Essays on the Economics of Health Care

Kole Reddig

Tepper School of Business, Carnegie Mellon University

April 13, 2020

Abstract

This dissertation has two chapters on health care provider decision-making. The first chapter, titled "Designing Contracts for Multitasking Groups: A Structural Model of Accountable Care Organizations," estimates a structural model of multitasking agents to investigate the cost-quality tradeoff in healthcare and design contracts for a large physician incentive program. The setting involves Medicare's Accountable Care Organizations (ACOs), which are groups of healthcare providers that receive incentive pay for spending below a cost target on shared patients. I incorporate three important aspects of this setting into the model: (i) healthcare providers make multitasking effort choices concerning quality of care and cost reduction; (ii) providers in the incentive program are paid based on group performance, so they act strategically and may free-ride; and (iii) a provider's decision to participate in the incentive program depends on anticipated earnings. By estimating the model, I identify the tradeoff between quality of care and reducing cost, and I show that multitasking plays a large role in determining agent decisions. Counterfactual simulations indicate the contract that maximizes only 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. Another counterfactual shows free-riding within ACOs decreases program savings by over $1 billion per year.

The second chapter, titled "Spillovers between Medicare and Medicaid: Evidence from the Supply-Side and Payment Parity," studies the effect of a large increase in Medicaid reimbursement rates on the volume and type of services physicians provide to Medicare beneficiaries. I find that in response to the Medicaid "fee bump," physicians that qualified for increased Medicaid fees decreased the number of Medicare beneficiaries they served by 0.3 percent. This spillover, however, was not uniform among Medicare beneficiaries: provision of services designated for established Medicare patients decreased by 7.2 percent, yet provision of services designated for new Medicare patients increased by 1.1 percent. These results are consistent with the predictions of a mixed-economy model of physician decision-making, and they indicate that while the Medicaid fee bump decreased service provision to some Medicare beneficiaries, it also facilitated increased service provision to others by decreasing the marginal cost of care.

Acknowledgments

I thank my dissertation committee members, Martin Gaynor, Maryam Saeedi, Rebecca Lessem, and Akshaya Jha. I would also like to thank Chris Cronin, Dennis Epple, Albert Ma, Hakkı Özdenören, and Amanda Starc. This dissertation has benefited from presentations at the CMU-Tepper Applied Microeconomics Lunch Seminar, the South Eastern Health Economics Study Group Conference, the University of Michigan H2D2 Research Day, the International Industrial Organization Conference, the American Society of Health Economists Conference, ITAM Business School, Analysis Group (Boston), the Federal Deposit Insurance Corporation, the Congressional Budget Office, the Department of Justice, The Brattle Group (Washington, D.C.), West Virginia University, and American University.

I am grateful for the help provided by Lawrence Rapp and Laila Lee. I thank my parents, Kory and Lisa Reddig, and my sisters, Carly and Cassie Reddig, for providing encouragement, support, and love, and I thank Dusty for constant companionship. Finally, I thank Renée Corbett for inspiring me to be the best economist, partner, and human that I can be.

Contents

List of Figures

List of Tables

Chapter 1

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

1.1 Introduction

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. Complicating the setting further, providers join ACOs voluntarily, so if anticipated earnings from the program are underwhelming, providers will not participate.

What role does multitasking play in the decisions of Medicare providers? How much is lost to free-riding in ACOs? How many providers will drop out of the program if incentive pay is decreased? Answers to these questions are central the design of the MSSP, and furthermore will inform incentive and contract design throughout the healthcare sector. I answer these questions by building and estimating a structural model of Medicare providers in ACOs. Motivated by reduced form empirical exercises, I model providers' decisions regarding ACOs in two stages: participation and performance. In the first stage, Medicare providers choose which, if any, ACO to join, taking into account the income they expect to earn from joining

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

each ACO. In the second stage, 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 multitasking choices and a crucial component to computing contracts that make a combination of monetary savings and quality of care the objective. I use techniques from Berry (1994) to estimate parameters describing provider utility from participating in an ACO using aggregate participation data and accounting for unobserved heterogeneity.

I find that a $100,000 decrease in an ACO's income decreases the number of Medicare providers participating in that ACO by approximately 3 from a mean of 34. In other words, a 1 percent decrease in an ACO's income decreases participation in that ACO by 0.5 percent. There is 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; Jacob, 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; Ho & Lee, 2017; 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), in the field of operations research, also 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 agents or free-riding within ACOs.

This paper continues as follows: Section 2.4 gives a brief overview of the MSSP and ACOs, including descriptive ACO statistics and motivating regression analysis. I outline my model of participation and performance in ACOs in Section 1.3. I describe identification and estimation of model primitives in Section 1.4, and estimation results are in Section 1.5. I present counterfactual analysis, including computation of savings-maximizing contracts between ACOs and Medicare, in Section 1.6, and Section 1.7 concludes.

1.2 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]

Table 1.1 displays statistics describing ACO participants and beneficiaries assigned to ACOs.[3] There is substantial heterogeneity in the number of providers that join an ACO—

[2]For the interested reader, Appendix A.1 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 1.1: **Summary ACO Statistics: Providers and Beneficiaries**

Variable	**Mean**	**S.D.**	**Min.**	**Med.**	**Max.**
Number of participants[a]	37.79	58.03	1.00	20.00	840.00
Hospital led[a]	0.40	0.49	0.00	0.00	1.00
Physician led[a]	0.20	0.40	0.00	0.00	1.00
Mixed leadership[a]	0.41	0.49	0.00	0.00	1.00
Total number of individual providers (1000s)	0.59	0.84	0.00	0.28	7.28
Proportion of providers PCP	0.41	0.18	0.03	0.36	1.00
Proportion of providers specialist	0.41	0.21	0.00	0.44	0.88
Number of states where beneficiaries reside	1.52	0.97	1.00	1.00	10.00
Number of assigned beneficiaries (1000s)	17.78	17.49	0.15	11.87	149.63
Average risk score	1.06	0.11	0.81	1.04	2.09
Percent of beneficiaries over age 75	39.10	6.04	13.20	39.18	66.25
Percent of beneficiaries male	42.66	2.06	34.57	42.69	57.50
Percent of beneficiaries nonwhite	16.88	15.34	1.49	12.45	94.98
Year	**2012/13**	**2014**	**2015**	**2016**	**2017**
Number of ACOs	220	338	404	433	480

$N = 1849$. This table shows summary statistics for ACOs for years 2013-2017. The superscript ^{a} indicates statistics are for 2014-2017 (due to data availability).

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. "Hospital led," "physician led," and "mixed leadership," are variables indicating the predominant type of participant in an ACO, derived from MSSP ACO Participant List datasets. 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 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

[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.

and distributes to its members the amount

$$\text{Sharing Rate} \cdot \text{Quality Score} \cdot (\text{Benchmark Expenditure} - \text{Expenditure}) \tag{1.1}$$

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

$$0.5 \cdot 0.9 \cdot (\$186 \text{ million} - \$180 \text{ million}) = \$2.7 \text{ million} \tag{1.2}$$

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.

