Designing Contracts for Multitasking Groups: A Structural Model of Accountable Care Organizations

Kole Reddig*

December 16, 2020

Abstract

This paper estimates a structural model of multitasking agents to investigate the cost-quality tradeoff in health care and design contracts for a large physician incentive program. The setting involves Medicare's Accountable Care Organizations, which are groups of healthcare providers that receive incentive pay for spending below a cost target on shared patients. Estimation of the structural model and counterfactual simulations reveal a substantial tradeoff between reducing cost and increasing quality. The contract that maximizes the monetary savings of the incentive program increases savings by over $700 million per year, but it decreases quality of care by two standard deviations.

*John Chambers School of Business and Economics, West Virginia University. Email: kole.reddig@mail.wvu.edu. I thank my advisors Martin Gaynor and Maryam Saeedi for support throughout this project. I also thank Chris Cronin, Dennis Epple, Akshaya Jha, Rebecca Lessem, Albert Ma, Hakkı Özdenören, and Amanda Starc for comments and suggestions on this paper, as well as participants at seminars at Carnegie Mellon University, ITAM, American University, West Virginia University, the South Eastern Health Economics Study Group Conference, the University of Michigan H2D2 Research Day, the International Industrial Organization Conference, and the American Society of Health Economists Conference.

In the United States healthcare sector, public and private insurers often implement physician incentive programs and pay-for-performance initiatives to control the cost of care. Designing payment contracts for these programs requires facing a fundamental challenge: physicians may decrease the quality of care they provide in order to reduce cost. This issue is an example of agent multitasking, which plays a critical role in decision-making and contract design in all sectors of the economy. In this paper, I estimate a structural model of multitasking agents to identify the extent to which healthcare providers decrease quality of care in order to reduce cost. I use the structural model to conduct counterfactual analysis that highlights the role of multitasking in incentive design in the context of a large physician incentive program.

The setting of this study is the Medicare Shared Savings Program (MSSP), a large incentive program that involves 11 million Medicare beneficiaries and $100 billion in healthcare expenditure each year.[1] The MSSP gives incentive pay to Accountable Care Organizations (ACOs), which are joint ventures of physicians, group practices, and hospitals that form to coordinate care of their shared patients. An ACO earns incentive pay through the MSSP if its members collectively reduce expenditure on health services. Because providers might decrease the quality of care they provide in order to reduce expenditure, both tasks (monetary savings and quality of care) determine ACO payment. Moreover, because the earnings of a provider in an ACO depend heavily on the decisions of others, free-riding within ACOs may severely limit performance.

What role does multitasking play in the decisions of Medicare providers? How much is lost to free-riding in ACOs? Answers to these questions are central the design of the MSSP, and furthermore will inform incentive and

[1] Source: The Centers for Medicare and Medicaid Services. "Shared Savings Program Fast Facts." `https://go.cms.gov/32VBOnZ`.

contract design throughout the healthcare sector. I answer these questions by building and estimating a structural model of Medicare providers in ACOs. In the model, providers participating in an ACO act strategically and choose efforts to put towards cost-saving and quality of care in order to maximize their own payoff. The choices of efforts of each member in an ACO form a Nash equilibrium that describes the ACO's overall performance and the income of providers.

In counterfactual analyses, I solve for the contracts between ACOs and Medicare that maximize two different objectives. In the first objective, Medicare (the principal) cares only about reducing healthcare expenditure, and seeks to maximize the money saved by the incentive program, less payment to ACOs. In the second, Medicare has preferences for both cost savings and quality of care, so its objective is to maximize quality-weighted incentive program savings, less payment to ACOs. I allow contracts vary along two policy-relevant dimensions: 1) the generosity of the contract for a given level of quality of care and expenditure, and 2) whether ACOs also make penalty payments to Medicare if their expenditure exceeds its target.

My research design exploits well-defined and observed contracts between Medicare and ACOs in the MSSP to identify structural parameters. The MSSP works by assigning an ACO an expected expenditure for healthcare services provided to its members' patients. If a year's Medicare expenditure on those beneficiaries is less than the expected amount, an ACO earns a portion of the difference, adjusted by a quality score, as incentive pay (hence "sharing savings" with Medicare). The form of these contracts is specified by Medicare and is public information, so I observe cross-ACO variation in the marginal dollar of group incentive pay for a given level of cost-savings and quality of care. Under an equilibrium assumption, this identifies a function describing the marginal cost (i.e., a supply curve) of reducing expenditure and improving quality. Furthermore, this yields an empirical estimate of the magnitude of the tradeoff between cost-savings and quality—a key factor driving multi-tasking choices and a crucial component to computing contracts that make a combination of monetary savings and quality of care the objective.

I find a strong tradeoff between Medicare savings and quality of care: a one standard deviation increase in an ACO's savings rate (an increase in 5 percentage points) increases the cost of increasing quality of care one standard deviation by $6,700 per participating provider. Currently, contracts between ACOs and Medicare allow ACOs to earn up to 75% of the money they save as incentive pay. By simulating equilibrium outcomes under alternative contracts, I compute that the optimal amount of savings to share with an ACO is 44%, where Medicare increases the savings of the program by more than $100 million per year. Under two-sided contracts, where ACOs must pay money back to Medicare if they spend too much, savings rates are four times higher, implying a 352% increase in savings to Medicare. Quality scores decrease under two-sided contracts because ACOs incur significantly higher costs of increasing quality when saving more. When a combination of program savings and ACO quality scores is the objective, neither contract (with or without penalties for spending too much) strictly dominates the other.

Because the earnings of a provider in an ACO depend on group performance, providers act strategically and make decisions based on the actions of others. As more providers join an ACO, any one provider's influence on ACO outcomes diminishes. The result of this is incentive dilution and free-riding, and the optimal effort choices of providers are less than the effort choices that would maximize the total surplus to all providers. I find that program savings would increase by over $1 billion per year without free-riding within

ACOs. Because of free-riding, Medicare must pay more to ACOs: if every ACO perfectly coordinated, the program would maximize its monetary savings by sharing 35% of savings with ACOs.

This paper contributes to economics literature concerning evidence of multitasking and agent response to incentive pay (Slade, 1996; Bai & Xu, 2005; Dumont, Fortin, Jacquemet, & Shearer, 2008; Mullen, Frank, & Rosenthal, 2010; Feng Lu, 2012; Hong, Hossain, List, & Tanaka, 2018). Along with the study by Kim, Sudhir, & Uetake (2019), this paper is one of the very first to estimate a structural model of multitasking agents. I also contribute to economics literature concerning health care provider payment systems and provider behavior in organizations (Gaynor, Rebitzer, & Taylor, 2004; Encinosa, Gaynor, & Rebitzer, 2007; Choné & Ma, 2011; Rebitzer & Votruba, 2011; Ho & Pakes, 2014; Frandsen & Rebitzer, 2015; Grassi & Ma, 2016; Frandsen, Powell, & Rebitzer, 2017). More generally, this paper aligns with the literature that studies the supply-side of health care, and examines the incentives faced and decisions made by physicians, hospitals, and insurers (Gaynor, 2006; Chandra, Cutler, & Song, 2011; Gaynor, Ho, & Town, 2015; Foo, Lee, & Fong, 2017; Einav, Finkelstein, & Mahoney, 2018; Eliason, Grieco, McDevitt, & Roberts, 2018; Hackmann, 2019).

Few studies in economics have discussed ACOs directly. Frandsen & Rebitzer (2015) calibrate a simple model of ACO performance to examine the size-variance tradeoff in group payment mechanisms like the MSSP, and they argue that ACOs will be unable to self-finance. That is, there is no contract with strong enough incentives to overcome the incentive to free-ride among a group of physicians. The authors conclude with a skeptical look at the MSSP, and mention the untenability of integrated organizations in the now very fractured US health care market. Frech et al. (2015) study county-level entry of private and public ACOs. The authors find small markets generally discourage ACO entry, and that public ACO entry is largely predicted by higher Medicare spending, higher population, and lower physician site concentration. Frandsen et al. (2017) discuss the MSSP's impact on health care in the United States in the context of common agency, where several payers motivate the same agent to improve care delivery and integration. The authors find that unique equilibrium contracts from payers are lower powered in the presence of shared savings payments, and ACO entry can possibly inspire other shared savings contracts in the private sector if they do not already exist. Aswani, Shen, & Siddiq (2019) study how to design MSSP ACO contracts. The authors focus on asymmetric information between Medicare and ACOs, and write contracts such that ACO payment is a function of ACO characteristics (such as the number of beneficiaries assigned to an ACO). Unlike this paper, Aswani et al. (2019) do not consider multitasking or free-riding within ACOs.

This paper continues as follows: Section 1 gives a brief overview of the MSSP and ACOs, including descriptive ACO statistics. I outline my model of performance in ACOs in Section 2. I describe identification and estimation of model primitives in Section 3, and estimation results are in Section 4. I present counterfactual analysis, including computation of savings-maximizing contracts between ACOs and Medicare, in Section 5, and Section 6 concludes.

1 Background and Data

The MSSP, a part of the Patient Protection and Affordable Care Act of 2010 (ACA), is a policy response to increasing healthcare costs in the United States. The premise of the program is that the United States is inefficient at providing healthcare because care delivery is *fragmented*. That is, unique to the United

States, patients tend to see several distinct providers that belong to separate businesses with little incentive to coordinate care. Patients therefore receive haphazard and often redundant care, implying increased utilization, cost, and risk of adverse health outcomes.

The MSSP gives providers financial motivation to integrate care delivery. To overcome institutional boundaries to care integration, the program explicitly evaluates and pays Medicare providers based on group performance. First, providers join Accountable Care Organizations, or ACOs, which are joint ventures of Medicare providers created to earn payment through the MSSP. Nearly any Medicare provider, including individual physicians, group practices, and large hospital systems, can start or participate in an ACO. Medicare fee-for-service (FFS) beneficiaries are then assigned to ACOs by Medicare according to their primary care provider (PCP).[2]

The first ten rows of Table 1 display statistics describing ACO participants and beneficiaries assigned to ACOs.[3] There is substantial heterogeneity in the number of providers that join an ACO—some large hospitals are able to form an ACO independently by employing enough PCPs to be assigned the legally required minimum of 5000 beneficiaries, and others are joint ventures of hundreds of providers. Every state has beneficiaries assigned to an ACO, though most ACOs concentrate on beneficiaries in just one state. "Average risk score" is the average Hierarchical Condition Category (HCC) risk score of

[2]For the interested reader, Appendix A gives a very detailed description of MSSP ACO formation, beneficiary assignment, and payment.

[3]The data for this table, and all analysis in this section, is from MSSP ACO Public Use Files, MSSP Participant Lists, MSSP ACO Performance Year Results, and Number of ACO Assigned Beneficiaries by County Public Use Files. In short, the data consists of ACO expenditures, benchmark expenditures, quality scores (along with every quality sub-measure), various assigned beneficiary demographics, and various participant and provider statistics. Little public information is available on the characteristics of specific ACO participants or providers.

Table 1: **Summary ACO Statistics: Providers and Beneficiaries**

Variable	**Mean**	**Std. Dev.**
Number of participants[a]	37.79	58.03
Total number of individual providers (1000s)	0.59	0.84
Proportion of providers PCP	0.41	0.18
Proportion of providers specialist	0.41	0.21
Number of states where beneficiaries reside	1.52	0.97
Number of assigned beneficiaries (1000s)	17.78	17.49
Average risk score	1.06	0.11
Percent of beneficiaries over age 75	39.10	6.04
Percent of beneficiaries male	42.66	2.06
Percent of beneficiaries nonwhite	16.88	15.34
Sharing rate 50%, one-sided	0.96	0.19
Sharing rate 60%, two-sided	0.01	0.11
Sharing rate 75%, two-sided	0.03	0.16
Benchmark Expenditure ($ billions)	0.19	0.18
Expenditure ($ billions)	0.19	0.18
Benchmark Expenditure - Expenditure ($ millions)	1.46	10.18
Savings Rate	0.01	0.05
Quality Score	0.87	0.12
$\mathbf{1}\{$Savings Rate $\geq$ Min. Savings Rate$\}$	0.31	0.46
Earned shared savings or losses	1.50	3.64
Earned shared savings, given qualified	4.95	5.11
Proportion of expenditure on inpatient services[a]	0.31	0.03
Proportion of expenditure on outpatient services[a]	0.20	0.06
Number of primary care services (1000s)	10.29	1.76
Number of inpatient admissions (1000s)	0.33	0.09

$N = 1849$. This table shows summary statistics for ACOs for years 2013-2017. The superscript ^{a} indicates statistics are for 2014-2017 (due to data availability). “Quality Score” is computed by the author from ACO quality sub-measures (public data codes Quality Score as 1 or “P4R” in an ACO’s first performance year.)

non-dual eligible beneficiaries assigned to an ACO. A beneficiary's risk score increases as predicted healthcare costs of that beneficiary increase.

