=\partial^{-}C_{N}\left(\bar{a}^{+}\right),\end{aligned}$$

and since $\partial^{+}C_{N}\left(a\right)$ is increasing in a, this implies that $\partial^{-}C_{N}\left(\bar{a}^{+}\right)<B\leq\partial^{+}C_{N}\left(\bar{a}^{+}\right)$, and therefore $\bar{a}^{+}>\bar{a}$ is also an equilibrium action.

For the last part of the proposition, note that if Condition W holds, then $MC_{N}\left(\bar{a}\right)$ is a singleton and is equal to $c'\left(\bar{a}\right)+N\bar{a}c''\left(\bar{a}\right)$, which is strictly increasing. $B\in MC\left(a^{*}\right)$ therefore has a unique solution a^{*}.■

PROPOSITION 3. *Suppose Condition CR holds. If there are multiple equilibrium actions,* a_{L}^{*} *and* $a_{H}^{*}>a_{L}^{*}$*, then* (*i*) *there exists an equilibrium with* $a^{*}=a_{H}^{*}$ *that Pareto dominates an equilibrium with* $a^{*}=a_{L}^{*}$*, and* (*ii*) *there does not exist an equilibrium with* $a^{*}=a_{L}^{*}$ *that Pareto dominates any equilibrium with* $a^{*}=a_{H}^{*}$*.*

Proof of Proposition 3. The first part of this proposition follows because symmetric equilibria are Pareto rankable. In a symmetric equilibrium, each payer receives $1/N$ of the total surplus of all the payers. This total surplus, $Ba-C\left(a\right)$ is increasing and convex by Condition CR, so it is higher for a_{H}^{*} than for a_{L}^{*}, since both are smaller than a^{SB} by Proposition 1. The provider is also better off under a_{H}^{*} than under a_{L}^{*} because incentive rents, $R\left(a\right)$, are increasing in a by Condition CR. The second part of the proposition also follows from the observation that $R\left(a\right)$ is increasing in a when Condition CR is satisfied: the lower equilibrium action necessarily makes the provider worse off.■

Accountable Care Organizations

PROPOSITION 4. *Suppose Condition W holds. Then for each* S*, there is a unique aggregate equilibrium contract* $b^{*}\left(S\right)$*, which is decreasing in* S*.*

Proof of Proposition 4. If Condition W holds, then there is a unique equilibrium action $a^{*}\left(S\right)$, which satisfies $B+S=c'\left(a^{*}\left(S\right)\right)+Na^{*}\left(S\right)c''\left(a^{*}\left(S\right)\right)$. Moreover, we will also have that $b^{*}\left(S\right)+S=c'\left(a^{*}\left(S\right)\right)$. Implicitly differentiating both expressions, we have

$$\frac{db^{*}\left(S\right)}{dS}=-N\frac{c''\left(a^{*}\left(S\right)\right)+a^{*}\left(S\right)c'''\left(a^{*}\left(S\right)\right)}{\left(1+N\right)c''\left(a^{*}\left(S\right)\right)+Na^{*}\left(S\right)c'''\left(a^{*}\left(S\right)\right)}<0,$$

establishing that $b^{*}\left(S\right)$ is decreasing in S.■

PROPOSITION 5. *Suppose Condition CR holds, and there is a sticking-point equilibrium. Then there exists a* B *and an ACO intervention* $S>0$ *such that* $b_{L}^{*}\left(S\right)>b_{L}^{*}\left(0\right)$*. Additionally, for any value* $\kappa>0$*, there exists a* B *for which the returns to an ACO intervention are greater than* κ *in the least equilibrium.*

Proof of Proposition 5. For both parts of the proposition, let $B=\partial^{+}C_{N}\left(a_{L}^{*}\left(0\right)\right)$. Since $a_{L}^{*}\left(0\right)=0<a_{H}^{*}\left(0\right)$, we have that $b_{L}^{*}\left(0\right)=0$. Set $S=\varepsilon>0$ small. Then $a_{L}^{*}\left(\varepsilon\right)>0$ and

$b_L^*(\varepsilon) + \varepsilon = \partial^- c(a_L^*(\varepsilon))$, so that $b_L^*(\varepsilon) = \partial^- c(a_L^*(\varepsilon)) - \varepsilon > 0$ for sufficiently small ε, which establishes the first claim.

For the second claim, note that there is some $\delta > 0$ such that for all $\varepsilon > 0$, $a_L^*(\varepsilon) \geq a_L^*(0) + \delta$. It follows that there is some $\Delta > 0$ such that for all $\varepsilon > 0$, $W_L(\varepsilon) - W_L(0) \geq \Delta$. $a_L^*(\varepsilon)$ is weakly increasing in ε and is bounded from above by one. We therefore have that for $\varepsilon < \Delta/\kappa$,

$$\frac{W_L(\varepsilon) - W_L(0)}{\varepsilon a_L^*(\varepsilon)} \geq \frac{\Delta}{a_L^*(\varepsilon)} \frac{1}{\varepsilon} \geq \frac{\Delta}{\varepsilon} > \kappa.$$

This establishes the second claim.■

Provider Fragmentation

The first lemma in this section shows that integration and non-integration weakly dominate any other governance structure. The second lemma in this section shows that our restriction to discrete monetary-payoff sharing rules is without loss of generality.