Table 1.2 contains statistics on ACO performance. The first three variables indicate the type of contract an ACO has with Medicare. The second block of variables describe ACO performance. ACO benchmark expenditures and realized expenditures are large: the mean is approximately \$190 million, with several ACOs having expenditure over \$1 billion.

Table 1.2: **Summary ACO Statistics: Savings and Quality**

Variable	Mean	S.D.	Min.	Med.	Max.
Sharing rate 50%, one-sided	0.96	0.19	0.00	1.00	1.00
Sharing rate 60%, two-sided	0.01	0.11	0.00	0.00	1.00
Sharing rate 75%, two-sided	0.03	0.16	0.00	0.00	1.00
Benchmark Expenditure ($ billions)	0.19	0.18	0.00	0.13	1.97
Expenditure ($ billions)	0.19	0.18	0.00	0.12	1.97
Benchmark Expenditure - Expenditure ($ millions)	1.46	10.18	-72.49	0.67	89.13
Savings Rate	0.01	0.05	-0.44	0.01	0.30
Quality Score	0.87	0.12	0.07	0.90	1.00
$\mathbf{1}\{\text{Savings Rate} \geq \text{Min. Savings Rate}\}$	0.31	0.46	0.00	0.00	1.00
Earned shared savings or losses	1.50	3.64	-4.66	0.00	41.91
Earned shared savings, given qualified	4.95	5.11	0.00	3.48	41.91
Proportion of expenditure on inpatient services^a	0.31	0.03	0.22	0.31	0.43
Proportion of expenditure on outpatient services^a	0.20	0.06	0.08	0.19	0.49
Number of primary care services (1000s)	10.29	1.76	5.39	9.98	26.16
Number of inpatient admissions (1000s)	0.33	0.09	0.17	0.32	1.86

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.)

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.

1.2.1 Economic Intuition and Motivating Regressions

Recall the previous example: this hypothetical ACO had a benchmark of $186 million, an actual expenditure of $180 million, a quality score of 0.9, and a sharing rate of 0.5. The ACO earned $2.7 million in incentive pay, and Medicare saved $3.3 million after paying the ACO. How would Medicare's savings change if the sharing rate increased to 0.75? ACO participants would have a larger incentive to improve efficiency of care, causing more savings and higher quality of care. However, Medicare would need to pay the ACO 50% more per dollar of savings.

The exact contract that maximizes the money providers save, less the payment to those providers, depends on the implicit and explicit costs incurred by ACO participants when saving money and increasing care quality. Savings-maximizing contracts share a larger proportion of ACO savings when ACO participants have larger marginal costs for spending reductions and improvements in quality. Because ACOs are paid as a group, the incentive for a given provider to decrease expenditure and increase quality of care is diluted, so free-riding in ACOs hinders performance. Free-riding effectively increases the marginal cost of savings and quality, so savings-maximizing contracts that take free-riding into account are more generous than contracts that do not. In the following sections of this paper, I build and estimate a model of provider participation and performance in ACOs. This yields empirical estimates of the cost of reducing expenditure and increasing quality incurred by ACO participants, and facilitates counterfactual simulations of contracts that maximize ACO savings, less payment to ACOs, while taking multitasking and free-riding into account. To write a model of ACOs, I make the assumption that ACO participants make decisions according to the incentives imposed by the program. I present evidence for this assumption by conducting two reduced form empirical exercises.

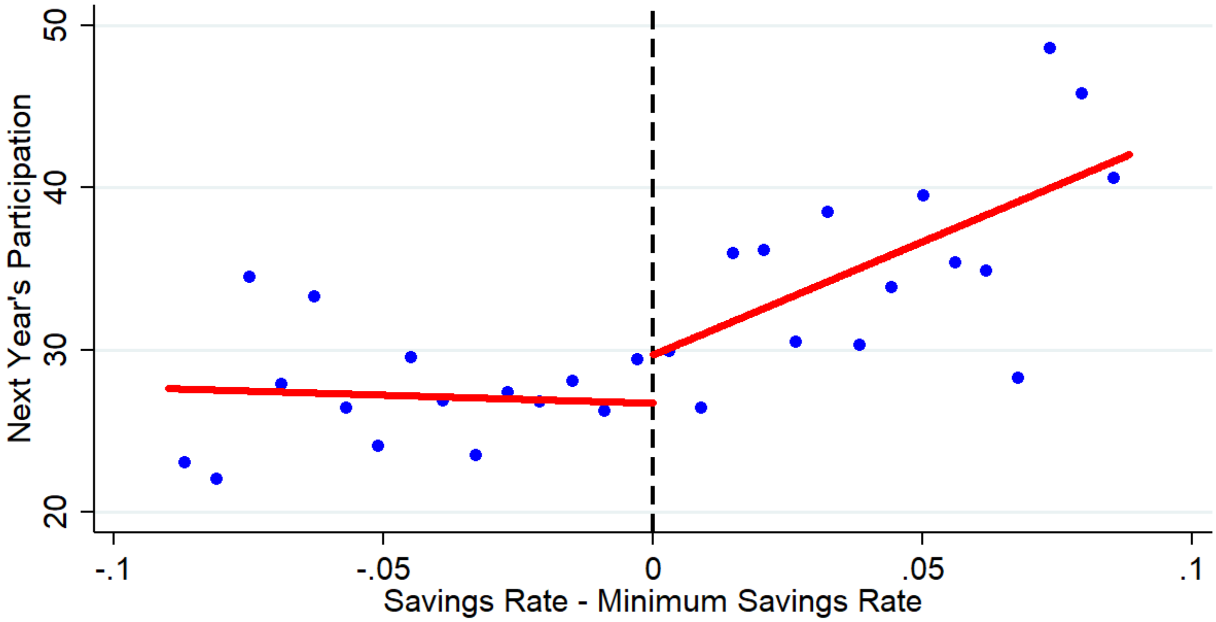
ACO Earnings and Participation

In the first exercise, I use a regression discontinuity design (RDD) to examine how an ACO's success in earning shared savings causes changes in participation in the following year. The purpose of this exercise is to show that the voluntary participation in the MSSP should be considered when modeling ACOs. Specifically, I estimate the model

$$n_{jt+1} = \alpha_0 + \alpha_1 \left(S_{jt} - \underline{S}_{jt}\right) + \beta_0 \mathbf{1}\{S_{jt} \geq \underline{S}_{jt}\} + \beta_1 \mathbf{1}\{S_{jt} \geq \underline{S}_{jt}\} \left(S_{jt} - \underline{S}_{jt}\right) + \varepsilon_{jt+1} \quad (1.3)$$

where n_{jt+1} is the number of Medicare providers participating in ACO j in year $t+1$, S_{jt} is the savings rate of ACO j in year t, and $\underline{S}_{jt}$ is the minimum savings rate of ACO j in year t. When ACO's do not earn incentive pay, α_1 is change in participation with respect to how close the ACO is to the threshold savings level. When ACO's qualify, $\alpha_1 + \beta_1$ is change in participation. Similarly, α_0 and β_0 determine the level of participation.