Payment of an ACO depends on a calendar year's Medicare expenditure on beneficiaries assigned to the ACO, a quality of care score, and the contract the ACO has with Medicare. Upon formation of an ACO, Medicare assigns a "benchmark expenditure" by forecasting Medicare expenditure for beneficiaries assigned to the ACO. After operating for a year, the ACO's payment is determined by the difference between the benchmark expenditure and realized expenditure on assigned beneficiaries and a composite quality score between 0 and 1.[4] If the ACO's savings rate, defined as $\frac{\text{Benchmark Expenditure}-\text{Expenditure}}{\text{Benchmark Expenditure}}$, exceeds a predetermined minimum, and if the ACO meets minimum quality of care standards, it earns and distributes to its members the amount

$$\text{Sharing Rate} \cdot \text{Quality Score} \cdot (\text{Benchmark Expenditure} - \text{Expenditure}) \quad (1)$$

where "Sharing Rate," a number between 0 and 1, is determined by the type of contract the ACO has with Medicare. While uncommon in the first few years of the MSSP, some contracts also penalize ACOs for having expenditure *larger* than benchmark expenditure. These are called "two-sided" contracts.

The overwhelming contract choice of ACOs, "Track 1," has a sharing rate of 50%. Under this contract, if a hypothetical ACO with a benchmark expenditure of \$186 million and minimum savings rate 0.02 had an expenditure of

[4] An ACO's overall quality score is determined by the combination of 30-40 sub-measures of care quality. These sub-measures fall into the domains of "Patient/Caregiver Experience," "Care Coordination/Patient Safety," "Preventative Health," and "At-Risk Population." Some sub-measures are survey responses (e.g., "ACO2: How Well Your Doctors Communicate"), while others are computed from Medicare Claims and aggregated to the ACO-level (e.g., "ACO21: Proportion of Adults who had blood pressure screened in past 2 years"). See `https://go.cms.gov/2xHy7Uo` for a full list of ACO quality scores for every performance year.

\$180 million with a quality score of 0.90, it would earn

$$0.5 \cdot 0.9 \cdot (\$186 \text{ million} - \$180 \text{ million}) = \$2.7 \text{ million} \tag{2}$$

in shared savings. Its savings rate is $(186 - 180)/186 = 0.03$, so the minimum savings rate is exceeded. Though paying a subsidy, Medicare saves money as well: on net, this ACO contributed a \$3.3 million decrease in Medicare expenditure, as it was paid \$2.7 million for saving \$6 million.

The last fifteen rows of Table 1 contain statistics on ACO performance. ACO benchmark expenditures and realized expenditures are large: the mean is approximately \$190 million, with several ACOs having expenditure over \$1 billion. From 2013 to 2017, ACOs saved money on average. However, less than one third of ACOs had a savings rate at least as large as their minimum savings rate, meaning most ACOs do not actually earn incentive pay. Average earned incentive pay is \$1.5 million per ACO, and given an ACO earns incentive pay, incentive pay is nearly \$5 million. Per participant, average earned incentive pay is \$189,108 unconditionally and \$609,654 among ACOs that qualify.

2 Model of Performance in ACOs

In this section, I develop a model of provider performance in ACOs. All decisions are made by Medicare providers, and occur in a static environment. I intentionally avoid modeling an ACO's management-level decisions—while ACO management does have influence over their members, it's ultimately the providers that see and treat assigned beneficiaries, so I assume participants are the relevant decision-makers. Any influence of management is modeled as unobserved heterogeneity, and I identify underlying structural parameters accordingly. ACO and provider participation are taken as given, where the set of all ACOs is $\mathcal{J}$ and the set of all providers is $\mathcal{I}$. Providers are indexed by i and ACOs are indexed by j.

Each member of an ACO chooses unobservable efforts to put towards savings and quality in order to maximize their own payoff.[5] ACO participants have full information on their peers, and each member in an ACO plays a simultaneous move game. An ACO's savings rate and quality score is the outcome of the Nash equilibrium strategies chosen by its participants. Though this model is written in a way such that decisions are made by individual participants, underlying structural parameters can be identified and estimated with aggregate, ACO level data. Section 3 details this process.

Suppose n_j participants are in ACO j, and let the set of all Medicare providers in ACO j be denoted I_j. All participants $i \in I_j$ simultaneously choose savings and quality efforts $s_{ij} \in [-1, 1]$ and $q_{ij} \in [0, 1]$.[6] These choices determine ACO savings rate S_j and overall quality score Q_j through the weighted sums

$$S_j = \sum_{i \in I_j} w_{ij} s_{ij} \qquad\qquad Q_j = \sum_{i \in I_j} w_{ij} q_{ij} \,. \tag{3}$$

Here, $\{w_{ij}\}_{i \in I_j}$ are exogenous influence weights such that $w_{ij} \geq 0$ for all $i \in I_j$ and $\sum_{i \in I_j} w_{ij} \equiv 1$. These weights account for heterogeneous influence of

[5]These participant-level efforts are unobservable. That is, ACO participants are not assigned a benchmark expenditure, and are not given quality scores, and so observable values of effort do not exist. However, participants act as if making an effort choice, effort choices map to ACO performance measures that are observed.

[6]Savings effort s_{ij} is restricted to the domain $[-1, 1]$—this implicitly restricts an ACO's total expenditure to be between zero and twice its benchmark expenditure. The upper bound on expenditure is arbitrary, and exists only so that strategy spaces of agents are compact. Quality effort q_{ij} is restricted to $[0, 1]$ so that overall quality score also falls between $[0, 1]$ (which is always the case in the MSSP).

participants' efforts on ACO performance.[7]

Each participant $i \in I_j$ solves the profit maximization problem

$$\max_{s_{ij}, q_{ij}} R_{ij}(S_j, Q_j) - c_{ij}\left(s_{ij}, q_{ij}\right) \tag{4}$$

where $R_{ij}(S_j, Q_j)$ is provider i's portion of shared savings earned by an ACO with savings S_j and quality score Q_j, and c_{ij} is the strictly convex and twice-continuously differentiable participant effort cost function. Specifically, $c_{ij}\left(s_{ij}, q_{ij}\right)$ is the explicit and implicit costs incurred by $i \in I_j$ when choosing s_{ij} and q_{ij}. For example, a physician that chooses very large values of s_{ij} and q_{ij} would incur significant cost—both in operational expenses as well as opportunity cost from forgone services to reduce expenditure on assigned beneficiaries. There is no direct utility gained for quality of care through provider altruism, which is common in models of physician decision-making (Glied & Hong, 2018; Hackmann, 2019). In this model, altruistic preference for increasing quality of care is absorbed by the cost function and effectively decreases the marginal cost of quality effort. Ultimately, c_{ij} places a natural restriction on how well participants, and hence ACOs, can perform.

Medicare determines contracts for ACOs: under the contract named Track 1, the incentive pay earned by ACO j takes the known and exogenous form

$$R_j(S_j, Q_j) = \begin{cases} 0.5 \cdot B_j S_j Q_j & \text{if } S_j \geq \underline{S}_j \text{ and } Q_j \geq \underline{Q} \\ 0 & \text{otherwise} \end{cases} \tag{5}$$

[7] For example, consider an ACO with $n_j = 2$ participants: a hospital with savings effort analogous to saving 2%, and an individual provider with savings effort analogous to saving 4%. This means $s_{1j} = 0.02$, $s_{2j} = 0.04$, and $\overline{s}_j = 0.03$. The ACO's savings rate, however, would be far closer to $S_j \approx 0.02$ since the hospital has a larger share of overall expenditure. See Appendix C for more details.

where B_j is the benchmark expenditure of ACO j, $\underline{S}_j$ is the minimum savings rate for ACO j, and $\underline{Q}$ is the quality reporting standard.[8] I assume shared savings is distributed to participants according to influence weights w_{ij}, so $R_{ij}(S_j, Q_j) = w_{ij} R_j(S_j, Q_j)$, and the two first order conditions for participant i are then

$$\frac{\partial c_{ij}}{\partial s_{ij}}(s_{ij}, q_{ij}) = \begin{cases} 0.5 \cdot B_j w_{ij}^2 Q_j & \text{if } S_j \geq \underline{S}_j \text{ and } Q_j \geq \underline{Q} \\ 0 & \text{otherwise} \end{cases} \tag{7}$$

$$\frac{\partial c_{ij}}{\partial q_{ij}}(s_{ij}, q_{ij}) = \begin{cases} 0.5 \cdot B_j w_{ij}^2 S_j & \text{if } S_j \geq \underline{S}_j \text{ and } Q_j \geq \underline{Q} \\ 0 & \text{otherwise} \end{cases}. \tag{8}$$

I've assumed with the specification of R_{ij} that ACOs split their earned shared savings with their participants according to influence weights w_{ij}, and not evenly between participants. Actual contracts between ACOs and ACO participants (known as "ACO Participant Agreements") are generally not publicly available. However, splitting shared savings according to influence on ACO outcomes is a good approximation of how ACOs actually split earnings.[9] For example, Gaynor et al. (2004) make the similar assumption that HMO group incentive pay is allocated among the group according to physician patient shares.

I prove existence of equilibrium in this game in Proposition B.1 in Appendix B. I also show, in general, there is not a unique equilibrium. There can be

[8]ACOs in their first performance year are "paid to report", and so shared savings takes the form

$$R_j(S_j, Q_j) = \begin{cases} 0.5 \cdot B_j S_j & \text{if } S_j \geq \underline{S}_j \text{ and } Q_j \geq \underline{Q} \\ 0 & \text{otherwise} \end{cases} \tag{6}$$

—in other words, Q_j is equivalently 1 when an ACO meets quality reporting standards in the first performance year.

[9]See `https://go.cms.gov/2HiHgus` for more detail.

up to two equilibria: one where the ACO qualifies or shared savings, and one where it does not. In either case, this is not an issue for estimation, since the equilibrium being played is observed in data. Intuitively, ACO participants are playing a simultaneous move coordination game. One equilibrium occurs when all participants $i \in I_j$ choose effort choices that solve

$$\max_{s_{ij}, q_{ij}} 0.5 \cdot w_{ij} B_j S_j Q_j - c_{ij}\left(s_{ij}, q_{ij}\right) , \tag{9}$$

and the ACO qualifies for shared savings. The other equilibrium occurs when all participants solve

$$\min_{s_{ij}, q_{ij}} c_{ij}\left(s_{ij}, q_{ij}\right) , \tag{10}$$

and the ACO does not qualify for shared savings. Let $\mathbf{s}_j = \left[s_{1j}, \ldots, s_{n_j j}\right]'$ and $\mathbf{q}_j = \left[q_{1j}, \ldots, q_{n_j j}\right]'$. Denote a Nash equilibrium strategy of participant i in ACO j as $\left(s_{ij}^*, q_{ij}^*\right)$ and a Nash equilibrium of the game as $\left(\mathbf{s}_j^*, \mathbf{q}_j^*\right)$. Accordingly, the ACO's saving rate and overall quality score resulting from the set of Nash equilibrium strategies are denoted S_j^* and Q_j^*.

3 Identification and Estimation

To estimate model primitives, I use ACO-level data on Track 1 ACOs from 2014 to 2017.[10] The first year of the program, 2013, is omitted because all ACOs are "paid to report" and quality scores are always set equal to one. Track 2 and 3 ACOs are omitted because these ACOs choose to face downside risk when in the MSSP, so selection may bias estimates if these ACOs

[10]Summary statistics for the data are presented in Table 1 in Section 1.

Table 2: **Summary of Model and Data used for Estimation**

Primitives	Data
Parameters in parameterized cost functions c_{ij}: $\boldsymbol{\theta} = \left\{\delta_S, \delta_Q, \boldsymbol{\gamma}_S, \boldsymbol{\gamma}_Q, \kappa\right\}$.	ACO savings rate: S_j ACO quality score: Q_j ACO benchmark exp.: B_j Summed-cubed share of expenditures: $W_j^{(3)}$ ACO characteristics : X_j

This table summarizes model primitives and the data used to identify each primitive.

were included. Identification and estimation of this model is complicated by the limited data available on ACO participants. Only aggregate data are observed: decisions in the model are made by Medicare providers, and available data describes the outcome of these agents' decisions aggregated to the ACO level. To overcome the challenge imposed by data availability, I use a novel aggregation procedure to map ACO characteristics and performance to model primitives.

Table 2 gives an overview of the model primitives and data used for estimation of the primitives. I estimate the parameters of an average (across participants) cost function, $\bar{c}_j(\cdot)$, up to fixed costs and provider-specific marginal costs. These parameters are denoted by $\boldsymbol{\theta}$. I assume that observed ACO savings rates and quality scores are savings rates S_j^* and quality scores Q_j^* from a Nash equilibrium. Equilibrium selection is not required for estimation since the equilibrium played (qualifying or not qualifying for shared savings) is observed.