In order to establish the first claim, we can define all sixteen governance structures by which items are controlled by doctor A. Denote these governance structures as $g_1 = (\pi, a, d_1, d_2)$, $g_2 = (\pi, a, d_1)$, $g_3 = (\pi, a, d_2)$, $g_4 = (\pi, a)$, $g_5 = (\pi, d_1, d_2)$, $g_6 = (\pi, d_1)$, $g_7 = (\pi, d_2)$, $g_8 = (\pi)$, $g_9 = (a, d_1, d_2)$, $g_{10} = (a, d_1)$, $g_{11} = (a, d_2)$, $g_{12} = (a)$, $g_{13} = (d_1, d_2)$, $g_{14} = (d_1)$, $g_{15} = (d_2)$, $g_{16} = \emptyset$. The set of possible governance structures is $G = \{g_1, \ldots, g_{16}\}$. Note that $g_1 = I$ and $g_2 = NI$.

LEMMA 10. *Given any aggregate bonus b, either $g^* = NI$ or $g^* = I$.*

Proof of Lemma 10. First, note that if doctor i possesses π, then total surplus is weakly higher if doctor i also possesses a. This is because if it is ever optimal for the doctors to choose $a > 0$, then the doctor possessing a will be willing to do so if she also possesses π. This step establishes that each of $g_5, \ldots, g_{12}$ is weakly dominated by one other governance structure. Next, note that if doctor A (B) possesses d_2 (d_1), then she should also possess d_1 (d_2). If, say, doctor A possesses d_2, then in any pure-strategy equilibrium, she will always choose $d_2 = d_1$. Any equilibrium choice involving $d_2 = d_1 = 0$ is also an equilibrium choice under the governance structure in which she possesses d_1, and any equilibrium choice involving $d_2 = d_1 = 0$ is also an equilibrium choice under the governance structure in which doctor B possesses both d_1 and d_2. This observation implies that g_3 and g_{14} are weakly dominated. Next, since $\alpha < 1$, governance structure g_4 is strictly dominated by g_1 and governance structure g_{16} is weakly dominated by g_2. Finally, governance structure g_{13} yields the same outcomes as g_1, and governance structure g_{15} yields the same outcomes as g_2. The details for all these claims are straightforward but tedious.■

LEMMA 11. *Given an aggregate contract b, $W^{NI}(b) = \max\{V(b) - 1, 0\}$ and $W^I = V(b) - \alpha$. There exists a $\hat{b}$, which may be 0 or ∞, such that for all $0 \leq b \leq \hat{b}$, non-integration is optimal, and for all $b \geq \hat{b}$, integration is optimal.*

Proof of Lemma 11. Given an aggregate contract b, under non-integration, doctor B will always choose $d_2 = 1$. Doctor A then solves

$$\max_{a,d_1} ba\left(1 - |d_1 - 1|\right) - d_1 - c\left(a\right).$$

If she chooses $d_1 = 1$, then the problem becomes

$$\max_{a} ba - c\left(a\right) - 1,$$

and by definition, the value of this problem is $V\left(b\right) - 1$. If she chooses $d_1 = 0$, then she receives 0. She will therefore coordinate with doctor B if $V\left(b\right) - 1 \geq 0$, and she will not otherwise. Doctor B receives 0 no matter what, so total welfare is given by $W^{NI}\left(b\right) = \max\left\{V\left(b\right) - 1, 0\right\}$.

Under integration, doctor A will choose $d_1 = d_2 = 0$ to minimize her private costs, while still coordinating. These choices yield a payoff of $-\alpha$ for doctor B. Further, doctor A will solve

$$\max_{a} ba - c\left(a\right),$$

and will therefore receive $V\left(b\right)$. Total welfare under integration is therefore $V\left(b\right) - \alpha$.

Since $\alpha < 1$, if $V\left(b\right) \leq \alpha$, $W^{I}\left(b\right) \leq 0 = W^{NI}\left(b\right)$. If $\alpha \leq V\left(b\right) \leq a$, $W^{I}\left(b\right) \geq 0 = W^{NI}\left(b\right)$, and if $V\left(b\right) \geq \alpha$, $W^{I}\left(b\right) \geq W^{NI}\left(b\right)$. By the envelope theorem, $V'\left(b\right) = a^{*}\left(b\right) \geq 0$, so $V\left(b\right)$ is increasing, which implies the last set of results.■

Now, suppose that in addition to choosing whether to unify control $\left(g = I\right)$ or divide control $\left(g = NI\right)$, parties can also decide on a sharing rule in which A receives $\lambda\pi$ in monetary payoffs, and B receives $\left(1 - \lambda\right)\pi$ in monetary payoffs. Lemma 12 shows that it is without loss of generality to consider governance structures in which $\lambda = 1$.