Figure 1.1: **Participation vs. Qualifying**



Note: This figure shows a binned scatter plot of ACO participation vs. the difference between ACO savings rate and minimum savings rate in the previous year. Plotted lines are from estimates of an RDD model, and show providers increase participation in response to increased ACO earnings.

Figure 1.1 plots a binned scatter plot of n_{jt+1} vs. $S_{jt} - \underline{S}_{jt}$ along with a line fitted using estimated values of α_0, α_1, β_0, and β_1.[5] When an ACO does not qualify for shared savings in year t, there is no relationship between the distance between savings rate and minimum savings rate in year t and ACO participation in year $t+1$. When an ACO qualifies, there is initially a small jump in participation, and the change in participation in year $t+1$ increases as the difference between savings rate and minimum savings rate grows. The results of this exercise indicate that providers are more likely to join ACOs that earn incentive pay. Moreover, providers are more likely to join an ACO as the amount of incentive pay earned by the ACO increases.

[5] Estimation is done with domain $|S_{jt} - \underline{S}_{jt}| < 0.09$. The parameter β_0 has p-value 0.33, and the parameter β_1 has p-value 0.051. Models with higher order polynomials of $S_{jt} - \underline{S}_{jt}$ have smaller F-statistics, so linear RDD is the preferred specification.

Tradeoff between Cost-Savings and Quality of Care

An essential part of designing incentives for physicians is accounting for quality of care. Medicare achieves this in the MSSP by assigning a quality of care score to ACOs and making it a determinant of incentive pay (see Equation 1.1). In this exercise, I study the empirical relationship between ACO savings rates and quality scores to show tradeoffs between spending reduction and quality scores. There are two unobserved factors correlated positively with both S_{jt} and Q_{jt} that confound estimating a tradeoff: 1) when an ACO has a high savings rate, the marginal benefit of quality score is larger, so quality scores are higher; 2) latent ACO efficiency drives both savings rates and quality scores higher. For these reasons, I estimate the regression

$$S_{jt} = \alpha_0 + \alpha_1 Q_{jt} + \beta_0 \left(S_{jt} - \underline{S}_{jt}\right) + \beta_1 Q_{jt} \left(S_{jt} - \underline{S}_{jt}\right) + \gamma_j + \delta_t + \varepsilon_{jt} \tag{1.4}$$

where Q_{jt} is ACO j's overall quality score in year t, S_{jt} is savings rate, $\underline{S}_{jt}$ is the minimum savings rate, γ_j is an ACO fixed effect, and δ_t is a year fixed effect.[6] This specification controls for the increasing marginal benefit of quality with respect to savings by including $S_{jt} - \underline{S}_{jt}$: when the difference between savings rate and minimum savings rate grows, marginal benefit of quality grows. By including ACO fixed effects, any ACO-specific and time invariant characteristics determining ACO efficiency are controlled for. The year fixed effects control for nationwide factors impacting ACO efficiency.

Table 1.3 shows results from estimating this regression. Column 1 estimates a univariate

[6] Note that I am interacting Q_{jt} with $S_{jt} - \underline{S}_{jt}$ and not with $\mathbf{1}\left\{S_{jt} \geq \underline{S}_{jt}\right\}$ in Equation 1.4. I do this because as $S_{jt} - \underline{S}_{jt}$ increases past zero, the marginal benefit of quality score continues to grow, and so the former interaction exploits more variation in the data to identify a savings-quality tradeoff.

Table 1.3: **Empirical Relationship between ACO Savings Rates and Quality Scores**

	Dependent Variable: S_{jt} (Savings Rate)			
	(1)	(2)	(3)	(4)
Q_{jt} (Quality Score)	0.0268*	0.0216	-0.0119***	-0.00805**
	(0.0107)	(0.0147)	(0.00190)	(0.00292)
$S_{jt} - \underline{S}_{jt}$ (Savings Rate – Min. Savings Rate)			0.838***	0.781***
			(0.0248)	(0.0291)
$Q_{jt}\left(S_{jt} - \underline{S}_{jt}\right)$			0.167***	0.226***
			(0.0285)	(0.0333)
Constant	-0.0147	-0.0102	0.0388***	0.0353***
	(0.00933)	(0.0127)	(0.00167)	(0.00255)
ACO and Year FE?	No	Yes	No	Yes
N	1849	1849	1849	1849

Standard errors in parentheses. * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

This table shows coefficient estimates for Equation 1.3.

regression of ACO savings on quality score, and the results show that confounding by the incentives imposed by the program and by latent ACO efficiency yield a positive raw relationship. When including ACO and year fixed effects (column 2) or controlling for marginal benefit of quality (column 3), the relationship between ACO savings and quality becomes smaller. Results of the full specification are in column 4. The estimates imply that when ACOs do not earn shared savings and $S_{jt} < \underline{S}_{jt}$, the relationship between ACO savings rates and quality scores is negative.

Table 1.4 presents the elasticity of savings rates with respect to quality scores, conditional on the how far an ACO's savings rate is from the minimum savings rate. When ACO's are far from earning shared savings and $S_{jt} - \underline{S}_{jt} = -0.10$, a 1 percent increase in quality score is associated with a 3.18 percent decrease in savings rate. At the cutoff point for earning shared savings ($S_{jt} - \underline{S}_{jt} = 0.00$) the relationship between savings and quality is approximately unit elastic with an elasticity of -0.84. For ACO's with sufficiently large values of $S_{jt} - \underline{S}_{jt}$, the

Table 1.4: **Elasticity of ACO Savings Rate with respect to Quality Score**

	$S_{jt} - \underline{S}_{jt}$ (Savings Rate – Min. Savings Rate)						
	-0.10	-0.05	-0.02	0.00	0.02	0.05	0.10
Elasticity $\frac{\partial S_{jt}}{\partial Q_{jt}} \cdot \frac{Q_{jt}}{S_{jt}}$	-3.18	-2.01	-1.31	-0.84	-0.37	0.34	1.51

This table shows the elasticity of ACO savings rate with respect to quality score, broken down by the distance savings rate is from the minimum savings rate. Each value is the percent change in savings rate when quality score increases by one percent, given savings rate minus minimum savings rate takes the value in the column header. These elasticities are computed at the means of savings rates and quality scores and use coefficient estimates in column (4) of Table 1.3.