3.1 Aggregation of First Order Conditions

The key estimates in this paper are identified from variation in savings rates and overall quality scores across ACOs *given the observed marginal incentive* *pay for each measure of performance.* For example, a large positive correlation of savings rates and quality scores across ACOs is not evidence of complementarity of savings and quality. If ACOs with large quality scores also tend to have large marginal incentive pay with respect to savings, then this positive correlation could exist in the presence of no complementarity or a tradeoff. Incorporating the structure of ACO participants' incentives is essential to obtain the results in this paper.

Recall participant first order conditions in Equations 7 and 8. These hold for any Nash equilibrium, so:

$$\frac{\partial c_{ij}}{\partial s_{ij}}\left(s_{ij}^*, q_{ij}^*\right) = \left(0.5 B_j w_{ij}^2 Q_j^*\right) \mathbf{1}\left\{S_j^* \geq \underline{S}_j\right\} \mathbf{1}\left\{Q_j^* \geq \underline{Q}\right\} \ \forall i \in I_j \tag{11}$$

$$\frac{\partial c_{ij}}{\partial q_{ij}}\left(s_{ij}^*, q_{ij}^*\right) = \left(0.5 B_j w_{ij}^2 S_j^*\right) \mathbf{1}\left\{S_j^* \geq \underline{S}_j\right\} \mathbf{1}\left\{Q_j^* \geq \underline{Q}\right\} \ \forall i \in I_j \,. \tag{12}$$

That is, for any Nash equilibrium, the marginal cost of savings effort (the left hand side of Equation 11) is equal to the marginal subsidy of savings effort (the right hand side of Equation 11) for all participants $i \in I_j$. Similarly, the marginal cost of quality effort is equal to the marginal benefit of quality effort (Equation 12). Pre-multiplying each side of the equations by w_{ij} and summing over $i \in I_j$, we get

$$\sum_{i \in I_j} w_{ij} \frac{\partial c_{ij}}{\partial s_{ij}}\left(s_{ij}^*, q_{ij}^*\right) = 0.5 W_j^{(3)} B_j Q_j^* \mathbf{1}\left\{S_j^* \geq \underline{S}_j\right\} \mathbf{1}\left\{Q_j^* \geq \underline{Q}\right\} \tag{13}$$

$$\sum_{i \in I_j} w_{ij} \frac{\partial c_{ij}}{\partial q_{ij}}\left(s_{ij}^*, q_{ij}^*\right) = 0.5 W_j^{(3)} B_j S_j^* \mathbf{1}\left\{S_j^* \geq \underline{S}_j\right\} \mathbf{1}\left\{Q_j^* \geq \underline{Q}\right\} \,, \tag{14}$$

where term $W_j^{(3)} \equiv \sum_{i \in I_j} w_{ij}^3 \equiv \sum_{i \in I_j} w_{ij}^3$ is a measure of influence concentration within an ACO (similar to a Herfindahl-Hirschman index [HHI]), and is computed from data as the sum of cubed shares of expenditure for each type of provider within an ACO. The term $W_j^{(3)}$ essentially discounts the marginal benefit of savings and quality at the ACO level according to the number of providers and dispersion of influence in an ACO. The computation of $W_j^{(3)}$ from data is discussed in detail in Appendix C.

Equations 13 and 14 are aggregate analogs of Equations 11 and 12. These state the weighted average of marginal cost of savings across ACO participants is equal to the weighted average marginal benefit of savings across ACO participants, and the weighted average of marginal cost of quality across ACO participants is equal to the weighted average marginal benefit of quality across ACO participants. For notational ease, let the left hand sides of Equations 13 and 14 be denoted MC_j^S and MC_j^Q, respectively, and similarly let the right hand sides of Equations 13 and 14 be denoted MB_j^S and MB_j^Q, respectively.

Here, data limitations become evident. Since within-ACO variation across participants is unavailable, its impossible to identify a specific marginal cost for each participant in each ACO. Weighted average marginal costs MC_j^S and MC_j^Q are the fullest description of marginal cost available. These values are still informative, however: because each is weighted by participant influence on ACO outcomes, they can be interpreted as the marginal costs of savings effort and quality effort for a representative participant in a given ACO.

Next, I make the following functional form assumption.

Assumption 3.1. The effort cost function c_{ij} takes the quadratic form

$$c_{ij}\left(s_{ij}, q_{ij}\right) = \frac{\delta_S}{2} s_{ij}^2 + \frac{\delta_Q}{2} q_{ij}^2 + \boldsymbol{\gamma}_S' x_{ij} s_{ij} + \boldsymbol{\gamma}_Q' x_{ij} q_{ij} + \kappa s_{ij} q_{ij} \,, \tag{15}$$

where $x_{ij} \in \mathbb{R}^k$ is a vector of participant and ACO specific characteristics, and the k-dimensional coefficient vectors $\boldsymbol{\gamma}_S$ and $\boldsymbol{\gamma}_Q$ map these characteristics to marginal cost of savings and quality, respectively.

Under Assumption 3.1, we have

$$\frac{\partial c_{ij}}{\partial s_{ij}}\left(s_{ij}^{*}, q_{ij}^{*}\right) = \delta_S s_{ij}^{*} + \boldsymbol{\gamma}_S' x_{ij} + \kappa q_{ij}^{*} \tag{16}$$

$$\frac{\partial c_{ij}}{\partial q_{ij}}\left(s_{ij}^{*}, q_{ij}^{*}\right) = \delta_Q q_{ij}^{*} + \boldsymbol{\gamma}_Q' x_{ij} + \kappa s_{ij}^{*}\,. \tag{17}$$

Using the definitions $S_j^* = \sum_{i \in I_j} w_{ij} s_{ij}^*$, $Q_j^* = \sum_{i \in I_j} w_{ij} q_{ij}^*$, and $\sum_{i \in I_j} w_{ij} = 1$, we can derive a functional form for MC_j^S that's linear in parameters:

$$MC_j^S = \sum_{i \in I_j} w_{ij} \frac{\partial c_{ij}}{\partial s_{ij}}\left(s_{ij}^{*}, q_{ij}^{*}\right) \tag{18}$$

$$= \delta_S \left(\sum_{i \in I_j} w_{ij} s_{ij}^{*}\right) + \boldsymbol{\gamma}_S' \left(\sum_{i \in I_j} w_{ij} x_{ij}\right) + \kappa \left(\sum_{i \in I_j} w_{ij} q_{ij}^{*}\right) \tag{19}$$

$$= \delta_S S_j^* + \boldsymbol{\gamma}_S' X_j + \kappa Q_j^*\,, \tag{20}$$

where $X_j \equiv \sum_{i \in I_j} w_{ij} x_{ij}$ is a k-dimensional vector of the weighted averages of provider characteristics, x_{ij}. Similarly for MC_j^Q:

$$MC_j^Q = \sum_{i \in I_j} w_{ij} \frac{\partial c_{ij}}{\partial q_{ij}}\left(s_{ij}^{*}, q_{ij}^{*}\right) \tag{21}$$

$$= \delta_Q \left(\sum_{i \in I_j} w_{ij} q_{ij}^{*}\right) + \boldsymbol{\gamma}_Q' \left(\sum_{i \in I_j} w_{ij} x_{ij}\right) + \kappa \left(\sum_{i \in I_j} w_{ij} s_{ij}^{*}\right) \tag{22}$$

$$= \delta_Q Q_j^* + \boldsymbol{\gamma}_Q' X_j + \kappa S_j^*\,. \tag{23}$$

Combining this with Equations 13 and 14, we get

$$MB_j^S = \delta_S S_j^* + \boldsymbol{\gamma}_S' X_j + \kappa Q_j^* \tag{24}$$

$$MB_j^Q = \delta_Q Q_j^* + \boldsymbol{\gamma}_Q' X_j + \kappa S_j^*\,. \tag{25}$$

I assume that MB_j^S and MB_j^Q are observed with additive errors terms

$$\nu_j^S(\boldsymbol{\theta}) = MB_j^S - \delta_S S_j^* + \boldsymbol{\gamma}_S' X_j + \kappa Q_j^* \tag{26}$$
$$\nu_j^Q(\boldsymbol{\theta}) = MB_j^Q - \delta_Q Q_j^* + \boldsymbol{\gamma}_Q' X_j + \kappa S_j^* \tag{27}$$

where $\boldsymbol{\theta} = \{\delta_S, \delta_Q, \boldsymbol{\gamma}_S, \boldsymbol{\gamma}_Q, \kappa\}$. Next, I make the following identifying assumption.

Assumption 3.2. The error terms $\nu_j^S(\boldsymbol{\theta})$ and $\nu_j^Q(\boldsymbol{\theta})$ have mean zero and are independent of S_j^*, Q_j^* and X_j.

Assumption 3.2 is necessary so that estimates of $\boldsymbol{\theta}$ are unbiased. Intuitively, these errors explain why two observationally identical ACOs may have different weighted marginal benefits of savings and quality. The reason these errors exist is well-established in health policy literature: some ACOs, after conditioning on observed characteristics, are favored by a larger assignment of benchmark expenditure B_j (McWilliams, 2014; McWilliams et al., 2018). Insofar as those unobserved factors driving larger benchmark expenditure assignment are uncorrelated with ACO-specific marginal cost, parameters describing marginal cost in $\boldsymbol{\theta}$ are identified.[11]

The elements of X_j, which can be interpreted as ACO-specific marginal cost shifters, are described in Table 7 in Appendix D. These elements include information concerning assigned beneficiaries, participating Medicare providers, as well as expenditure and service statistics.

[11] A common concern of identification strategies that leverage first order conditions is that errors are realized before optimization and are observed by agents. If this is the case, choices of agents are necessarily correlated with errors, rendering biased estimates. This paper assumes that errors are unobserved by both the econometrician and agents, and those marginal cost shocks that are otherwise unaccounted for are nonetheless "controlled for" by marginal cost shifters in X_j. If this assumption does not hold, a solution is to use instruments for S_j^* and Q_j^* (lagged values are a common choice) to obtain unbiased estimates.

To estimate $\boldsymbol{\theta}$, I apply Assumption 3.2 and use the moment conditions

$$\mathbb{E}\left[\begin{array}{c} \nu_j^S(\boldsymbol{\theta}) \\ \nu_j^Q(\boldsymbol{\theta}) \end{array} \middle| S_j^*, Q_j^*, X_j\right] = 0 \tag{28}$$

and the Generalized Method of Moments (GMM) estimator (Hansen, 1982). That is,

$$\hat{\boldsymbol{\theta}} = \arg\min_{\theta\in\Theta} \left(\boldsymbol{\nu}(\boldsymbol{\theta})' \left[\begin{array}{ccc} \mathbf{S} & \mathbf{Q} & \mathbf{X}^{perf} \\ \mathbf{S} & \mathbf{Q} & \mathbf{X}^{perf} \end{array}\right]\right) \mathbf{W} \left(\boldsymbol{\nu}(\boldsymbol{\theta})' \left[\begin{array}{ccc} \mathbf{S} & \mathbf{Q} & \mathbf{X}^{perf} \\ \mathbf{S} & \mathbf{Q} & \mathbf{X}^{perf} \end{array}\right]\right)' \tag{29}$$

where $\boldsymbol{\nu}(\boldsymbol{\theta}) = \left[\begin{array}{c} \nu_j^S(\boldsymbol{\theta}) \\ \nu_j^Q(\boldsymbol{\theta}) \end{array}\right]_{j\in\mathcal{J}}$ is a $2J \times 1$ dimensional vector of stacked errors; $\mathbf{S}$, $\mathbf{Q}$, and $\mathbf{X}^{perf}$ are similarly vectors and the matrix of stacked observations of S_j^*, Q_j^*, and X_j; and $\mathbf{W}$ is a $2(k+2) \times 2(k+2)$ positive definite weighting matrix.[12]

3.2 Illustrative Example

The functional form assumption on cost is nessecary to identify a cost function from aggregate outcomes. Nonetheless, several results of this paper remain for any cost function. In particular, an empirical estimate of the savings-quality tradeoff is by definition the change in marginal cost of savings with respect to quality, or $\frac{\partial \hat{MC}_j^S}{\partial Q_j^*}$. While this is given a single parameter, κ, above, the tradeoff would nonetheless be identified for any cost function satisfying model

[12]Note that because the cross equation restriction that κ appears in both $\nu_j^S(\boldsymbol{\theta})$ and $\nu_j^Q(\boldsymbol{\theta})$, $\boldsymbol{\theta}$ is over-identified. There are $2k+4$ moments and $2k+3$ parameters in $\boldsymbol{\theta}$. In a separate estimation, I remove the cross equation restriction and allow κ to differ in each equation. The resulting parameter estimates are not significantly different, which is consistent with the structural interpretation of κ that $\frac{\partial^2 c_{ij}}{\partial s_{ij}\partial q_{ij}} \equiv \frac{\partial^2 c_{ij}}{\partial q_{ij}\partial s_{ij}}$.

assumptions. To gain intuition for identifying the shape of ACO participants' cost functions (including a savings-quality tradeoff), let's consider a simple example. This example is written without assuming an explicit functional form for c_{ij} so that it's clear that results regarding the shape of participant's cost functions stem from variation in ACO outcomes and not a specific functional form assumption.