LEMMA 12. *Given any aggregate bonus b, $\lambda^{*} = 1$.*

Proof of Lemma 12. Suppose $\lambda \in (0, 1]$, so that doctor A's payoffs are

$$\lambda\left(ba\left(1 - |d_1 - d_2|\right) - c\left(a\right)\right) - d_1,$$

and doctor B's payoffs are

$$\left(1 - \lambda\right)\left(ba\left(1 - |d_1 - d_2|\right) - c\left(a\right)\right) - \alpha\left(1 - d_2\right).$$

We will show that for any governance structure with $\lambda < 1$, there is a governance structure with $\lambda = 1$ that yields weakly higher expected total surplus.

Suppose $g = I$. A will choose $d_1 = d_2 = 0$ and will choose a to maximize $\lambda\left(ba - c\left(a\right)\right)$, yielding expected total monetary payoffs of $V\left(b\right)$ (a fraction λ of which go to A, and the rest of which go to B), so joint surplus is $V\left(b\right) - \alpha$ for any λ, and hence $\lambda^{*} = 1$ is weakly optimal if $g = I$.

Suppose $g = NI$. In any equilibrium with $d_1 = d_2$, doctor A will choose a to maximize $\lambda\left(ba - c\left(a\right)\right)$, again yielding expected total monetary payoffs of $V\left(b\right)$. Under $g = NI$, there are three potential (pure) equilibrium configurations for (d_1, d_2): $d_1 = d_2 = 0$, $d_1 = d_2 = 1$, and $d_1 = 0, d_2 = 1$.

The first configuration, $d_1^* = d_2^* = 0$, is an equilibrium configuration if $V(b) \geq \alpha/(1-\lambda)$ (which ensures that B does not want to deviate to $d_2 = 1$), and it yields total surplus $V(b) - \alpha$. The second configuration, $d_1^* = d_2^* = 1$, is an equilibrium configuration if $V(b) \geq 1/\lambda$ (which ensures that A does not want to deviate to $d_1 = 0$), and it yields total surplus $V(b) - 1$. The third configuration, $d_1^* = 0, d_2^* = 1$, is an equilibrium configuration if $V(b) \leq 1/\lambda$, and it yields total surplus 0. The configuration $d_1 = 1, d_2 = 0$ is never an equilibrium configuration, since doctor B would always do strictly better by deviating to $d_2 = 1$.

For any b such that $V(b) \in [\alpha, 1]$, $g = NI$, $\lambda \in (0, 1]$ yields, for any equilibrium configuration (d_1^*, d_2^*), weakly less total surplus than does $g = I, \lambda = 1$. For any b such that $V(b) \leq \alpha$ or $V(b) \geq 1$, $g = NI$, $\lambda \in (0, 1)$ yields, for any equilibrium configuration (d_1^*, d_2^*), weakly less total surplus than does $g = NI, \lambda = 1$. Finally, if $\lambda = 0$, the same results as above hold, except that monetary payoffs of less than $V(b)$ are possible, since doctor A would be indifferent among different action choices. We therefore have that any governance structure in which $\lambda < 1$ is weakly dominated by a governance structure in which $\lambda^* = 1$.■

Appendix B: Extensions

Provider Bargaining Power

In this section, we extend the model to allow the provider to have an outside option that yields profits $\bar{u} \geq 0$, which he can exercise if he rejects all the contracts the payers offer him. Given aggregate contract b, if he accepts the contracts, his profits will be $V(b) = \max_a ba - c(a)$. He will therefore accept the contracts if b is sufficiently large. The possibility that he may reject payers' contracts if the aggregate contract is not large enough can give rise to sticking-point equilibria, as the following Proposition highlights.

PROPOSITION 6. *Suppose Condition CR is satisfied. Given $B > 0$, there is a sticking-point equilibrium if $V(B/N) \leq \bar{u} \leq V(B)$.*

Proof of Proposition 6. First, we will present some preliminary results that are useful for establishing this proposition. By the envelope theorem, we can write $V(b) = \int_0^b a(t)\,dt$, where $a(b) = \max \operatorname{argmax}_{a \in \mathcal{A}} ba - c(a)$. By Topkis's theorem, $a(b)$ is increasing in b, and $a(0) = 0$, so there exists some minimal $\hat{b}$, which may be infinite, for which $a\left(\hat{b}\right) = 0$ for all $b \leq \hat{b}$. We therefore have $V(b) = 0$ for all $b \leq \hat{b}$, and $V(b)$ is strictly increasing on $\left(\hat{b}, \infty\right)$. As long as $\hat{b}$ is finite, for any $\bar{u} > 0$, there exists a unique $\bar{b}(\bar{u})$ satisfying $V\left(\bar{b}(\bar{u})\right) = \bar{u}$, so that the provider will accept the aggregate contract b if and only if $b \geq \bar{b}(\bar{u})$. The function $\bar{b}(\bar{u})$ is strictly increasing in $\bar{u}$. Let $\bar{a}(\bar{u}) = a\left(\bar{b}(\bar{u})\right)$ denote the smallest positive action the provider will choose in any aggregate contract he is willing to accept.