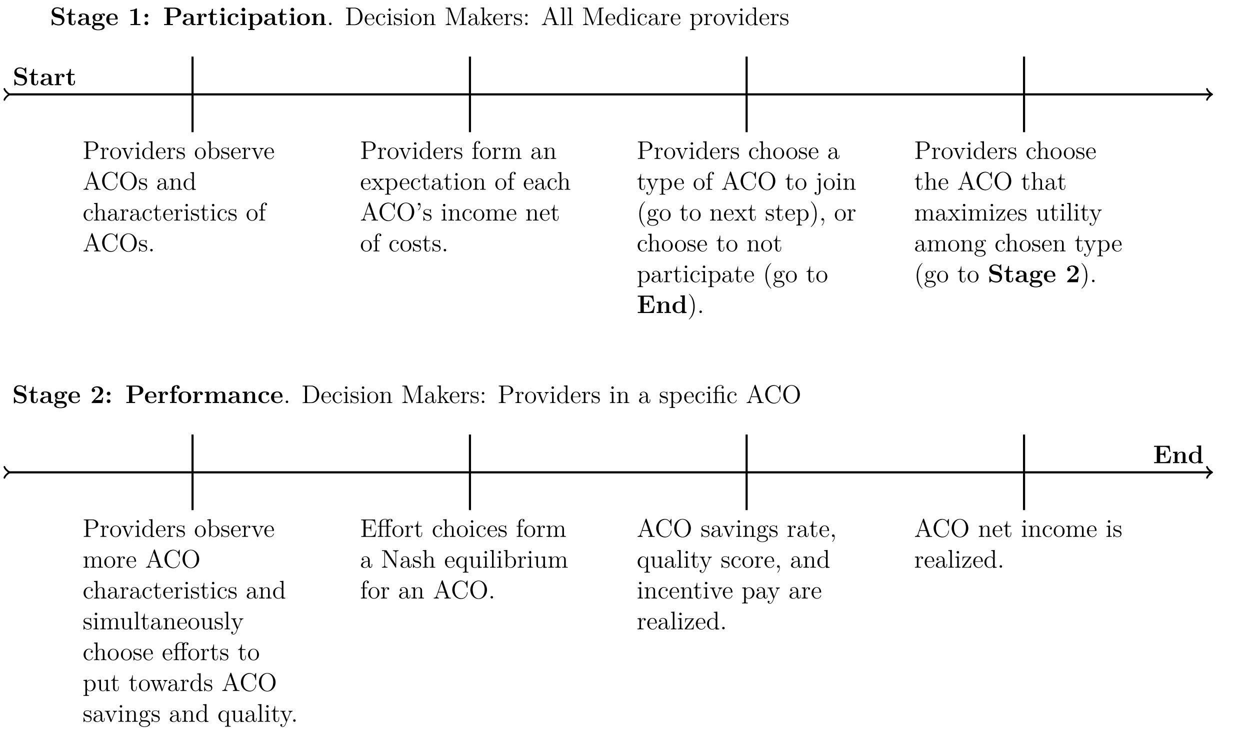
relationship between ACO savings rates and quality scores is positive, since for this range providers in an ACO have a large incentive to increase savings and quality simultaneously.

These empirical exercises have shown original concerns about voluntary participation and multitasking of healthcare providers are warranted. The limitation of these exercises is they do not clearly indicate the extent to which these concerns impact contracts that maximize the monetary savings of the incentive program. In the next section, I build a model that includes voluntary participation, multitasking, and strategic behavior so that these complications can be accounted for when designing incentives.

1.3 Model of Participation and Performance in ACOs

I model participation and performance in ACOs as a two-stage decision process. 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

Figure 1.2: **Model Timing**



Note: This figure shows timing of the model of ACOs. In the first stage, all Medicare providers choose which, if any, ACO to join. In the second stage, providers joining an ACO play a simultaneous move game that determines the ACO's performance.

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.

The timing of decisions in the model is pictured in Figure 1.2. In the first stage, a potential participant chooses which ACO to join. The decision is in two steps: first, they choose between joining a hospital led ACO, a physician led ACO, an ACO with mixed leadership, or no ACO at all.[7] If the provider chooses a type, they then choose an ACO to join among that type. With this structure, I allow for correlation of utilities of participants in ACOs in the same type, accounting for the possibility that providers are ex-ante more

[7]ACO leadership types are described in Table 1.1.

likely to join an ACO of a given type. A Medicare provider chooses to join the ACO that offers the highest utility, which is a function of their expected income from participating (net of costs) and observed characteristics of the ACO. ACO characteristic observed in the stage of the model are things determined *at the time of participation*; for example, the number of beneficiaries assigned to an ACO and the number of individual providers in the ACO are observed, but the expenditure of the ACO is not observed because its not determined when participation decisions are made.

In the second stage, participation is taken as given, and each member chooses unobservable efforts to put towards savings and quality in order to maximize their own payoff. 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 1.4 details this process.

In the model's first stage, the decision makers in this model are Medicare providers that qualify as a participant in the MSSP. Medicare providers that are participants in an ACO are the decision makers in the model's second stage. These are heterogeneous groups, each including individual providers, group practices, and hospitals. The set of potential participants $\mathcal{I}$ and set of all ACOs $\mathcal{J}$ are exogenous.

1.3.1 Participation

The model starts with providers simultaneously choosing which, if any, ACO to join. Providers have preferences for ACO income, net of costs, and for specific ACO characteristics. Let $j \in \{1, \ldots, J\} = \mathcal{J}$ index ACOs, and $j = 0$ denote the outside option to not join any ACO. The potential participant $i \in \mathcal{I}$ gets the following utility from joining ACO $j \neq 0$ in nest d:

$$u_{ij} = \alpha_i y_j + \boldsymbol{\beta}' X_j^{part} + \xi_j + \rho\zeta_{id} + (1-\rho)\epsilon_{ij} \,. \tag{1.5}$$

The variable y_j is the net income of an ACO, and X_j^{part} is a vector of ACO characteristics that observed at the time of participation, including assigned beneficiary demographics and provider characteristics. The variable ξ_j is unobserved ACO heterogeneity that is possibly correlated with y_j and X_j^{part}. The term $\rho\zeta_{id}$ is i's specific preference for participating in an ACO in the nest d, where nests are the leadership type of an ACO (and the outside option): $d \in \{hosp, phys, mix, 0\}$. The variable ϵ_{ij} is an idiosyncratic utility shock. Note that ξ_j, $\rho\zeta_{id}$, and ϵ_{ij} are observed by providers i, but all are unobserved by the econometrician. Following Berry (1994) and Cardell (1997), I assume ϵ_{ij} is distributed Type I Extreme Value, and ζ_{id} has the unique distribution such that $\rho\zeta_{id} + (1-\rho)\epsilon_{ij}$ is distributed Type I Extreme Value. I normalize the utility of the outside option, $j = 0$, to $u_{i0} = \rho\zeta_{i0} + (1-\rho)\epsilon_{i0}$.

Net income y_j captures the expected pecuniary benefit to a provider for participating in ACO j. I use the net income of ACO j, and not the earned shared savings of ACO j, to account for the implicit and explicit effort cost incurred by spending less on Medicare beneficiaries and providing higher quality of care. If earned shared savings payments were used instead of net income, pecuniary benefit for participation would be overstated. In order to measure net income, I subtract the estimated increase in cost incurred by providers while operating in an ACO from the earned shared savings of an ACO. I discuss net income in detail in Section 1.3.3.

The parameters in the first stage of this model are α_i, individual i's marginal utility of net income; $\boldsymbol{\beta}$, a vector describing mean preferences over ACO characteristics; and the nesting parameter $\rho \in [0, 1]$, which measures the correlation of utilities of members in the same nest. As ρ increases, the influence an ACO's nest has over a participant's decision increases. The set of parameters in the first stage of this model is denoted $\boldsymbol{\theta}_1 = \{\alpha_i, \boldsymbol{\beta}, \rho\}$.