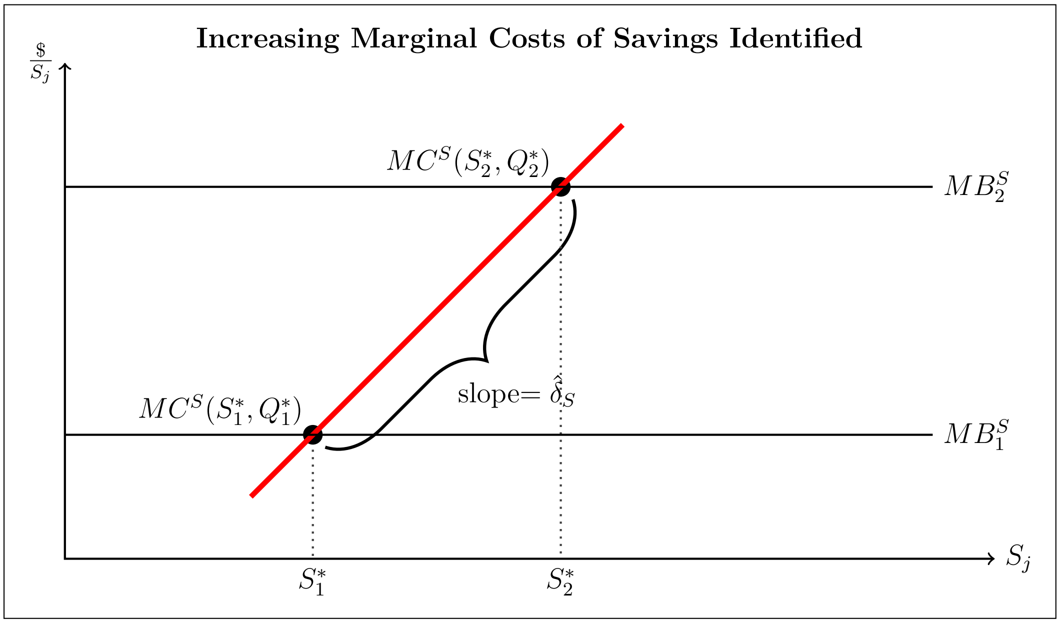
First suppose there are two ACOs with weighted average marginal benefit of savings such that $MB_1^S < MB_2^S$, where $S_1^* < S_2^*$ and $Q_1^* = Q_2^*$. Since marginal benefit is (on average) equal to marginal cost in equilibrium, we have $S_1^* < S_2^*$ and $MC^S(S_1^*, Q_1^*) < MC^S(S_2^*, Q_2^*)$. Marginal cost is of savings increasing in savings, so cost is convex in savings. The average increase in this case would then be equal to

$$\frac{\partial^2 c_{ij}}{\partial s_{ij}^2} \approx \frac{MC^S(S_2^*, Q_2^*) - MC^S(S_1^*, Q_1^*)}{S_2^* - S_1^*} \equiv \hat{\delta}_S \,. \tag{30}$$

The argument is pictured in Figure 1. Dollars are on the y-axis, and ACO savings rate is on the x-axis. The slope of the line connecting points at $\left(S_1^*, MC^S(S_1^*, Q_1^*)\right)$ and $\left(S_2^*, MC^S(S_2^*, Q_2^*)\right)$ is $\hat{\delta}_S$. The variation in weighted average marginal benefit identifies the slope (with respect to ACO savings rate) of the marginal cost of savings.

In this example, I've assumed identical quality scores, so variation across just these two hypothetical ACOs does not identify a savings-quality tradeoff. To show variation that identifies a savings-quality tradeoff, suppose there is another ACO with the same marginal benefit as ACO 2, but a different savings rate and quality score. Specifically, suppose ACO 3 is observed such that $MB_1^S < MB_2^S = MB_3^S$, $S_1^* < S_2^* < S_3^*$, and $Q_1^* = Q_2^* > Q_3^*$. Define $\hat{\delta}_S$ as before. The change in marginal cost of savings with respect to quality (the

Figure 1: **Identification of Marginal Cost**



Note: This figure shows how a convex cost function is identified from observed values of marginal benefit of savings. Given two ACOs with different marginal benefits of savings MB_1^S and MB_2^S (that are otherwise identical), the observed difference between their chosen savings rates S_1^* and S_2^* identifies the change in marginal cost of savings with respect to savings.

savings-quality tradeoff) is equal to

$$\frac{\partial^2 c_{ij}}{\partial s_{ij} \partial q_{ij}} \approx \frac{MC^S(S_2^*, Q_2^*) - MC^S(S_2^*, Q_3^*)}{Q_2^* - Q_3^*}\,. \tag{31}$$

Then, applying the defintion of $\hat{\delta}_S$, we have

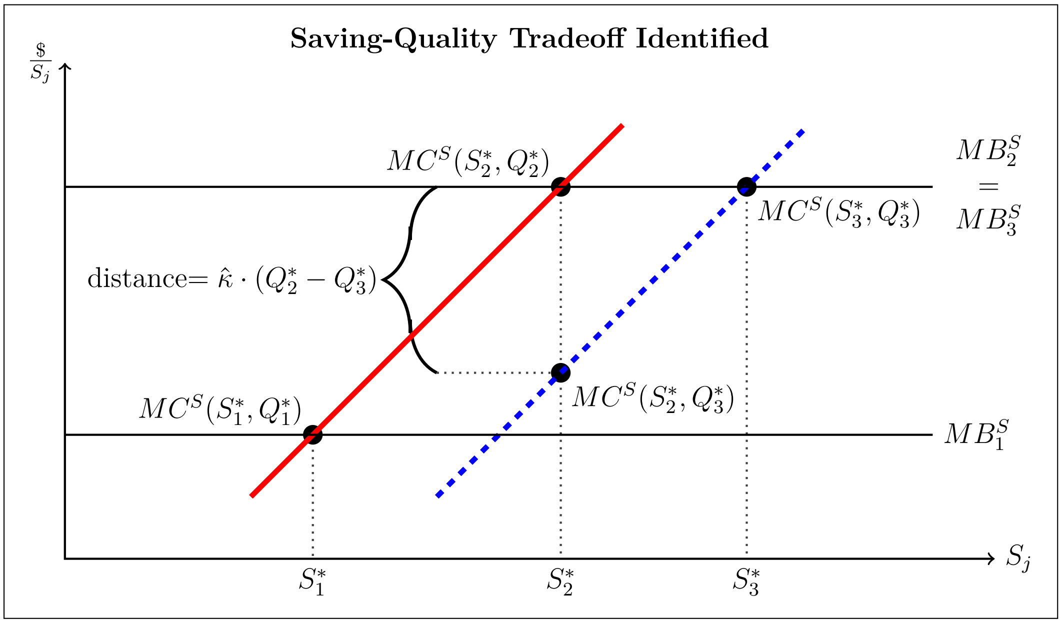
$$\begin{aligned}
\frac{\partial^2 c_{ij}}{\partial s_{ij} \partial q_{ij}} &\approx \frac{MC^S(S_2^*, Q_2^*) - MC^S(S_2^*, Q_3^*)}{Q_2^* - Q_3^*} && (32) \\
&= \frac{MC^S(S_2^*, Q_2^*) - \left[MC^S(S_3^*, Q_3^*) - \hat{\delta}_S(S_3^* - S_2^*)\right]}{Q_2^* - Q_3^*} && (33) \\
&= \frac{MC^S(S_2^*, Q_2^*) - \left[MC^S(S_2^*, Q_2^*) - \hat{\delta}_S(S_3^* - S_2^*)\right]}{Q_2^* - Q_3^*} && (34) \\
&= \hat{\delta}_S \frac{S_3^* - S_2^*}{Q_2^* - Q_3^*} \equiv \hat{\kappa}\,. && (35)
\end{aligned}$$

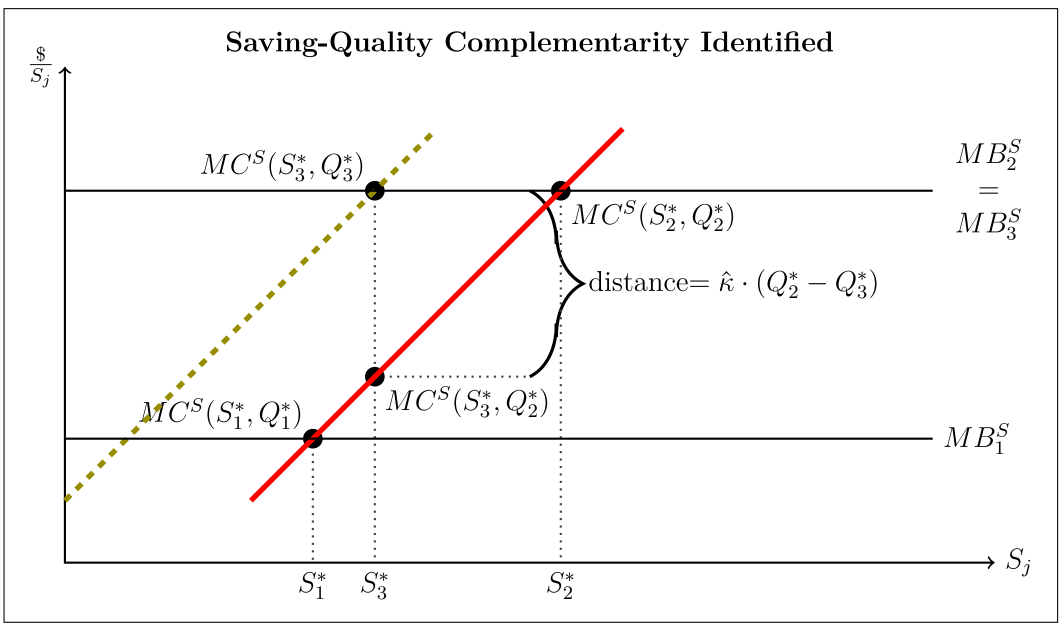
Since $Q_2^* > Q_3^*$ and $S_3^* > S_2^*$, this means the marginal cost of savings is increasing in quality.

Figure 2 shows this process for identifying a tradeoff in the top panel. The solid red line is the same that was found in Figure 1. Since ACOs 2 and 3 have the same marginal revenue of savings (and thus marginal cost), I can compute the marginal cost of savings for an ACO with the savings rate of ACO 2 and quality score of ACO 3, $MC^S(S_2^*, Q_3^*)$. Then, the difference between $MC^S(S_2^*, Q_2^*)$ and $MC^S(S_2^*, Q_3^*)$ (the vertical difference between the red solid line and blue dashed line) is the increase in marginal cost of savings for an increase in quality from Q_3^* to Q_2^*.

The bottom panel of Figure 2 shows how complementarity of savings and quality can be identified. The setup remains the same, except the savings rate of ACO 2 is now greater than the quality score of ACO 3: $MB_1^S < MB_2^S = MB_3^S$, $S_1^* < S_2^* > S_3^*$, and $Q_1^* = Q_2^* > Q_3^*$. Note that $\hat{\kappa}$ is negative in this case,

Figure 2: **Identification of Savings-Quality Tradeoff or Complementarity**





Note: In the top panel, this figure shows how a savings-quality tradeoff is identified from observed values of marginal benefit of savings. Given three ACOs such that marginal benefits of savings are $MB_1^S < MB_2^S = MB_3^S$, savings rates are $S_1^* < S_2^* < S_3^*$, and quality scores are $Q_1^* = Q_2^* > Q_3^*$, the observed difference between savings rates S_2^* and S_3^* identifies the change in marginal cost of savings with respect to quality. The bottom panel shows the analogous case for identifying complementarity between savings and quality, where $S_1^* < S_3^* < S_2^*$.

since increasing the quality score from Q_3^* to Q_2^* decreases marginal cost by $\hat{\kappa} \cdot (Q_2^* - Q_3^*)$.