With these preliminaries out of the way, we will first show that $a = 0$ is an equilibrium action if $V(B/N) \leq \bar{u}$. Notice that $a = 0$ is an equilibrium action if, when P_i believes $b_j = 0$

for all $j \neq i$, i would prefer choosing $b_i = \bar{b}(\bar{u})$ and inducing $\bar{a}(\bar{u})$. Note that conditional on inducing some $a > 0$, the payer may do even better inducing an action $a > \bar{a}(\bar{u})$, but by Condition CR, the marginal cost of inducing higher actions is increasing. We therefore have that $a = 0$ is an equilibrium action if $\left(B/N - \bar{b}(\bar{u})\right)\bar{a}(\bar{u}) \leq 0$, which is equivalent to $V(B/N) \leq \bar{u}$.

We will now show that if $\bar{u} \leq V(B)$, there is some equilibrium action $a \geq \bar{a}(\bar{u})$. Suppose P_i believes all other payers are supporting action $\bar{a}(\bar{u})$, that is, i believes $b_j = \bar{b}(\bar{u})/N$ for all $j \neq i$. Then i prefers to support action $\bar{a}(\bar{u})$ over $a = 0$ if $\left(B/N - \bar{b}(\bar{u})/N\right)\bar{a}(\bar{u}) \geq 0$, which is equivalent to $V(B) \geq \bar{u}$. If this is the case, then there is some equilibrium action $a \geq \bar{a}(\bar{u})$. To see why, define the operator

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

By the argument above, any $a \in \hat{a}(\bar{a}(\bar{u}), \bar{u})$ must exceed $\bar{a}(\bar{u})$, and this is a monotone operator. By Zhou's (1994) extension of Tarski's (1955) fixed-point theorem, it has a solution $a^* \geq \bar{a}(\bar{u})$, which is an equilibrium action by Theorem 1.

Putting these results together, we have that if $V(B/N) \leq \bar{u} \leq V(B)$, both 0 and some $a \geq \bar{a}(\bar{u}) > 0$ are equilibrium actions, so there is a sticking-point equilibrium.■

Proposition 6 shows how the provider's outside option can lead to a sticking-point equilibrium. It has a couple important implications. First, if there is no sticking-point equilibrium when $\bar{u} = 0$, then a stronger provider can lead to a sticking-point equilibrium. Second, the proof of the proposition shows that when $\bar{u} \leq V(B)$, there is always an equilibrium action $a \geq \bar{a}(\bar{u}) > 0$ for some increasing function $\bar{a}(\bar{u})$. One implication of this observation is that the set of *positive* equilibrium actions is weakly, and indeed may be strictly, increasing in $\bar{u}$. If the provider is powerful, there will be a tendency for payers to "go big or go home."

Asymmetric Payers

In the baseline model, payers are symmetric and receive benefits B/N if there is a success. In this section, we will allow for asymmetries in the distribution of benefits across payers. In particular, suppose payers $P_1, \ldots, P_N$ receive benefits $\lambda_i By$, for $\lambda_i \geq 0$ for all i, and $\sum_{i=1}^{N} \lambda_i = 1$. This section establishes three sets of results. First, introducing asymmetries among payers does not add sticking-point equilibria. Second, such asymmetries may eliminate sticking-point equilibria if the payers are sufficiently asymmetric. Finally, if the provider's cost function is well-behaved, introducing asymmetries can increase the equilibrium action.

The first proposition shows that if there is no sticking-point equilibrium when payers are symmetric, there will not be a sticking-point equilibrium when payers are asymmetric either.

PROPOSITION 7. *Suppose there is no sticking-point equilibrium when $\lambda_i = 1/N$ for all i. Then there is no sticking-point equilibrium for any $(\lambda_1, \ldots, \lambda_N)$ satisfying $\lambda_i \geq 0$ for all i, and $\sum_{i=1}^{N} \lambda_i = 1$.*

Proof of Proposition 7. Since there is no sticking-point equilibrium when $\lambda_i = 1/N$ for all i, for each payer i, there exists a $b_i^* > 0$ such that $\frac{1}{N} Ba(b_i^*) - b_i^* a(b_i^*) > 0$. In other words, each payer would want to deviate to support a positive action if they thought that each other payer was going to support action 0.

Now, suppose payer benefits are $\lambda_i By$ for $\lambda_i \geq 0$ for all i, and $\sum_{i=1}^N \lambda_i = 1$. Then we must have that, for some i, $\lambda_i \geq 1/N$, and therefore

$$\lambda_i Ba(b_i^*) - b_i^* a(b_i^*) \geq \frac{1}{N} Ba(b_i^*) - b_i^* a(b_i^*) > 0.$$

We therefore have that $a = 0$ cannot be an equilibrium action, since P_i would want to deviate.■

For the second result, let us specialize the class of payer asymmetries we are looking at. In particular, let P_1's benefit be $\left(\frac{1}{N} + \lambda\right) By$, and P_j's benefit, for $j = 2, \ldots, N$, be $\left(\frac{1}{N} - \frac{\lambda}{N-1}\right) By$. Then the following proposition shows that if there is a sticking-point equilibrium when payers are symmetric, introducing enough asymmetries in payer benefits eliminates such an equilibrium.