The parameter of primary interest is α_i. If positive, then Medicare providers are less likely to join an ACO when net income is lower, warranting concerns about voluntary participation when designing contracts. Though plausible, this fact has not been established in health or economics literature. For reference, Ryan et al. (2015), Yasaitis et al. (2016), and Mansour et al. (2017) discuss physician income and ACO participation, though participation in response to income is inconclusive.

Let I_j be the set of $i \in \mathcal{I}$ that choose to join ACO j. Formally:

$$I_j = \left\{ i \in \mathcal{I} : j = \arg\max_{j' \geq 0} u_{ij'} \right\} \tag{1.6}$$

Because ϵ_{ij} and $\rho\zeta_{id} + (1-\rho)\epsilon_{ij}$ are distributed Type I Extreme Value, the probability of observing $i \in I_j$ where j is in nest d is

$$a_{ij} = \frac{\exp\left(\frac{\alpha_i y_j + \beta' X_j^{part} + \xi_j}{1-\rho}\right)}{\left[\sum_{j' \in d} \exp\left(\frac{\alpha_i y_{j'} + \beta' X_{j'}^{part} + \xi_{j'}}{1-\rho}\right)\right]^{\rho} \cdot \sum_d \left[\sum_{j'' \in d}\left(\frac{\alpha_i y_{j''} + \beta' X_{j''}^{part} + \xi_{j''}}{1-\rho}\right)\right]^{1-\rho}}. \tag{1.7}$$

The total number of providers in ACO j is $n_j \equiv M a_{ij}$, where $M = |\mathcal{I}|$.

1.3.2 Performance

In the second stage of this model, participating Medicare providers in ACOs choose their own savings and quality efforts, which in turn determines each ACO's overall savings rate and quality score.[8] Participant savings and quality efforts are chosen strategically to maximize a participant's profit from participating in an ACO.

Recall, n_j participants decide to join ACO j, and the set of all Medicare providers in ACO j is denoted I_j. All participants $i \in I_j$ simultaneously choose savings and quality efforts $s_{ij} \in [-1, 1]$ and $q_{ij} \in [0, 1]$.[9] These choices determine ACO savings rate S_j and overall quality score Q_j through the weighted sums

$$S_j = \sum_{i \in I_j} w_{ij} s_{ij} \qquad\qquad Q_j = \sum_{i \in I_j} w_{ij} q_{ij}\,. \tag{1.8}$$

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 participants' efforts on ACO performance.[10]

[8] 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.

[9] 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).

[10] 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 $\bar{s}_j = 0.03$. 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 A.3 for more details.

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

$$\max_{s_{ij}, q_{ij}} R_{ij}(S_j, Q_j) - c_{ij}(s_{ij}, q_{ij}) \tag{1.9}$$

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, $c_{ij}(s_{ij}, q_{ij})$ 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

$$R_j(S_j, Q_j) = \begin{cases} 0.5 \cdot B_j S_j Q_j & \text{if } S_j \geq \underline{S}_j \text{ and } Q_j \geq \underline{Q} \\ 0 & \text{otherwise} \end{cases} \tag{1.10}$$

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.[11] 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

$$\frac{\partial c_{ij}}{\partial s_{ij}}(s_{ij}, q_{ij}) = \begin{cases} 0.5 \cdot B_j w_{ij}^2 Q_j & \text{if } S_j \geq \underline{S}_j \text{ and } Q_j \geq \underline{Q} \\ 0 & \text{otherwise} \end{cases} \tag{1.12}$$

$$\frac{\partial c_{ij}}{\partial q_{ij}}(s_{ij}, q_{ij}) = \begin{cases} 0.5 \cdot B_j w_{ij}^2 S_j & \text{if } S_j \geq \underline{S}_j \text{ and } Q_j \geq \underline{Q} \\ 0 & \text{otherwise} \end{cases} . \tag{1.13}$$

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.[12] 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.

The shared savings function R_j is written in a way such that it *may* generate a simultaneous move game with strategic complementarity. These games have the property that the best response function of a player is increasing in the strategies of the other players (Bulow et al., 1985; Milgrom & Roberts, 1990). In the context of ACOs, this means that the optimal

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

$$R_j(S_j, Q_j) = \begin{cases} 0.5 \cdot B_j S_j & \text{if } S_j \geq \underline{S}_j \text{ and } Q_j \geq \underline{Q} \\ 0 & \text{otherwise} \end{cases} \tag{1.11}$$

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

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

savings and quality effort choices of a physician are larger when a different physician chooses higher savings and quality efforts. The game played by ACO participants is supermodular if the following holds: 1) $\frac{\partial^2 c_{ij}}{\partial s_{ij} \partial q_{ij}} \leq \frac{w_{ij}^3}{2} B_j$, 2) $S_j \gg \underline{S}_j$, and 3) $Q_j \gg \underline{Q}$.[13]

While strategic complementarity does not hold in general, the concept is useful to guide one's understanding of the game that's played within ACOs. Specifically, because savings and quality enter multiplicatively into an ACO's pay, if there is a small enough tradeoff between savings and quality for a given physician, then an increase in effort of one member has the response of higher effort of another member. When one member decreases their effort, others decrease theirs as well. The game will tend to have two equilibria: one with coordination where the ACO earns shared savings, and one without coordination where providers minimize cost and the ACO fails to earn shared savings.

Conditions 1), 2), and 3) above are not usually satisfied (upon estimation of c_{ij}, $\frac{\partial^2 c_{ij}}{\partial s_{ij} \partial q_{ij}}$ is too large), so I prove existence of equilibrium in this game without relying on the presence of supermodularity in Proposition A.2.3 in Appendix A.2. I also show, in general, there is not a unique equilibrium. There can be 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

$$\max_{s_{ij}, q_{ij}} 0.5 \cdot w_{ij} B_j S_j Q_j - c_{ij}\left(s_{ij}, q_{ij}\right) \,, \tag{1.14}$$

[13] I prove that the game is supermodular under these conditions in Appendix A.2.

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

$$\min_{s_{ij},q_{ij}} c_{ij}\left(s_{ij}, q_{ij}\right) \,, \tag{1.15}$$

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.

1.3.3 Net Income

An ACO's net income is the realized increase in earnings by all members of an ACO in a given performance year by participating in the MSSP. By joining an ACO and earning shared savings, a participant acts differently than they otherwise would, which carries explicit and implicit costs. If the ACO does not qualify for shared savings, net income is defined as zero—participants choose savings and quality efforts in the same way they would were they not participating in an ACO. That is, the solution to the problem in Equation 1.15 is also the counterfactual effort choice of participant i when it does not join an ACO. Thus, if an ACO does qualify for shared savings, then net income is the total earned subsidy of the ACO, minus the increase in cost incurred by participants for having savings and quality efforts higher than they would otherwise be.