4 Estimation Results

The estimated cost function parameters, $\hat{\boldsymbol{\theta}}$, are presented in Table 3. The three parameters controlling the shape of the cost function are estimated precisely, and the resulting cost function satisfies the properties required for an equilibrium to exist in the game played by ACO participants in every ACO. The parameter κ, which is the cross partial derivative of cost with respect to savings rate and quality, has a considerably high estimate. Increasing savings effort by one standard deviation makes a one standard deviation increase in quality effort nearly \$6,700 more costly per participant. Increasing quality effort by one standard deviation increases the cost of increasing savings effort by one standard deviation by more than \$7,500 per participant. This means there is a significant tradeoff between producing ACO savings and increasing quality of care. Other parameter estimates in Table 3 indicate several determinants of the marginal cost of savings and quality.

5 Counterfactuals

In this section, I use the estimated model of participation and performance to evaluate the available contracts between ACOs and CMS. In order to perform counterfactual exercises, I must solve for equilibrium values of S_j^* and Q_j^* using the estimate of $\boldsymbol{\theta}$. The following systems of two equations hold for any equilibrium and can be solved for the two unknowns, $\hat{S}_j^*$ and $\hat{Q}_j^*$. The first

Table 3: **Cost Function Parameter Estimates**

$$c(s,q) = (\delta_S/2)s^2 + (\delta_Q/2)q^2 + \boldsymbol{\gamma}_S s + \boldsymbol{\gamma}_Q q + \kappa sq$$

Coef.	Variable	Estimate	Std. Err.	P-value
	δ_S	271.130	37.115	0.000
	δ_Q	1.693	0.417	0.000
	κ	15.533	6.049	0.010
$\boldsymbol{\gamma}_S$	# states	4.460	2.067	0.031
	# beneficiaries	0.210	0.143	0.142
	average risk score	116.170	38.847	0.003
	% over 75	0.131	0.242	0.587
	% nonwhite	-0.124	0.061	0.041
	% male	1.097	1.082	0.310
	# providers	-2.195	2.017	0.276
	fraction PCP	5.292	7.131	0.458
	fraction inpatient	-134.000	60.540	0.027
	fraction outpatient	-113.930	30.472	0.000
	# PC services	-4.279	1.335	0.001
	# admissions	-60.862	27.000	0.024
	fraction PC served by PCP	6.711	10.569	0.525
	all group	29.339	5.204	0.000
$\boldsymbol{\gamma}_Q$	# states	-5.031	3.003	0.094
	# beneficiaries	0.327	0.198	0.099
	average risk score	0.011	0.011	0.318
	% over 75	8.352	3.910	0.033
	% nonwhite	0.009	0.017	0.594
	% male	-0.004	0.004	0.294
	# providers	0.055	0.081	0.497
	fraction PCP	-0.152	0.153	0.320
	fraction inpatient	0.403	0.475	0.396
	fraction outpatient	-10.024	4.461	0.025
	# PC services	-6.613	2.015	0.001
	# admissions	-0.298	0.104	0.004
	fraction PC served by PCP	-4.148	2.426	0.087
	all group	0.598	0.680	0.380
	N		1486	

Standard errors and p-values are from bootstrapping with 1000 rep. Estimates include year and Census Division FE. δ_S, δ_Q, and κ are scaled estimates.

system is when the ACO qualifies for incentive pay:

$$F \cdot W_j^{(3)} B_j \hat{Q}_j^* = \hat{\delta}_S \hat{S}_j^* + \hat{\boldsymbol{\gamma}}'_S X_j + \hat{\kappa} \hat{Q}_j^* \tag{36}$$

$$F \cdot W_j^{(3)} B_j \hat{S}_j^* = \hat{\delta}_Q \hat{Q}_j^* + \hat{\boldsymbol{\gamma}}'_Q X_j + \hat{\kappa} \hat{S}_j^* \tag{37}$$

The variable F is the sharing rate for the ACO. The second system is when the ACO does not qualify for incentive pay:

$$0 = \hat{\delta}_S \hat{S}_j^* + \hat{\boldsymbol{\gamma}}'_S X_j + \hat{\kappa} \hat{Q}_j^* \tag{38}$$

$$0 = \hat{\delta}_Q \hat{Q}_j^* + \hat{\boldsymbol{\gamma}}'_Q X_j + \hat{\kappa} \hat{S}_j^* \,. \tag{39}$$

These systems both have guaranteed solutions given assumptions on c_{ij}. For each ACO, I check if the first solution is an equilibrium and satisfies $\hat{S}_j^* \geq \underline{S}_j$ and $\hat{Q}_j^* \geq \underline{Q}$. In the event solutions to both of these systems are equilibria, I impose the following equilibrium selection rule.

Assumption 5.1.

The equilibrium that's played is the utilitarian equilibrium where the ACO qualifies for shared savings.

Note that while Assumption 5.1 must be imposed to simulate ACO savings rates and quality scores, it's not imposed to estimate the model, because the equilibrium that is played is observed directly in the data.

5.1 Evaluating Existing Contracts between Medicare and ACOs

From the beginning of the Medicare Shared Savings Program in 2012 and until June 2019, ACOs had four contract options: Track 1, Track 1+, Track

Table 4: **ACO Contract Options**

Sharing Rate	Only Earns Shared Savings	Also Pays Shared Losses
0.40	Basic Track A, B	
0.50	Track 1	Track 1+, Basic Track C, D, E
0.60		Track 2
0.75		Track 3, Enhanced Track

This table displays sharing rates and downside risk presence for existing ACO contract options. Tracks 1, 1+, and 2 were replaced by Basic Tracks A, B, C, D, and E in July 2019. Track 3 was renamed to Enhanced Track in July 2019.

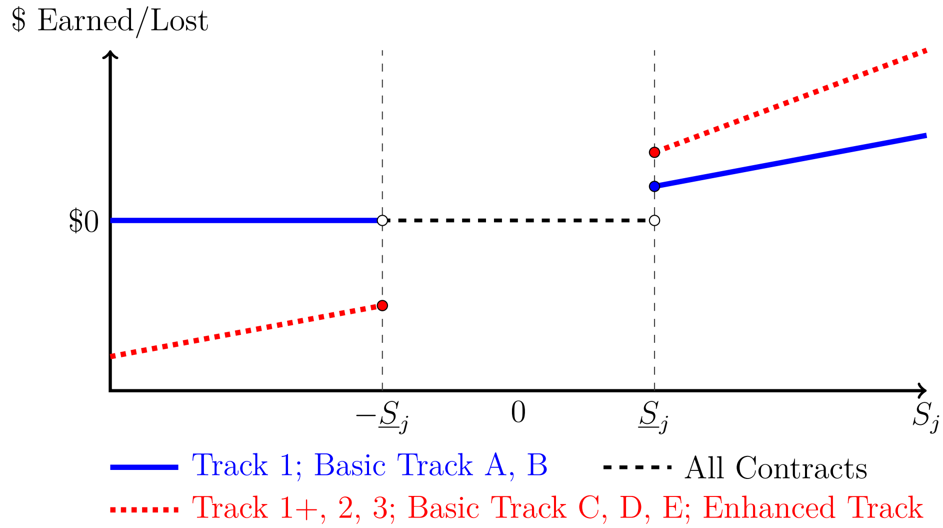
2, and Track 3. Starting in July 2019, these four contracts were replaced with the Basic Track (and its five levels A, B, C, D, and E) and the Enhanced Track. These contracts vary along two general dimensions: the proportion of savings that's shared with an ACO, and the requirement to pay shared losses to Medicare if savings is too low.

Table 4 shows where each contract falls. ACOs under Track 1 and levels A and B of the Basic track do not face downside risk. This contract structure, faced by over 90% of ACOs between 2012 and 2017, offers shared savings when the savings rate of the ACO is above the minimum savings rate. In the model's notation, this is when $S_j^* \geq \underline{S}_j$. Under every other contract option, there is two-sided risk, and ACOs are required to repay Medicare if their savings rate is below the symmetric minimum loss rate: $S_j^* \leq -\underline{S}_j$.

Figure 3 shows how the various contracts differ in power by graphing an ACO's earned shared savings or losses as a function of its savings rate, S_j. Under any risk model, an ACO earns shared savings when $S_j \geq \underline{S}_j$ and earns nothing when $S_j \in \left(-\underline{S}_j, \underline{S}_j\right)$. Two-sided ACOs (dotted line) typically earn a higher proportion of savings to encourage exposure to downside risk.

The estimation of the cost function and utility from participation uses only

Figure 3: **Risk Models**



Note: This figure shows the various contract options (or "Risk Models") offered by Medicare to ACOs. All ACOs earn shared savings when their savings rate S_j is above the minimum savings rate $\underline{S}_j$. Some ACOs pay shared losses when their savings rate S_j is below the minimum loss rate $-\underline{S}_j$.

Track 1 ACOs, where the shared savings earned by ACO j is

$$R_j(S_j, Q_j) = \begin{cases} F \cdot B_j S_j Q_j & \text{if } S_j \geq \underline{S}_j \text{ and } Q_j \geq \underline{Q} \\ 0 & \text{otherwise} \end{cases} \tag{40}$$

where $F = 0.5$. Other contract types are omitted from the estimation sample, but I can predict their behavior by altering the revenue function and using same cost function (which is invariant to contract changes). For the following predictions, two-sided ACOs have the shared savings formula

$$R_j^{TS}(S_j, Q_j) = \begin{cases} FQ_j \cdot B_j S_j & \text{if } S_j \geq \underline{S}_j \text{ and } Q_j \geq \underline{Q} \\ (1 - FQ_j) \cdot B_j S_j & \text{if } S_j \leq -\underline{S}_j \\ 0 & \text{otherwise} \end{cases} \tag{41}$$

where $F = 0.5$ for Track 1+ and Basic Track C, D, and E ACOs, $F = 0.6$ for Track 2 ACOs and $F = 0.75$ for Track 3 and Enhanced Track ACOs.[13]

Table 5 displays the simulation results. The table contains predictions of average ACO performance for one-sided and two-sided incentive structures and for varying proportions of savings shared with an ACO. The model's prediction for the estimation sample is in the middle-left cells, where $F = 0.50$ and payment is one-sided (italicized font). The row with $F = 0.50^d$ contains statistics from data. The model fits the data very well for average ACO savings rate, quality score, and total program savings. Predictions of the proportion of ACOs that save above or below the minimum savings and loss rates are less accurate: this occurs because some ACOs in data save just below $\underline{S}_j$, but the model predicts savings just above $\underline{S}_j$. This is likely a consequence of the equilibrium selection procedure.

The model predicts very large increases in average savings rates of ACOs under the two-sided model. For example, under Track 1, the equilibrium for some ACOs is to minimize cost at a savings rate below $-\underline{S}_j$. This is not optimal under Track 3 because they are penalized for doing so. Looking at columns "Prop. Qualified" and "Prop. Failed" in Table 5, we can see that out of 1486 observations, 609 (41%) qualify for shared savings under Track 1, and 684 (46%) under Track 3. Moreover, under Track 3, just 45 (3%) pay shared losses to CMS, compared to 550 (37%) Track 1 ACOs with a sharing rate less than the minimum loss rate.

Under both one-sided and two-sided incentives, quality scores almost always increase as F increases.[14] For a fixed F, however, ACOs facing one-sided

[13] For two-sided ACOs, the so-called "final loss rate" is defined as $1 - FQ_j$. It is bounded below at 0.4. for Track 3/Enhanced Track ACOs. For Track 2 ACOs, it's bounded above by 0.6; on Track 3/Enhanced Track, its bounded above by 0.75.

[14] The lone exception to this is when F changes from 0.6 to 0.75 under two-sided risk.

Table 5: **ACO Performance Predictions**

	One-Sided Risk Model				
F	S_j^*	Q_j^*	Prop. Qualified	Prop. Failed	Program Savings
0.25	-0.0088	0.8843	0.3351	0.3950	-0.1432
0.30	-0.0048	0.8867	0.3526	0.3869	-0.0301
0.40	0.0023	0.8920	0.3863	0.3769	0.1096
0.50	*0.0078*	*0.8953*	*0.4118*	*0.3668*	*0.1182*
0.50^d	*0.0084*	*0.8840*	*0.3201*	*0.2011*	*0.0919*
0.60	0.0137	0.8998	0.4388	0.3580	0.0354
0.75	0.0202	0.9046	0.4623	0.3520	-0.2984

	Two-Sided Risk Model				
F	S_j^*	Q_j^*	Prop. Qualified	Prop. Failed	Program Savings
0.25	0.0117	0.7167	0.3338	0.0168	0.3971
0.30	0.0150	0.7211	0.3546	0.0168	0.4883
0.40	0.0213	0.7293	0.3937	0.0168	0.5837
0.50	0.0317	0.7660	0.4468	0.0249	0.7408
0.60	0.0360	0.7987	0.4704	0.0511	0.5340
0.75	**0.0375**	**0.7620**	**0.4643**	**0.0289**	**0.0403**

These tables show model simulations for various ACO contract options. F is the proportion of savings shared with an ACO, S_j^* is average ACO savings rate, Q_j^* is average ACO quality score, Prop. Qualified is the proportion of ACOs that qualify for shared savings ($S_j^* \geq \underline{S}_j$), and Prop. Failed is the proportion of ACOs with savings rare below minimum loss rate ($S_j^* \leq -\underline{S}_j$). Program Savings is the total dollar amount saved in $ billions, defined in Equation 42. The superscript d indicates values observed in data. Italicized numbers are performance statistics under estimation sample. Bold numbers are the model's predictions for Track 3/Enhanced Track ACOs.

incentives have a significantly higher quality score than ACOs facing two-sided incentives. Since there is a large tradeoff between savings and quality (i.e., $\hat{\kappa}$ is very large), ACOs must choose a lower quality score to avoid paying shared losses to CMS.