PROPOSITION 8. *Suppose, when $\lambda = 0$, there is a sticking-point equilibrium, and total surplus is strictly positive under the equilibrium action a_H^*. Then there exists $0 \leq \bar{\lambda} < 1-1/N$ such that for all $\lambda > \bar{\lambda}$, there is no sticking-point equilibrium.*

Proof of Proposition 8. Since equilibrium action a_H^* yields strictly positive total surplus,

$$Ba\left(b_{a_H^*}^*\right) - b_{a_H^*}^* a\left(b_{a_H^*}^*\right) > 0,$$

and since 0 is an equilibrium action under $\lambda = 0$,

$$\frac{1}{N} Ba\left(b_{a_H^*}^*\right) - b_{a_H^*}^* a\left(b_{a_H^*}^*\right) \leq 0.$$

If the second condition holds with equality, set $\bar{\lambda} = 0$. If the second condition holds with strict inequality, there exists $\kappa \in (0, 1)$ such that

$$\begin{aligned} 0 &= (1-\kappa) \frac{1}{N} Ba\left(b_{a_H^*}^*\right) + \kappa Ba\left(b_{a_H^*}^*\right) - b_{a_H^*}^* a\left(b_{a_H^*}^*\right) \\ &= \left(\frac{1}{N} + \kappa\left(1 - \frac{1}{N}\right)\right) Ba\left(b_{a_H^*}^*\right) - b_{a_H^*}^* a\left(b_{a_H^*}^*\right). \end{aligned}$$

In this case, set $\bar{\lambda} = \kappa(1 - 1/N) < 1 - 1/N$. For all $\lambda > \bar{\lambda}$, we have

$$\left(\frac{1}{N} + \lambda\right) Ba\left(b_{a_H^*}^*\right) - b_{a_H^*}^* a\left(b_{a_H^*}^*\right) > 0 = \left(\frac{1}{N} + \lambda\right) Ba(0) - 0a(0),$$

which implies that 0 is not an equilibrium action, and there is therefore no sticking-point equilibrium.■

Finally, when the provider's cost function is well-behaved, small asymmetries among the payers do not affect the set of equilibrium actions, but if payers are sufficiently asymmetric, an increase in asymmetry can increase the equilibrium action. Suppose again that P_1's benefit is $\left(\frac{1}{N}+\lambda\right)By$, and P_i's benefit, for $i=2,\ldots,N$, is $\left(\frac{1}{N}-\frac{\lambda}{N-1}\right)By$. We can parametrize this distribution of benefits by λ.

PROPOSITION 9. *Suppose Condition W is satisfied. Then there exists a unique equilibrium action a^*_λ given parameter λ. There exists $0<\bar{\lambda}<1-1/N$ such that for all $0\leq\lambda\leq\bar{\lambda}$, $a^*_\lambda=a^*_0$, and for all $\lambda>\bar{\lambda}$, a^*_λ is increasing in λ, and $a^*_{1-1/N}=a^{SB}$.*

Proof of Proposition 9. First, we will establish a preliminary result. Let $(\lambda_1,\ldots,\lambda_N)$, with $\lambda_1\geq\lambda_2\geq\cdots\geq\lambda_N>0$ and $\sum_{i=1}^N\lambda_i=1$ denote the distribution of benefits across payers. Suppose a subset of $\mathcal{I}$ payers chooses a strictly positive contract in equilibrium. For those who do so, they will choose $b_i^*>0$ to satisfy

$$\lambda_i Ba'(b^*)-a(b^*)-b_i^*a'(b^*)=0,$$

so that if we sum up over the I payers who choose $b_i^*>0$, we get

$$\left(\sum_{i\in\mathcal{I}}\lambda_i\right)B=c'(a(b^*))+Ia(b^*)c''(a(b^*)).$$

For all payers j who do not contribute, it has to be the case that

$$0\in\underset{b_i\geq 0}{\operatorname{argmax}}\,\lambda_i Ba(b^*+b_i)-b_ia(b^*+b_i)$$

so that $\lambda_i Ba'(b^*)-a(b^*)\leq 0$ or

$$\lambda_i<\frac{1}{B}a(b^*)c''(a(b^*)),$$

and for all i who do contribute, it has to be the case that

$$\lambda_i=\frac{1}{B}a(b^*)c''(a(b^*))+\frac{1}{B}b_i^*>\frac{1}{B}a(b^*)c''(a(b^*)).$$

So if I payers contribute, it is the I biggest payers who contribute, and if $\lambda_i=\lambda_j$, then $b_i^*=b_j^*$.