Let the profit of participant i in ACO j be

$$\pi_{ij}\left(\mathbf{s}_j, \mathbf{q}_j\right) = R_{ij}(S_j, Q_j) - c_{ij}\left(s_{ij}, q_{ij}\right) . \tag{1.16}$$

Let y_j denote net income. Define

$$(\tilde{s}_{ij}, \tilde{q}_{ij}) = \arg\min_{s_{ij}, q_{ij}} c_{ij}\left(s_{ij}, q_{ij}\right) . \tag{1.17}$$

Then,

$$y_j = \sum_{i \in I_j} \{R_{ij}\left(S_j^*, Q_j^*\right) - \overbrace{\left[c_{ij}\left(s_{ij}^*, q_{ij}^*\right) - c_{ij}\left(\tilde{s}_{ij}, \tilde{q}_{ij}\right)\right]}^{\text{Increase in cost from earning shared savings}}\} = \sum_{i \in I_j}\left[\pi_{ij}\left(\mathbf{s}_j^*, \mathbf{q}_j^*\right) + c_{ij}\left(\tilde{s}_{ij}, \tilde{q}_{ij}\right)\right] \tag{1.18}$$

When the equilibrium profile of ACO j is such that ACO participants minimize cost, net income is equivalently zero. I discuss the computation of y_j from data in Section 1.4.

1.4 Identification and Estimation

To estimate model primitives, I use ACO-level data on Track 1 ACOs from 2014 to 2017.[14] The first year of the program, 2013, is omitted because participation information is not available. 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 were included. Identification and estimation of this model is complicated by the limited data available on ACO

[14]Summary statistics for the data are presented in Tables 1.1 and 1.2 in Section 2.4.

Table 1.5: **Summary of Model and Data used for Estimation**

Model Stage	Primitives	Data
1 - Participation	Parameters $\boldsymbol{\theta}_1$ in utility from participation u_{ij}.	Share of participation: a_j Estimated net income: $y_j(\hat{\boldsymbol{\theta}}_2)$ ACO characteristics determined during participation: X_j^{part}
2 - Performance	Parameters $\boldsymbol{\theta}_2$ in parameterized cost functions c_{ij}.	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 determined during performance: X_j^{perf}

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

participants. Only aggregate data are observed: decisions in both stages of 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 methods from empirical industrial organization (Berry, 1994; Berry, Levinsohn, & Pakes, 1995; Nevo, 2000) to map ACO characteristics and performance to model primitives.

Table 1.5 gives an overview of the model primitives and data used for estimation of the primitives. Computing ACO net income y_j requires subtracting effort cost from earned shared savings, so I estimate the model backwards. First, 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 $\boldsymbol{\theta}_2$. Using estimated cost function parameters, I compute an estimate of net income $y_j\left(\hat{\boldsymbol{\theta}}_2\right) \equiv \hat{y}_j$ for each ACO, and finally I use the estimate of net income $\hat{y}_j$ to estimate parameters describing average utility from participation, $\boldsymbol{\theta}_1$. 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. I also assume providers have rational expectations over the net income of an ACO, so variation estimated net income $\hat{y}_j$ identifies average marginal utility of net income.

1.4.1 Identification and Estimation of Participant Cost Functions

Overview

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.[15] Incorporating the structure of ACO participants' incentives is essential to obtain the results in this paper.

Recall participant first order conditions in Equations 1.12 and 1.13. These hold for any Nash equilibrium, so:

$$\frac{\partial c_{ij}}{\partial s_{ij}}\left(s_{ij}^*, q_{ij}^*\right) = \left(0.5 B_j w_{ij}^2 Q_j^*\right) \mathbf{1}\left\{S_j^* \geq \underline{S}_j\right\} \mathbf{1}\left\{Q_j^* \geq \underline{Q}\right\} \ \forall i \in I_j \tag{1.19}$$

$$\frac{\partial c_{ij}}{\partial q_{ij}}\left(s_{ij}^*, q_{ij}^*\right) = \left(0.5 B_j w_{ij}^2 S_j^*\right) \mathbf{1}\left\{S_j^* \geq \underline{S}_j\right\} \mathbf{1}\left\{Q_j^* \geq \underline{Q}\right\} \ \forall i \in I_j \,. \tag{1.20}$$

That is, for any Nash equilibrium, the marginal cost of savings effort (the left hand side of Equation 1.19) is equal to the marginal subsidy of savings effort (the right hand side of

[15] Indeed, this seems to be the case: see the discussion of Table 1.3 Section 2.4.

Equation 1.19) for all participants $i \in I_j$. Similarly, the marginal cost of quality effort is equal to the marginal benefit of quality effort (Equation 1.20). Pre-multiplying each side of the equations by w_{ij} and summing over $i \in I_j$, we get

$$\sum_{i \in I_j} w_{ij} \frac{\partial c_{ij}}{\partial s_{ij}} \left(s_{ij}^*, q_{ij}^* \right) = 0.5 W_j^{(3)} B_j Q_j^* \mathbf{1} \left\{ S_j^* \geq \underline{S}_j \right\} \mathbf{1} \left\{ Q_j^* \geq \underline{Q} \right\} \tag{1.21}$$

$$\sum_{i \in I_j} w_{ij} \frac{\partial c_{ij}}{\partial q_{ij}} \left(s_{ij}^*, q_{ij}^* \right) = 0.5 W_j^{(3)} B_j S_j^* \mathbf{1} \left\{ S_j^* \geq \underline{S}_j \right\} \mathbf{1} \left\{ Q_j^* \geq \underline{Q} \right\} , \tag{1.22}$$

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 A.3.

Equations 1.21 and 1.22 are aggregate analogs of Equations 1.19 and 1.20. 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 1.21 and 1.22 be denoted MC_j^S and MC_j^Q, respectively, and similarly let the right hand sides of Equations 1.21 and 1.22 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 1.4.1. The effort cost function c_{ij} takes the quadratic form

$$c_{ij}\left(s_{ij}, q_{ij}\right) = \frac{\delta_S}{2} s_{ij}^2 + \frac{\delta_Q}{2} q_{ij}^2 + \boldsymbol{\gamma}_S' x_{ij}^{perf} s_{ij} + \boldsymbol{\gamma}_Q' x_{ij}^{perf} q_{ij} + \kappa s_{ij} q_{ij}\,, \tag{1.23}$$

where $x_{ij}^{perf} \in \mathbb{R}^k$ 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 1.4.1, we have

$$\frac{\partial c_{ij}}{\partial s_{ij}}\left(s_{ij}^*, q_{ij}^*\right) = \delta_S s_{ij}^* + \boldsymbol{\gamma}_S' x_{ij}^{perf} + \kappa q_{ij}^* \tag{1.24}$$

$$\frac{\partial c_{ij}}{\partial q_{ij}}\left(s_{ij}^*, q_{ij}^*\right) = \delta_Q q_{ij}^* + \boldsymbol{\gamma}_Q' x_{ij}^{perf} + \kappa s_{ij}^*\,. \tag{1.25}$$

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:

$$MC_j^S = \sum_{i \in I_j} w_{ij} \frac{\partial c_{ij}}{\partial s_{ij}} \left(s_{ij}^*, q_{ij}^* \right) = \delta_S \left(\sum_{i \in I_j} w_{ij} s_{ij}^* \right) + \boldsymbol{\gamma}'_S \left(\sum_{i \in I_j} w_{ij} x_{ij}^{perf} \right) + \kappa \left(\sum_{i \in I_j} w_{ij} q_{ij}^* \right) \tag{1.26}$$

$$= \delta_S S_j^* + \boldsymbol{\gamma}'_S X_j^{perf} + \kappa Q_j^* \,, \tag{1.27}$$