The contract faced by ACOs plays a large impact on total savings to CMS. Their savings from the program changes since 1) ACOs have different savings rates, 2) the amount of subsidy paid for a given savings rate is different, and 3) CMS may recoup excess payment when ACOs perform poorly. Column "Program Savings" is the total money saved (or lost) over the benchmark expenditure, less the amount shared with ACOs. The values indicate we should expect the total savings to CMS to increase significantly were all ACOs under the contract design of Track 1+ or Basic Tracks C, D, and E, from \$118 million to \$740 million. Incentives are too strong, however, for Track 3 and the Enhanced Track. In spite of a much larger average savings rate, since so much of savings is paid to ACOs, overall program savings decreases by 66%.

5.2 Solving for Contracts between ACOs and Medicare

The results in the previous section indicate that the proportion of savings shared with an ACO and the presence of downside risk both play a large role in determining the success of the Medicare Shared Savings Program. To find the contract that maximizes savings of the MSSP, I compute the *savings-optimal*

Quality score decreases here because of the cap on the final loss rate $(1 - FQ_j)$. This is also the reason that the proportion of ACOs with $S_j^* \leq -\underline{S}_j$ becomes smaller over the same increment.

sharing rate F by solving the problem

$$\max_{F\in[0,1]} \sum_{j\in\mathcal{J}} \left\{ \overbrace{B_j S_j^*(F)}^{\$\text{ saved by ACO } j} - \overbrace{F \cdot B_j S_j^*(F) Q_j^*(F) \mathbf{1}\left\{S_j^*(F) \geq \underline{S}_j\right\} \mathbf{1}\left\{Q_j^*(F) \geq \underline{Q}\right\}}^{\$\text{ paid to ACO } j} \right\} \tag{42}$$

$$\text{s.t. } \left(s_{ij}^*(F), q_{ij}^*(F)\right) = \arg\max_{s_{ij}, q_{ij}} \pi_{ij}\left(s_{ij}, q_{ij}\right) \text{ for all } i \in I_j \text{ and } j \in \mathcal{J}.$$

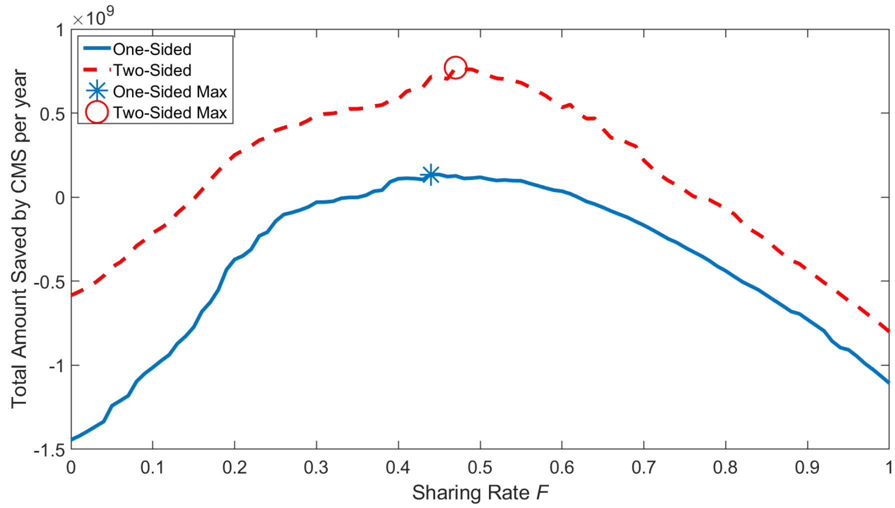
The objective function is the total amount of money saved by the Medicare Shared Savings Program. Note that an ACO's savings rate S_j^* and quality score Q_j^* are written as a function of the sharing rate F, since ACOs save more when F is higher. The tradeoff, of course, is that CMS only receives a fraction of what's saved from the benchmark. The constraint in the above problem is the incentive compatibility constraint, which states that given the contract imposed, ACO participants choose savings and quality efforts from a Nash equilibrium.

The solid line of Figure 4 plots the objective function of CMS when maximizing total savings with one-sided ACOs (Equation 42), and the dashed line plots the objective function of CMS with two-sided ACOs.[15] CMS saves the most money under a one-sided incentive scheme at $F^* = 0.44$. The amount saved is just \$16.6 million higher than under the estimation sample, where $F = 0.5$. If payment is two-sided, the optimal saving fraction is nearly the same at $F^* = 0.47$. This again implies incentives are too powerful under Track 3 and the Enhanced Track, where F is 0.75. Compared to these higher powered incentives, the amount saved at $F^* = 0.47$ is \$730 million larger.[16]

[15]CMS's objective for two-sided ACOs is slightly different than Equation 42 and includes an extra term for shared losses paid back to CMS.

[16]The sharing rate is higher for two-sided ACOs under current law in order to encourage ACOs to choose those Tracks—my analysis does not account for this choice.

Figure 4: **Savings-Optimal Sharing Rate**



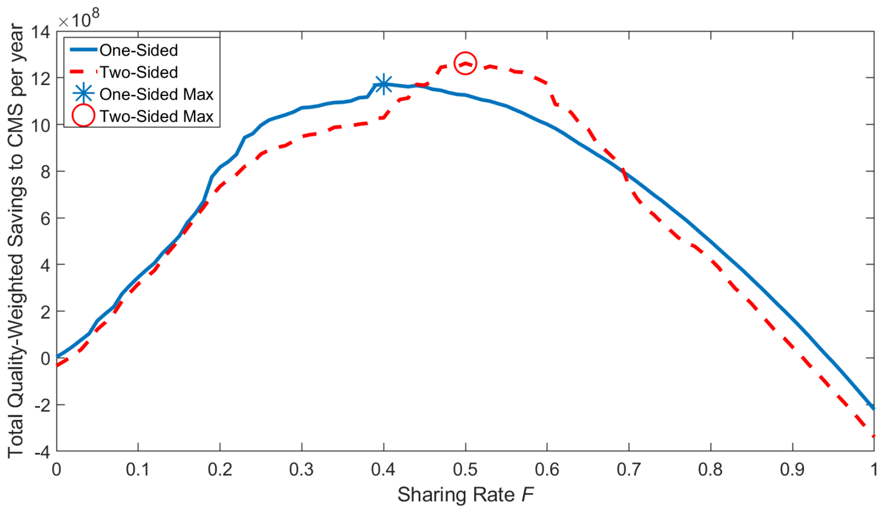
Note: This figure shows model simulations for the total amount of program savings for various sharing rates (horizontal axis) and with and without penalties for exceeding benchmark expenditure (solid vs. dashed line).

The objective function in Equation 42 is written such that the solution maximizes total program savings, so the solution is *savings-optimal*. Importantly, that objective is *decreasing* in the quality score of ACOs, since a higher quality score increases the amount paid to ACOs. To examine how this impacts the optimal sharing rate, I also compute the *savings-quality-optimal* sharing rate, where the objective is to maximize savings weighted by quality score. Formally, the problem is

$$\max_{F\in[0,1]} \sum_{j\in\mathcal{J}} \left\{ \left[B_j S_j^*(F) - F \cdot B_j S_j^*(F) Q_j^*(F) \mathbf{1}\left\{ S_j^*(F) \geq \underline{S}_j \right\} \mathbf{1}\left\{ Q_j^*(F) \geq \underline{Q} \right\} \right] Q_j^*(F) \right\} \tag{43}$$

s.t. $\left(s_{ij}^*(F), q_{ij}^*(F)\right) = \arg\max_{s_{ij},q_{ij}} \pi_{ij}\left(s_{ij}, q_{ij}\right)$ for all $i \in I_j$ and $j \in \mathcal{J}$.

Figure 5: **Savings-Quality-Optimal Sharing Rate**



Note: This figure shows model simulations of quality-weighted program savings for various sharing rates (horizontal axis) and with and without penalties for exceeding benchmark expenditure (solid vs. dashed line).

when $S_j^* \geq 0$ and

$$\max_{F \in [0,1]} \sum_{j \in \mathcal{J}} \left\{ \left[B_j S_j^*(F) - F \cdot B_j S_j^*(F) Q_j^*(F) \mathbf{1} \left\{ S_j^*(F) \geq \underline{S}_j \right\} \mathbf{1} \left\{ Q_j^*(F) \geq \underline{Q} \right\} \right] \left[1 - Q_j^*(F) \right] \right\} \tag{44}$$

$$\text{s.t.} \quad \left(s_{ij}^*(F), q_{ij}^*(F) \right) = \arg\max_{s_{ij}, q_{ij}} \pi_{ij} \left(s_{ij}, q_{ij} \right) \text{ for all } i \in I_j \text{ and } j \in \mathcal{J}.$$

when $S_j^* < 0$. The objective is weighted by $(1 - Q_j^*)$ when savings is negative.

Figure 5 plots the objective function of CMS when maximizing total savings weighted by quality score with one-sided ACOs (Equations 43 and 44) and the objective function of CMS with two-sided ACOs. The apparent dominance of the two-sided risk model disappears once we weight program savings by quality scores. In fact, outside of the interval $[0.44, 0.70]$, the one-sided risk model has higher quality-weighted savings than the two-sided risk model. The maximum value occurs at $F = 0.40$ for the one-sided risk model and $F = 0.50$

for the two-sided risk model. The two-sided risk model has an objective value just 7.6% higher at its maximum.

These counterfactual exercises offer strong evidence that the optimal sharing rate for the MSSP is between 0.4 and 0.5—very close to some current contract options. The push to two-sided incentive structures is well-founded if maximizing program savings is the objective, however, these savings come at the cost of quality of care. The model predicts that program savings will not increase as ACOs shift to higher powered incentives, because the increase in savings is wiped out by the additional incentive pay given to ACOs.

5.3 Performance Loss due to Free-Riding within ACOs

In this section, I consider the problem where a governing body with complete control over ACO participant behavior chooses participant savings and quality in order to maximize the total profit of all participants in an ACO. The maximization problem is

$$\max_{\mathbf{s}_j, \mathbf{q}_j} R_j\left(S_j, Q_j\right) - \sum_{i \in I_j} c_{ij}\left(s_{ij}, q_{ij}\right). \tag{45}$$

The difference between this problem and the game played by participants is that cost is now shared between participants: agents with low margins that operate at a loss are compensated by those with high margins. I solve this for every ACO, and present the means in Table 6. Under perfect cooperation, average ACO savings rate increases by nearly four percentage points, or about one standard deviation. Quality scores increase by just 0.02, or 0.22 standard deviations. Table 6 also indicates that as coordination increases, the incentives imposed by Medicare should be weakened. While 44% of savings is

Table 6: **Performance Loss from Non-Cooperative Behavior**

	Strategic Behavior			Perfect Coordination		
F	S_j^*	Q_j^*	Program Savings	S_j^*	Q_j^*	Program Savings
0.25	-0.0088	0.8843	-0.1432	-0.0088	0.8843	-0.1432
0.30	-0.0048	0.8867	-0.0301	0.0299	0.9098	1.6643
0.40	0.0023	0.8920	0.1096	0.0392	0.9168	1.6687
0.50	0.0078	0.8953	0.1182	0.0466	0.9239	1.4698
0.60	0.0137	0.8998	0.0354	0.0540	0.9286	1.2576
0.75	0.0202	0.9046	-0.2984	0.0623	0.9360	0.6744

This table shows model simulations for various ACO contract options. F is the proportion of savings shared with an ACO, S_j^* is average ACO savings rate, and Q_j^* is average ACO quality score. Program Savings is the total dollar amount saved in $ billions, defined in Equation 42.

the optimal amount to share with free-riding, under perfect coordination the optimal amount is 35%.

6 Conclusion

Incentive design is used by firms, governments, households, educators, and many others to achieve a variety of ends. Accordingly, designing effective incentives is a popular topic of study in all fields of economics. Incentive design is particularly important in the United States healthcare sector, where physician incentive programs and pay-for-performance initiatives impact the quality of life and spending of individuals in 3.5 trillion dollar industry. In this paper, I investigate the empirical role of multitasking in the context of the Medicare Shared Savings Program and Accountable Care Organizations in order to design contracts that maximize the money saved by the incentive program while accounting for free-riding of healthcare providers, the savings-quality tradeoff, and voluntary participation.