Next, let $\left(\frac{1}{N}+\lambda,\frac{1}{N}-\frac{\lambda}{N-1},\ldots,\frac{1}{N}-\frac{\lambda}{N-1}\right)$ be the distribution of benefits across payers. By the previous argument, it will either be the case that only P_1 will choose $b_1^*>0$ or all P_i will choose $b_i^*>0$. Suppose only P_1 chooses $b_1^*>0$. Then the equilibrium action satisfies

$$\left(\frac{1}{N}+\lambda\right)B=c'(\tilde{a}^*_\lambda)+\tilde{a}^*_\lambda c''(\tilde{a}^*_\lambda), \tag{1}$$

and indeed, $b_j^* = 0$ if

$$\left(\frac{1}{N} - \frac{\lambda}{N-1}\right) B < \tilde{a}_\lambda^* c''\left(\tilde{a}_\lambda^*\right). \tag{2}$$

First, notice that $\tilde{a}_\lambda^*$ is increasing in λ, since $c'(a) + ac''(a)$ is increasing in a, and the left-hand side of the equilibrium condition is increasing in λ. Define $\tilde{\lambda}$ to satisfy (2) with equality, so that

$$\left(\frac{1}{N} - \frac{\tilde{\lambda}}{N-1}\right) B = \tilde{a}_{\tilde{\lambda}}^* c''\left(\tilde{a}_{\tilde{\lambda}}^*\right) = \left(\frac{1}{N} + \tilde{\lambda}\right) B - c'\left(\tilde{a}_{\tilde{\lambda}}^*\right),$$

and therefore $c'\left(\tilde{a}_{\tilde{\lambda}}^*\right) = \frac{N}{N-1}\tilde{\lambda}B$. For all $\lambda > \tilde{\lambda}$, $a_\lambda^* = \tilde{a}_\lambda^*$, so a_λ^* is increasing in λ. Moreover, if $\lambda = 1 - 1/N$, by (1), $\tilde{a}_{1-1/N}^* = a^{SB}$.

Next, suppose all P_i choose $b_i^* > 0$. Then the equilibrium action satisfies

$$B = c'\left(\hat{a}_\lambda^*\right) + N\hat{a}_\lambda^* c''\left(\hat{a}_\lambda^*\right),$$

and indeed, all P_i will in fact choose $b_i^* > 0$ if

$$b_i^* = \left(\frac{1}{N} - \frac{\lambda}{N-1}\right) B - \hat{a}_\lambda^* c''\left(\hat{a}_\lambda^*\right) > 0. \tag{3}$$

Let $\hat{\lambda}$ satisfy (3) with equality. For all $\lambda \leq \hat{\lambda}$, $a_\lambda^* = \hat{a}_\lambda^*$, which is independent of λ, and therefore $a_\lambda^* = a_0^*$ for all $0 \leq \lambda \leq \hat{\lambda}$. Finally, note that

$$\left(\frac{1}{N} - \frac{\hat{\lambda}}{N-1}\right) B = \hat{a}_{\hat{\lambda}}^* c''\left(\hat{a}_{\hat{\lambda}}^*\right) = \frac{1}{N}B - \frac{1}{N}c'\left(\hat{a}_{\hat{\lambda}}^*\right)$$

so that $c'\left(\hat{a}_{\hat{\lambda}}^*\right) = \frac{N}{N-1}\hat{\lambda}B$. We therefore have $\hat{\lambda} = \tilde{\lambda}$. Let $\bar{\lambda}$ be equal to this value. The results of the proposition follow.■

Alternative Timing for Provider Integration

Under the timing of the model in Section 5, $P_1, \ldots, P_N$ first simultaneously offer contracts $b_i \geq 0$ to the doctors. The doctors then bargain over a governance structure $g \in \{I, NI\}$ to maximize their joint surplus, and then actions a^* and horizontal coordination decisions d_1^* and d_2^* are chosen. In this section, we consider an alternative timing of the model in which the doctors first bargain over a governance structure $g \in \{I, NI\}$, $P_1, \ldots, P_N$ then simultaneously offer $b_i \geq 0$ to the doctors, and then a^*, d_1^*, and d_2^* are chosen.

Under this timing, the doctors must form a conjecture $\hat{b}$ about the aggregate contract that $P_1, \ldots, P_N$ will offer them, depending on their governance structure g. Define $V\left(\hat{b}\right) = \max_{a \in \mathcal{A}} \hat{b}a - c(a)$. Given a conjecture $\hat{b}^I$ under $g = I$, the doctors' joint surplus will be

$W^I\left(\hat{b}^I\right) = V\left(\hat{b}^I\right) - \alpha$. This is because under $g = I$, doctor A will choose $d_1 = d_2 = 0$, and she will choose $a \in \text{argmax}_{a\in\mathcal{A}}\, \hat{b}^I a - c(a)$. Given a conjecture $\hat{b}^{NI}$ under $g = NI$, the doctors' joint surplus will be $W^{NI}\left(\hat{b}^{NI}\right) = \max\left\{V\left(\hat{b}^{NI}\right) - 1, 0\right\}$. This is because under $g = NI$, if $V\left(\hat{b}^{NI}\right) < 1$, we will have $d_1 = 0$ and $d_2 = 1$, and if $V\left(\hat{b}^{NI}\right) \geq 1$, we will have $d_1 = d_2 = 1$, and doctor A will choose $a \in \text{argmax}_{a\in\mathcal{A}}\, \hat{b}^{NI} a - c(a)$.