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

$$MC_j^Q = \sum_{i \in I_j} w_{ij} \frac{\partial c_{ij}}{\partial q_{ij}} \left(s_{ij}^*, q_{ij}^* \right) = \delta_Q \left(\sum_{i \in I_j} w_{ij} q_{ij}^* \right) + \boldsymbol{\gamma}'_Q \left(\sum_{i \in I_j} w_{ij} x_{ij}^{perf} \right) + \kappa \left(\sum_{i \in I_j} w_{ij} s_{ij}^* \right) \tag{1.28}$$

$$= \delta_Q Q_j^* + \boldsymbol{\gamma}'_Q X_j^{perf} + \kappa S_j^* \,. \tag{1.29}$$

Combining this with Equations 1.21 and 1.22, we get

$$MB_j^S = \delta_S S_j^* + \boldsymbol{\gamma}'_S X_j^{perf} + \kappa Q_j^* \tag{1.30}$$

$$MB_j^Q = \delta_Q Q_j^* + \boldsymbol{\gamma}'_Q X_j^{perf} + \kappa S_j^* \,. \tag{1.31}$$

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

$$\nu_j^S \left(\boldsymbol{\theta}_2 \right) = MB_j^S - \delta_S S_j^* + \boldsymbol{\gamma}'_S X_j^{perf} + \kappa Q_j^* \tag{1.32}$$

$$\nu_j^Q \left(\boldsymbol{\theta}_2 \right) = MB_j^Q - \delta_Q Q_j^* + \boldsymbol{\gamma}'_Q X_j^{perf} + \kappa S_j^* \tag{1.33}$$

where $\boldsymbol{\theta}_2 = \{\delta_S, \delta_Q, \boldsymbol{\gamma}_S, \boldsymbol{\gamma}_Q, \kappa\}$. Next, I make the following identifying assumption.

Assumption 1.4.2. The error terms $\nu_j^S(\boldsymbol{\theta}_2)$ and $\nu_j^Q(\boldsymbol{\theta}_2)$ have mean zero and are independent of S_j^*, Q_j^* and X_j^{perf}.

Assumption 1.4.2 is necessary so that estimates of $\boldsymbol{\theta}_2$ 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 $\boldsymbol{\theta}_2$ are identified.[16]

The elements of X_j^{perf}, which can be interpreted as ACO-specific marginal cost shifters, appear in Table 1.6. Note that X_j^{perf} includes information concerning assigned beneficiaries, participating Medicare providers, as well as expenditure and service statistics.

To estimate $\boldsymbol{\theta}_2$, I apply Assumption 1.4.2 and use the moment conditions

$$\mathbb{E}\left[\begin{array}{c} \nu_j^S(\boldsymbol{\theta}_2) \\ \nu_j^Q(\boldsymbol{\theta}_2) \end{array}\middle| S_j^*, Q_j^*, X_j^{perf}\right] = 0 \tag{1.34}$$

[16]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^{perf}. 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.

Table 1.6: **Elements of X_j^{perf} and X_j^{part}**

Abbreviated Variable Name	**Description**	**In X_j^{perf}**	**In X_j^{part}**
# 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 and participation utility parameters. Not listed: Constant term, year and census division fixed effects. A checkmark in the third or fourth column indicates that variable is in the vector of characteristics X_j^{perf} or X_j^{part}, respectively.

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

$$\hat{\boldsymbol{\theta}}_2 = \arg\min_{\theta\in\Theta} \left(\boldsymbol{\nu}\left(\boldsymbol{\theta}_2\right)' \begin{bmatrix} \mathbf{S} & \mathbf{Q} & \mathbf{X}^{perf} \\ \mathbf{S} & \mathbf{Q} & \mathbf{X}^{perf} \end{bmatrix} \right) \mathbf{W} \left(\boldsymbol{\nu}\left(\boldsymbol{\theta}_2\right)' \begin{bmatrix} \mathbf{S} & \mathbf{Q} & \mathbf{X}^{perf} \\ \mathbf{S} & \mathbf{Q} & \mathbf{X}^{perf} \end{bmatrix} \right)' \tag{1.35}$$

where $\boldsymbol{\nu}\left(\boldsymbol{\theta}_2\right) = \begin{bmatrix} \nu_j^S\left(\boldsymbol{\theta}_2\right) \\ \nu_j^Q\left(\boldsymbol{\theta}_2\right) \end{bmatrix}_{j\in\mathcal{J}}$ 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^{perf}; and $\mathbf{W}$ is a $2(k+2) \times 2(k+2)$ positive definite weighting matrix.[17]

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 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.

[17]Note that because the cross equation restriction that κ appears in both $\nu_j^S\left(\boldsymbol{\theta}_2\right)$ and $\nu_j^Q\left(\boldsymbol{\theta}_2\right)$, $\boldsymbol{\theta}_2$ is over-identified. There are $2k+4$ moments and $2k+3$ parameters in $\boldsymbol{\theta}_2$. 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 $\frac{\partial^2 c_{ij}}{\partial s_{ij}\partial q_{ij}} \equiv \frac{\partial^2 c_{ij}}{\partial q_{ij}\partial s_{ij}}$.

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

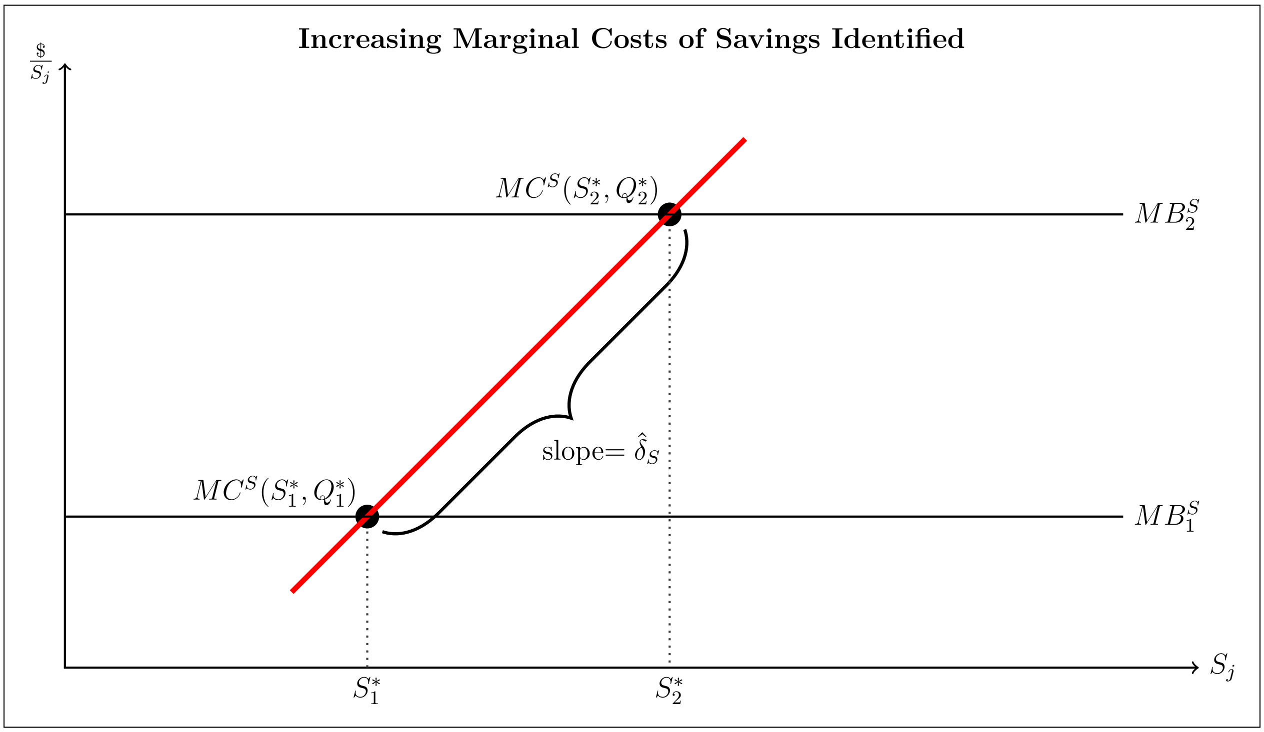
$$\frac{\partial^2 c_{ij}}{\partial s_{ij}^2} \approx \frac{MC^S(S_2^*, Q_2^*) - MC^S(S_1^*, Q_1^*)}{S_2^* - S_1^*} \equiv \hat{\delta}_S \, . \tag{1.36}$$