In this paper, I develop and estimate a structural model of performance in ACOs. I find that providers face a large tradeoff between increasing savings and increasing quality of care. Counterfactual policy analysis shows that if ACOs are required to pay penalties to Medicare for spending too much, savings increases drastically, though quality falls. The optimal proportion of savings to share with an ACO (both when and when not weighting by quality score) falls between 0.4 and 0.5. Another counterfactual shows performance improves significantly were ACOs able to perfectly coordinate, and over $1 billion per year is lost to free-riding.

References

Aswani, A., Shen, Z.-J. M., & Siddiq, A. 2019. Data-driven incentive design in the medicare shared savings program. *Operations Research*, *67*(4), 1002-1026.

Bai, C.-E., & Xu, L. C. 2005. Incentives for ceos with multitasks: Evidence from chinese state-owned enterprises. *Journal of Comparative Economics*, *33*(3), 517 - 539. (Symposium: Monetary and Exchange Rate Policies in the CIS Countries: Between the EU and Russia)

Chandra, A., Cutler, D., & Song, Z. 2011. Chapter six - who ordered that? the economics of treatment choices in medical care. In M. V. Pauly, T. G. Mcguire, & P. P. Barros (Eds.), *Handbook of health economics* (Vol. 2, p. 397 - 432). Elsevier.

Choné, P., & Ma, A. 2011. Optimal health care contract under physician agency. *Annals of Economics and Statistics*(101/102), 229–256.

Dumont, E., Fortin, B., Jacquemet, N., & Shearer, B. 2008. Physicians' multitasking and incentives: Empirical evidence from a natural experiment. *Journal of Health Economics*, *27*(6), 1436 - 1450.

Einav, L., Finkelstein, A., & Mahoney, N. 2018. Provider incentives and healthcare costs: Evidence from long-term care hospitals. *Econometrica*, *86*(6), 2161-2219.

Eliason, P. J., Grieco, P. L. E., McDevitt, R. C., & Roberts, J. W. 2018, November. Strategic patient discharge: The case of long-term care hospitals. *American Economic Review*, *108*(11), 3232-65.

Encinosa, W. E., Gaynor, M., & Rebitzer, J. B. 2007. The sociology of groups and the economics of incentives: Theory and evidence on compensation systems. *Journal of Economic Behavior & Organization*, *62*(2), 187 - 214.

Feng Lu, S. 2012. Multitasking, information disclosure, and product quality: Evidence from nursing homes. *Journal of Economics & Management Strategy*, *21*(3), 673-705.

Foo, P. K., Lee, R. S., & Fong, K. 2017. Physician prices, hospital prices, and treatment choice in labor and delivery. *American Journal of Health Economics*, *3*(3), 422-453.

Frandsen, B., Powell, M., & Rebitzer, J. B. 2017. Sticking points: Common-agency problems and contracting in the u.s. healthcare system..

Frandsen, B., & Rebitzer, J. B. 2015. Structuring incentives within accountable care organizations. *Journal of Law, Economics, and Organization*, *31*(suppl. 1), 77-103.

Frech, H. E., Whaley, C., Handel, B. R., Bowers, L., Simon, C. J., & Scheffler, R. M. 2015, Sep 01. Market power, transactions costs, and the entry of accountable care organizations in health care. *Review of Industrial Organization*, *47*(2), 167–193.

Gaynor, M. 2006, June. *What do we know about competition and quality in health care markets?* (Working Paper No. 12301). National Bureau of Economic Research.

Gaynor, M., Ho, K., & Town, R. J. 2015, June. The Industrial Organization of Health-Care Markets. *Journal of Economic Literature*, *53*(2), 235-284.

Gaynor, M., Rebitzer, J., & Taylor, L. 2004. Physician incentives in health maintenance organizations. *Journal of Political Economy*, *112*(4), 915–931.

Glied, S., & Hong, K. 2018. Health care in a multi-payer system: Spillovers of health care service demand among adults under 65 on utilization and outcomes in medicare. *Journal of Health Economics*, *60*, 165 - 176.

Grassi, S., & Ma, A. 2016. Information acquisition, referral, and organization. *The RAND Journal of Economics*, *47*(4), 935-960.

Hackmann, M. B. 2019, May. Incentivizing better quality of care: The role of medicaid and competition in the nursing home industry. *American Economic Review*, *109*(5), 1684-1716.

Hansen, L. P. 1982. Large sample properties of generalized method of moments estimators. *Econometrica*, *50*(4), 1029–1054.

Ho, K., & Pakes, A. 2014. Hospital choices, hospital prices, and financial incentives to physicians. *The American Economic Review*, *104*(12), 3841–3884.

Hong, F., Hossain, T., List, J. A., & Tanaka, M. 2018. Testing the theory of multitasking: Evidence from a natural field experiment in chinese factories. *International Economic Review*, *59*(2), 511-536.

Kim, M., Sudhir, K., & Uetake, K. 2019, September. *A structural model of a multitasking salesforce: Job task allocation and incentive plan design* (Discussion Paper No. 2199). Cowles Foundation.

McWilliams, J. M. 2014, Oct 01. Aco payment models and the path to accountability. *Journal of General Internal Medicine*, *29*(10), 1328–1330.

McWilliams, J. M., Hatfield, L. A., Landon, B. E., Hamed, P., & Chernew, M. E. 2018. Medicare spending after 3 years of the medicare shared savings program. *New England Journal of Medicine*, *379*(12), 1139-1149.

Mullen, K. J., Frank, R. G., & Rosenthal, M. B. 2010. Can you get what you pay for? pay-for-performance and the quality of healthcare providers. *The RAND Journal of Economics*, *41*(1), 64-91.

Rebitzer, J. B., & Votruba, M. E. 2011, October. *Organizational economics and physician practices* (Working Paper No. 17535). National Bureau of Economic Research.

Slade, M. E. 1996. Multitask agency and contract choice: An empirical exploration. *International Economic Review*, *37*(2), 465–486.

For Online Publication

A Institutional Details

This appendix gives detailed background information on the Medicare Shared Savings Program. ACOs began operating in the MSSP in 2012. Nearly any Medicare provider, including individual physicians, group practices, and large hospital systems, can start an ACO and recruit other Medicare providers to participate in their joint venture.[17] Once an ACO shows they have established a governing board that oversees clinical and administrative aspects of operation and shows the presence of formal contracts between itself and its member participants (including the distribution of any earned incentive pay), it then enters into a five year agreement with the Centers for Medicare and Medicaid Services (CMS).[18] Medicare fee-for-service (FFS) beneficiaries are assigned to ACOs by CMS: if a given Medicare beneficiary receives the plurality of primary care services from a primary care provider who is (or is employed by) an ACO participant, that beneficiary is assigned to that participant's ACO.[19]

There are two separate components of assessing ACO performance, and both determine the amount ACOs are paid. The first is an overall quality score, which is a composite score between 0 and 1 of several sub-measures of care quality. These sub-measures fall into the domains of "Patient/Caregiver

[17] A participant can be nearly any health care provider that accepts and bills Medicare. Participants are legally defined by their Tax ID Number (TIN) or CMS Certification number (CCN).

[18] Before July 2019, agreements lasted three years.

[19] When a Medicare beneficiary receives the plurality of primary care services from a primary care provider not associated with an ACO, they are not assigned to an ACO. This assignment methodology results in roughly one fifth to one third of all FFS beneficiaries assigned to ACOs each year. An ACO must be assigned at least 5000 beneficiaries to operate and earn shared savings payments.

Experience," "Care Coordination/Patient Safety," "Preventative Health," and "At-Risk Population." Some sub-measures are survey responses (e.g., "ACO2: How Well Your Doctors Communicate"), while others are computed from Medicare Claims and aggregated to the ACO-level (e.g., "ACO21: Proportion of Adults who had blood pressure screened in past 2 years").[20]

The second component is ACO savings. CMS first establishes an ACO's benchmark expenditure by forecasting per-beneficiary Medicare expenditure for beneficiaries that would have been assigned to the ACO in the three years prior to the agreement period. For performance years after the first, the benchmark is updated based on projected growth of per-beneficiary Medicare expenditure.[21] The savings rate of an ACO in a performance year is then the difference between its benchmark expenditure and the actual expenditure on assigned beneficiaries divided by its benchmark expenditure.

A.1 ACO payment from 2012 until June 2019

For the first six performance years of the MSSP, ACOs had a choice between four payment contracts called "Tracks." The contracts vary in power and exposure to downside risk. Track 1, available to ACOs only in their first six years of operation, is lowered powered and requires no loss sharing with CMS (i.e., it's one-sided). Accordingly, each performance year the shared savings paid by CMS to an ACO on Track 1 is

$$\frac{1}{2} \cdot (\text{Benchmark Expenditure} - \text{Expenditure}) \cdot \text{Quality Score} \tag{46}$$

[20]See `https://go.cms.gov/2xHy7Uo` for a full list of ACO quality scores for every performance year.

[21]Regional adjustments to benchmarks were introduced in 2017 for ACOs in their fourth year of operation.

when an ACO's savings rate meets or exceeds its minimum savings rate and its quality score meets or exceeds quality reporting standards. Otherwise, an ACO earns \$0 in shared savings. For example, consider an ACO with a benchmark expenditure of \$186 million (the average over 2012-2017) and a minimum savings rate of 0.02. If that ACO has an expenditure of \$160 million with a quality score of 0.90, it would earn

$$\frac{1}{2} \cdot (\$186 \text{ million} - \$160 \text{ million}) \cdot 0.90 = \$11.7 \text{ million} \tag{47}$$

in shared savings. Its savings rate is $(18.6 - 16)/16 = 0.1625$, and hence the minimum savings rate is exceeded. Though paying a subsidy, Medicare saves money as well: on net, it saves \$14.3 million, as it paid \$11.7 million to save \$26 million.

Like Track 1, Track 1+ offers ACOs up to 50% of savings as incentive pay. It differs by introducing downside risk, requiring ACOs to pay 30% of losses to Medicare if expenditure is much larger than benchmark expenditure and savings is below the minimum loss rate. Track 2 and Track 3 ACOs face both higher powered incentives and downside risk. Track 2 and Track 3 give 60% and 75% of savings back to ACOs, respectively. If savings is below the minimum loss rate, these ACOs must pay money back to Medicare at a rate of $\left(1 - \frac{3}{5} \cdot \text{Quality Score}\right) \cdot 100\%$ and $\left(1 - \frac{3}{4} \cdot \text{Quality Score}\right) \cdot 100\%$ of losses for Tracks 2 and 3, respectively.

Track 1 has been the overwhelming contract choice of ACOs. In 2013, 2014, and 2015, between 97% and 99% of the 200-400 operating ACOs chose Track 1. In 2016 and 2017, 95% and 92% of the 432 and 472 ACOs operating that year chose Track 1.

A.2 ACO payment from July 2019 until the present

The first six years of the Medicare Shared Savings Program produced modest decreases in Medicare expenditure (McWilliams et al., 2018). In an attempt to improve ACO performance, CMS made several changes to the MSSP with its final rule named “Pathways to Success” (or “Pathways”).

Changes in Pathways pertinent to this paper regard the contracts between ACOs and Medicare. Tracks 1, 1+, 2, and 3 are replaced with two Tracks: “Basic” and “Enhanced.” Under the Basic Track, there are five levels, “A” through “E.” Under levels A and B, ACOs earn up to 40% of savings and do not pay shared losses if expenditure exceeds benchmark expenditure. Under levels C, D, and E, ACOs earn up to 50% of savings and pay an increasing amount of shared losses if expenditure exceeds benchmark expenditure. An ACO is automatically advanced one level (e.g., from level A to B) after each performance year. The Enhanced Track is equivalent to Track 3.

Various other changes were made to the MSSP in Pathways, including beneficiary assignment methodology, benchmark calculation, and assigning new ACO classifications (“low-renvenue” and “experienced”) that impact the payment contracts available to an ACO.

In this paper, counterfactual predictions consider two dimensions of contracts: the fraction of savings shared with an ACO and the presence of downside risk. These dimensions broadly account for all previous (Tracks 1, 1+, 2, and 3) and current (Basic and Enhanced Tracks) contract options.

B Existence of Equilibrium

The following proposition establishes existence of equilibrium.