Given these preliminary results, we can now state and prove our main result: if there is a sticking-point equilibrium under provider integration, there is an equilibrium in which providers choose not to integrate, payers offer fee-for-service contracts, and the providers do not choose a positive action.

PROPOSITION 10. *Suppose there is a sticking-point equilibrium in the subgame following $g = I$. Then there exists a subgame-perfect Nash equilibrium in which $g^* = NI$, $b^* = 0$, and $a^* = 0$.*

Proof of Proposition 10. Suppose there is a sticking-point equilibrium in the subgame following $g = I$. Then $B \in MC_N(0)$, and $b^* = 0$ is an equilibrium aggregate contract in that subgame.

Next, define $C_N^\alpha(a, \bar{a}, N)$ as the multiple-payer analog of $C(a, \alpha)$ in Figure 5, and let $MC_N^\alpha(\bar{a})$ be the corresponding multiple-payer's marginal cost correspondence. Then $MC_N(0) \subseteq MC_N^\alpha(0)$ for all $\alpha \geq 0$. This is because, when $\alpha = 0$, $C_N^0(a, \bar{a}, N)$ is exactly equal to $C_N(a, \bar{a}, N)$, and when $\alpha > 0$, the minimum aggregate bonus necessary to get the doctors to choose any $a > 0$ is strictly higher than it is when $\alpha > 0$.

Under $g = NI$, $a^* = 0$ is an equilibrium action if and only if $B \subseteq MC_N^1(0)$. Since $B \subseteq MC_N(0) \subseteq MC_N^1(0)$, we therefore have that $a^* = 0$ is an equilibrium action under $g = NI$, and $b^* = 0$ is an equilibrium aggregate contract.

Finally, suppose doctors correctly conjecture $\hat{b}^I = \hat{b}^{NI} = 0$. Then, since $W^{NI}(0) > W^I(0)$, the doctors will choose $g^* = NI$. Given $g^* = NI$, the providers will in fact offer aggregate contract $b = 0$, and the doctors will choose $a = 0$. We have therefore shown that there is a SPNE in which $g^* = NI$, $b^* = 0$, and $a^* = 0$.$\blacksquare$

References

[1] Agha, Leila, Brigham Frandsen, and James B. Rebitzer. 2017. Causes and Consequences of Fragmented Care Delivery: Theory, Evidence, and Public Policy. NBER Working Paper No. 23078.

[2] Athey, Susan and Scott Stern. 1998. An Empirical Framework for Testing Theories About Complimentarity in Organizational Design. NBER Working Paper No. 6600.

[3] 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*.

[4] Baker, Laurence, Kate Bundorf, Aileen Devlin, and Daniel Kessler. 2016. Why Don't Commercial Health Plans Use Prospective Payment? NBER Working Paper 22709. *NBER Working Paper Series*, Cambridge, Mass.

[5] 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.

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

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

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

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

[10] Burns, Lawton Robert, Jeff Goldsmith and Aditi Sen. 2013. Horizontal and Vertical Integration of Physicians: A Tale of Two Tails. *Annual Review of Health Care Management: Revisiting the Evolution of Health Systems Organization, Advances in Health Care Management*, 15: 39-117.

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

[12] Burns, Lawton Robert and Mark V. Pauly. 2018. Transformation of the Health Care Industry: Curb Your Enthusiasm? *The Milbank Quarterly*, 96(1): 57-109.

[13] Carter, Grace M., Peter D. Jacobson, Gerald F. Kominski and Mark J. Perry. 1994. Use of Diagnosis-Related Groups by Non-Medicare Payers. *Health care financing review*, 16(2): 127-158.

[14] 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.

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

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

[17] Cutler, David and Dan Ly. 2011. The (Paper)Work of Medicine: Understanding International Medical Costs. *Journal of Economic Perspectives*, 25(2): 3-25.

[18] Cutler, David M. and Richard J. Zeckhauser. 2000. Chapter 11 - the Anatomy of Health Insurance. *Handbook of Health Economics*, 1: 563-643.

[19] Dafny, Leemore S. 2005. How Do Hospitals Respond to Price Changes? *American Economic Review*, 95(5): 1525-1547.

[20] 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.

[21] Gitterman, Daniel P., Bryan J. Weiner, Marisa E. Domino, Aaron N. McKethan, and Alain C. Enthoven. 2003. The Rise and Fall of a Kaiser Permanente Expansion Region. *Milbank Quarterly*, 81(4): 567-601.

[22] 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.

[23] Goldsmith, Jeff and Lawton Burns. 2016. Fail to Scale: Why Great Ideas in Healthcare Don't Thrive Everywhere. *TheHealthCareBlog*, Sept. 29, 2016, http://thehealthcareblog.com/blog/2016/09/29/fail-to-scale-why-great-ideas-in-healthcare-dont-thrive-everywhere/.

[24] Goldsmith, Jeff, Lawton R. Burns, Aditi Sen, and Trevor Goldsmith. 2015. Integrated Delivery Networks: In Search of Benefits and Market Effects. National Academy of Social Insurance Research Report.

[25] Gottlieb, Joshua D., Adam Hale Shapiro, and Abe Dunn. 2018. The Complexity Of Billing And Paying For Physician Care. *Health Affairs*. 37(4): 619-626.

[26] Hart, Oliver and Bengt Holmstrom. 2010. A Theory of Firm Scope. *Quarterly Journal of Economics*, 125(2): 483-513.

[27] Ho, Katherine. 2009. Barriers to Entry of a Vertically Integrated Health Insurer: An Analysis of Welfare and Entry Costs. *Journal of Economics and Management Strategy*, 18(2): 487-545.

[28] Hsiao, Chun-Ju and Esther Hing. 2014. Use and Characteristics of Electronic Health Record Systems Among Office-based Physician Practices: United States, 2001–2013. *NCHS Data Brief*, 143.

[29] 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.

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

[31] Institute Of Medicine Committee On Quality Of Health Care In America. 2001. *Crossing the Quality Chasm: A New Health Care System for the 20th Century*. Washington D.C: National Academy Press.

[32] Kandel, Eugene and Edward P. Lazear. 1992. Peer Pressure and Partnerships. *Journal of Political Economy*, 100(4): 801-817.

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

[34] Kiselman, Christer and Shiva Samieinia. 2017. Convexity of Marginal Functions in the Discrete Case. In *Analysis Meets Geometry. The Mikael Passare Memorial Volume*. Eds. Andersson, Boman, Kiselman, Kurasov, and Sigurdsson.

[35] Legros, Patrick and Andrew F. Newman. 2013. A Price Theory of Vertical and Lateral Integration. *Quarterly Journal of Economics*, 128(2): 725-770.

[36] McWilliams, J. Michael. 2016. Changes in Medicare Shared Savings Program Savings From 2013 to 2014. *Journal of the American Medical Association*. 316(16): 1711-1713.

[37] McWilliams, J. Michael, Michael E. Chernew, and Bruce E. Landon. 2017. Medicare ACO Program Savings Not Tied To Preventable Hospitalizations Or Concentrated Among High-Risk Patients. *Health Affairs*. 36(12): 2085-2093.

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

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

[40] Mehrotra, Ateev, Christopher B. Forrest, and Caroline Y. Lin. 2011. Dropping the Baton: Specialty Referrals in the United States. *The Milbank Quarterly*, 89(1): 39-68.

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

[42] Moses, Hamilton, David Matheson, Ray Dorsey, Benjamin George, David Sadoff, and Satoshi Yoshimura. 2013. The Anatomy of Health Care in the United States.*JAMA*, 310(18): 1947-1964.

[43] Peiris, David, Madeleine C. Phipps-Taylor, Courtney A. Stachowski, Lee-Sien Kao, Stephen M. Shortell, Valerie A. Lewis, Meredith B. Rosenthal, and Carrie H. Colla. ACOs Holding Commercial Contracts are Larger and More Efficient than Non-Commercial ACOs. *Health Affairs*, 35(10): 1849-1856.

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

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

[46] 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.

[47] Rosenthal, Meredith B., Bruce E. Landon, Sharon-Lise T. Normand, Richard G. Frank, and Arnold M. Epstein. 2006. Pay for Performance in Commercial HMOs. *New England Journal of Medicine*, 355: 1895-1902.

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

[49] Simon, Melora, Niteesh K. Choudhry, Jim Frankfort, David Margolius, Julia Murphy, Luis Paita, Thomas Wang, and Arnold Milstein. 2017. Exploring Attributes of High-Value Primary Care. *Annals of Family Medicine* 15:529-534.

[50] Simon, Steven R., Rainu Kaushal, Paul D. Cleary, Chelsea A. Jenter, Lynn A. Volk, Eric G. Poon, John E. Orav, Helen G. Lo, Deborah H. Williams, and David W. Bates. 2007. Correlates of electronic health record adoption in office practices: a statewide survey. *Journal of the American Medical Informatics Association*, 14(1): 110–117.

[51] Starr, Paul. 1984. *The Social Transformation of American Medicine*. Basic Books.

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

[53] The National Commission on Physician Payment Reform. 2013. Report of the National Commission on Physician Payment Reform.

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

[55] Welch, W. Pete, Allison Evans Cuellar, Sally C. Stearns, and Andrew B. Bindman. 2013. Proportion of Physicians in Large Group Practices Continued to Grow in 2009-11. *Health affairs*, 32(9): 1659-1666.

[56] Zhou, Lin. 1994. A Theorem on Open Coverings of a Simplex and Scarf's Core Existence Theorem through Brouwer's Fixed Point Theorem. *Economic Theory*, 4(3): 473-477.

[57] Zuvekas, Samuel and Joel Cohen. 2016. Fee-for-Service, while Much Maligned, Remains the Dominant Payment Method for Physician Visits. *Health Affairs*, 35(3): 411.