Proposition B.1. *Let* $\frac{\partial^2 c_{ij}}{\partial s_{ij}^2} \cdot \frac{\partial^2 c_{ij}}{\partial q_{ij}^2} \geq \left(\frac{\partial^2 c_{ij}}{\partial s_{ij} \partial q_{ij}} - \frac{w_{ij}^3}{2} B_j \right)^2$*. Then, there is a Nash equilibrium in pure strategies.*

Proof. First, the assumption that $\frac{\partial c_{ij}}{\partial s_{ij}} \cdot \frac{\partial c_{ij}}{\partial q_{ij}} \geq \left(\frac{\partial^2 c_{ij}}{\partial s_{ij} \partial q_{ij}} - \frac{w_{ij}^3}{2} B_j \right)^2$ and that c_{ij} is strictly convex guarantees that there is a unique solution to both of the problems (fixing $\mathbf{s}_{-ij}$ and $\mathbf{q}_{-ij}$)

$$\underbrace{\begin{array}{ll} \max\limits_{\substack{s_{ij} \in [-1,1] \\ q_{ij} \in [0,1]}} & \pi_{ij}^{Q}(\mathbf{s}_j, \mathbf{q}_j) \end{array}}_{\text{Problem A}} \qquad \underbrace{\begin{array}{ll} \min\limits_{\substack{s_{ij} \in [-1,1] \\ q_{ij} \in [0,1]}} & c_{ij}\left(s_{ij}, q_{ij}\right) \end{array}}_{\text{Problem B}}$$

for all $i \in I_j$, where

$$\pi_{ij}^{Q}(\mathbf{s}_j, \mathbf{q}_j) = 0.5 \cdot w_{ij} B_j \cdot S_j \cdot Q_j - c_{ij}\left(s_{ij}, q_{ij}\right) \tag{48}$$

In any equilibrium, every participant is solving Problem A, or every participant is solving Problem B. Otherwise, there would be at least one participant not maximizing π_{ij}. Let $\left(\mathbf{s}_j^B, \mathbf{q}_j^B\right)$ be the a tuple of vectors such that the elements of the vectors solve Problem B for all $i \in I_j$, and similarly define $\left(\mathbf{s}_j^A, \mathbf{q}_j^A\right)$. I will show that equilibrium exists, and it is always one of these tuples.

$\left(\mathbf{s}_j^B, \mathbf{q}_j^B\right)$ is an equilibrium when there is no $i \in I_j$ such that i is better off choosing $\left(s_{ij}^A, q_{ij}^A\right)$ while others choose $\left(\mathbf{s}_{-ij}^B, \mathbf{q}_{-ij}^B\right)$. In other words, the cost-minimizing equilibrium exists when no participant is so influential (high w_{ij}) with low enough marginal costs such that it's still optimal for that participant to push the entire ACO to earn shared savings.

Suppose there is a participant with such characteristics, and $\left(\mathbf{s}_j^B, \mathbf{q}_j^B\right)$ is not an equilibrium. What's left to establish is that there is at least one such $\left(\mathbf{s}_j^A, \mathbf{q}_j^A\right)$ that is an equilibrium. This is trivial, however, because by assuming

$\frac{\partial^2 c_{ij}}{\partial s_{ij}^2} \cdot \frac{\partial^2 c_{ij}}{\partial q_{ij}^2} \geq \left(\frac{\partial^2 c_{ij}}{\partial s_{ij} \partial q_{ij}} - \frac{w_{ij}^3}{2} B_j \right)^2$, and since the strategy space of each participant a closed, bounded, and convex set, then this is a concave game. Concave games have unique Nash equilibria, so $\left(\mathbf{s}_j^A, \mathbf{q}_j^A\right)$ exists. □

C Influence weights w_{ij}

I've defined influence weights $\{w_{ij}\}_{i \in I_j}$ such that

$$\sum_{i \in I_j} w_{ij} s_{ij} = S_j \qquad \sum_{i \in I_j} w_{ij} q_{ij} = Q_j \tag{49}$$

where $\sum_{i \in I_j} w_{ij} = 1$. Note that for participant savings efforts s_{ij} have to have a definition analogous to that of S_j, we would have

$$S_j = \frac{BE_j - E_j}{BE_j} = \frac{\sum_{i \in I_j} BE_{ij} - \sum_{i \in I_j} E_{ij}}{\sum_{i \in I_j} BE_{ij}} = \sum_{i \in I_j} w_{ij} \frac{BE_{ij} - E_{ij}}{BE_{ij}} = \sum_{i \in I_j} w_{ij} s_{ij} \tag{50}$$

where BE_j and E_j are the benchmark expenditure and expenditure of ACO j (both real quantities observed in data) and BE_{ij} and E_{ij} are the benchmark expenditure and expenditure of participant i in ACO j (both theoretical quantities). Thus, a definition of w_{ij} consistent with the above is $w_{ij} = \frac{BE_{ij}}{BE_j}$, or simply participant i's share of ACO benchmark expenditure. Intuitively, this means that a very influential participant i in ACO j will have a relatively large share of expected expenditure on assigned beneficiaries.

In data, I measure w_{ij} as shares of expenditure for each *type* of provider within an ACO. To be specific, suppose provider i has type k. Then,

$$w_{ij} = \frac{\text{Total Spending by type } k}{(\text{Total \# of } i \text{ with type } k) \times (\sum_\ell \text{Total Spending by type } \ell)} \tag{51}$$

The numerator and both terms in the denominator are observed for the general types k.

This measure of w_{ij} has two important requirements. First, it requires that providers of the same type have similar shares of overall expenditure within an ACO. This is likely the case, since ACOs tend to be predominantly hospital based or group practice based. Second, this measure requires that the *ratio* BE_{ij}/BE_j is close to the ratio E_{ij}/E_j, since w_{ij} as defined in Equation 51 is the latter ratio.

D Elements of X_j

See Table 7 for elements of X_j.

E Uncertainty in Savings and Quality

In the second stage of this model, participating Medicare providers in ACOs choose their own savings and quality efforts to overall ACO performance, though the mapping from participant choices to overall performance is deterministic. To check the robustness of this paper's results with respect to the assumption of certainty, this section briefly discusses a model and estimation where uncertainty is included. This model is a slight generalization of Frandsen & Rebitzer (2015), since I allow for heterogeneous participants and payment functions that depend on quality score.

Table 7: **Elements of X_j**

Abbreviated Variable Name	**Description**
# states	Number of states where beneficiaries assigned to the ACO reside.
# beneficiaries	Number of beneficiaries assigned to the ACO in thousands.
average risk score	Average CMS HCC risk score of aged, non-dual beneficiaries assigned to the ACO.
% over 75	Percent of assigned beneficiaries over age 75.
% male	Percent of assigned beneficiaries that are male.
% nonwhite	Percent of assigned beneficaries that are non-white.
# providers	Total number of individual providers in an ACO in thousands.
fraction PCP	Proportion of individual providers that are primary care physicians.
fraction inpatient	Proportion of expenditures that are inpatient expenditures (includes short term, long term, rehabilitation, and psychiatric).
fraction outpatient	Proportion of expenditures that are outpatient expenditures.
# PC services	Total number of primary care services in thousands.
# admissions	Total number of inpatient hospital discharges in thousands.
fraction PC served by PCP	Proportion of primary care services provided by primary care physician.
all group	Indicates every participant in ACO is a group practice or hospital.

This table shows control variables used for estimation of marginal cost parameters. Not listed: Constant term, year and census division fixed effects.

Define s_{ij}, q_{ij}, S_j, and Q_j as before, except that *realized* efforts of participants are i.i.d. random variables

$$\hat{s}_{ij} \sim N\left(s_{ij}, \sigma_S^2\right) \qquad \hat{q}_{ij} \sim N\left(q_j, \sigma_Q^2\right) \tag{52}$$

where $N(\cdot)$ is the normal distribution. Defining $\hat{S}_j = \sum_{i \in I_j} w_{ij} \hat{s}_{ij}$ and $\hat{Q}_j = \sum_{i \in I_j} w_{ij} \hat{q}_{ij}$, each participant $i \in I_j$ solves the expected profit maximization problem

$$\max_{s_{ij}, q_{ij}} \mathbb{E}\left[R_{ij}\left(\hat{S}_j, \hat{Q}_j\right)\right] - c\left(s_{ij}, q_{ij}; x_{ij}, \boldsymbol{\theta}\right) \tag{53}$$

where $R_{ij}(\hat{S}_j, \hat{Q}_j)$ is the per-participant shared savings earned by an ACO with savings $\hat{S}_j$ and quality score $\hat{Q}_j$ (defined in Section **??**). The objective function in Equation 53 becomes

$$E_\Pi^j(s_{ij}, q_{ij}, S_j, Q_j) = 0.5 \cdot w_{ij} B_j \cdot E_S(S_j) \cdot E_Q(Q_j) - c_{ij}\left(s_{ij}, q_{ij}\right) \tag{54}$$

where

$$E_S(S_j) = \mathbb{E}\left[\hat{S}_j \mathbf{1}\left\{\hat{S}_j \geq \underline{S}_j\right\}\right] = S_j \Phi\left(\frac{S_j - \underline{S}_j}{\sqrt{W_j^{(2)}}\sigma_S}\right) + \sqrt{W_j^{(2)}}\sigma_S \phi\left(\frac{S_j - \underline{S}_j}{\sqrt{W_j^{(2)}}\sigma_S}\right) \tag{55}$$

$$E_Q(Q_j) = \mathbb{E}\left[\hat{Q}_j \mathbf{1}\left\{\hat{Q}_j \geq \underline{Q}\right\}\right] = Q_j \Phi\left(\frac{Q_j - \underline{Q}}{\sqrt{W_j^{(2)}}\sigma_Q}\right) + \sqrt{W_j^{(2)}}\sigma_Q \phi\left(\frac{Q_j - \underline{Q}}{\sqrt{W_j^{(2)}}\sigma_Q}\right). \tag{56}$$

and $W_j^{(2)} = \sum_{i \in I_j} w_{ij}^2$ (see Appendix C). The functions ϕ and Φ are the standard normal probability and cumulative density functions, respectively, and $\mathbf{1}\{\cdot\}$ is the indicator function that takes a value of one if the statement in the brackets is true and zero otherwise.

E.1 Existence of Equilibrium

Proposition E.1. *Consider the simultaneous move game played by participants i in ACO j. If $D^2 E_{\Pi}^{j}$ is negative semidefinite, then there exists a Nash equilibrium in pure strategies. This equilibrium is unique.*

Proof. If the Hessian matrix $D^2 E_{\Pi}$ is negative semidefinite, then each participant i has a unique pair $\left(s_{ij}^*, q_{ij}^*\right)$ that maximizes $E_{\Pi}(\cdot)$ given values of s_{-ij} and q_{-ij}. Note it is possible that $\left|\frac{\partial c}{\partial q_{ij}}\right|$ is large enough that a corner solution for q_{ij}^* occurs.

What's left to determine is if the values $\left\{\left(s_{ij}^*, q_{ij}^*\right)\right\}_{i \in I_j}$ constitute a Nash equilibrium. This is obvious—any choice of participants must satisfy their FOCs (or corner solution). Given s_{ij}^* and q_{ij}^* are the best responses to S_j^* and Q_j^*, any deviation would suboptimal. Hence, equilibrium exists, and it is unique. □

E.2 Identification and Estimation

Identification and estimation of $\boldsymbol{\theta}$ in this model (with uncertainty) is nearly identical to their identification and estimation outlined in Section 3 for the model without uncertainty. There are two additional parameters to estimate, σ_S and σ_Q. These parameters are identified by variation in $W_j^{(2)}$ *or* if c has linear marginal cost in savings and quality.

E.3 Results

Table 8 shows the estimates of parameters in $\boldsymbol{\theta}$ that describe the shape of

Table 8: **Cost Function Parameter Estimates (Uncertainty Model)**

$c(s,q) = \frac{\delta_S}{2}s^2 + \frac{\delta_Q}{2}q^2 + \gamma_S s + \gamma_Q q + \kappa sq$

Model	Coef.	Estimate	Std. Err.	P-value	95% CI	
Baseline	δ_S	271.130	37.115	0.000	216.230	337.640
	δ_Q	1.693	0.417	0.000	0.997	2.373
	κ	15.533	6.049	0.010	3.620	23.680
w/ Uncertainty	δ_S	353.940	47.718	0.000	260.910	418.170
	δ_Q	1.591	0.565	0.005	0.970	2.324
	κ	21.489	6.248	0.001	3.693	24.086
	σ_S	0.011	0.013	0.370	0.000	0.021
	σ_Q	0.010	0.004	0.023	0.002	0.014
N		1486				

Standard errors, p-values, and CIs are from bootstrapping with 1000 rep. Estimates include year and Census Division FE. δ_S, δ_Q, and κ are scaled estimates.

the cost function as well as $\hat{\sigma}_S$ and $\hat{\sigma}_Q$. The parameters estimated from the model with uncertainty are well within a reasonable range of the parameters estimated from the model without uncertainty, albeit some with less precision. The estimate of σ_S is very imprecise, while σ_Q is estimated with some precision.