The argument is pictured in Figure 1.3. 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 savings-quality tradeoff) is equal to

$$\frac{\partial^2 c_{ij}}{\partial s_{ij} \partial q_{ij}} \approx \frac{MC^S(S_2^*, Q_2^*) - MC^S(S_2^*, Q_3^*)}{Q_2^* - Q_3^*} \, . \tag{1.37}$$

Figure 1.3: **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.

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

$$\frac{\partial^2 c_{ij}}{\partial s_{ij} \partial q_{ij}} \approx \frac{MC^S(S_2^*, Q_2^*) - MC^S(S_2^*, Q_3^*)}{Q_2^* - Q_3^*} \tag{1.38}$$

$$= \frac{MC^S(S_2^*, Q_2^*) - \left[MC^S(S_3^*, Q_3^*) - \hat{\delta}_S(S_3^* - S_2^*)\right]}{Q_2^* - Q_3^*} \tag{1.39}$$

$$= \frac{MC^S(S_2^*, Q_2^*) - \left[MC^S(S_2^*, Q_2^*) - \hat{\delta}_S(S_3^* - S_2^*)\right]}{Q_2^* - Q_3^*} \tag{1.40}$$

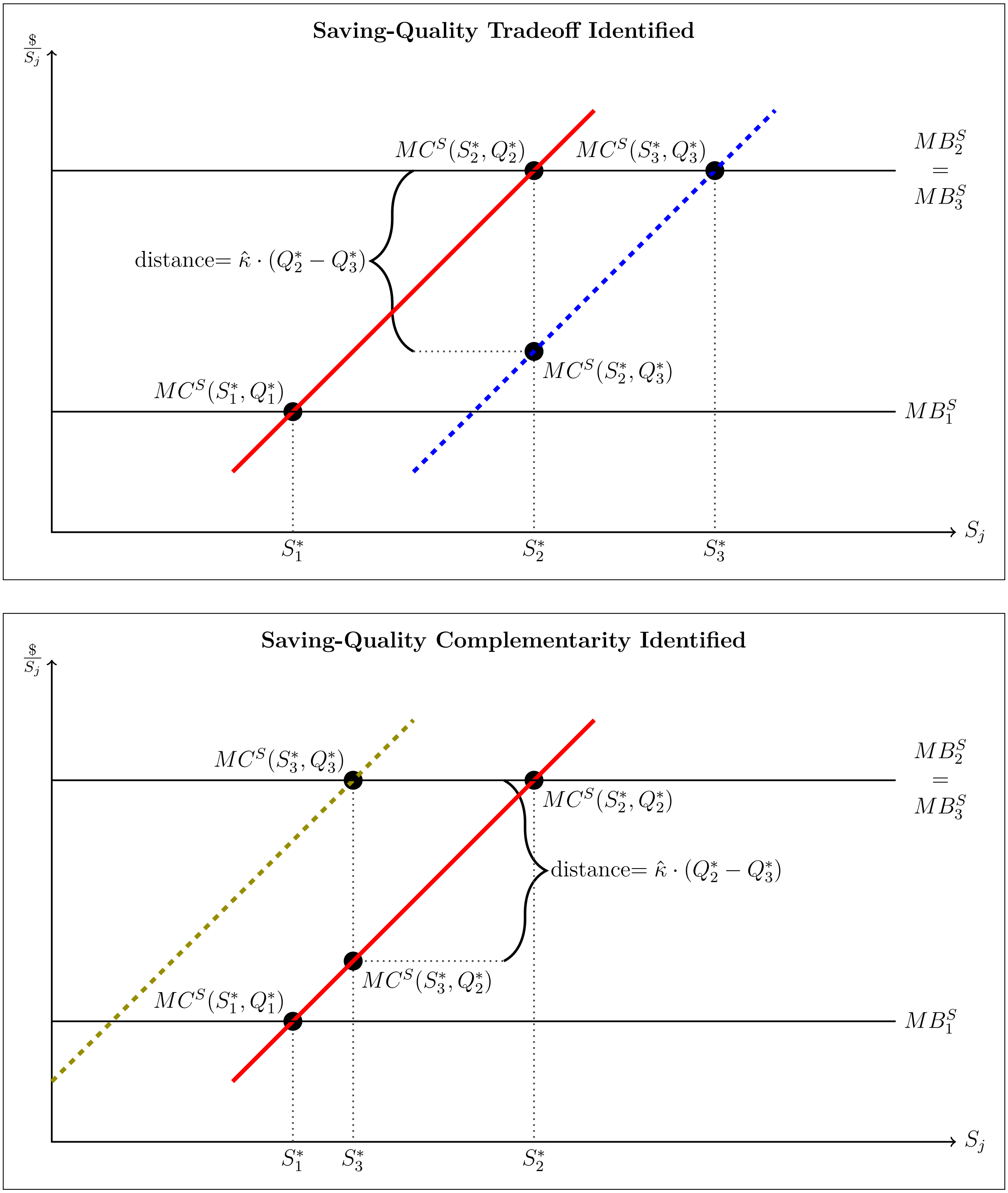
$$= \hat{\delta}_S \frac{S_3^* - S_2^*}{Q_2^* - Q_3^*} \equiv \hat{\kappa}\,. \tag{1.41}$$

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

Figure 1.4 shows this process for identifying a tradeoff in the top panel. The solid red line is the same that was found in Figure 1.3. 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 1.4 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, since increasing the quality score from Q_3^* to Q_2^* decreases marginal cost by $\hat{\kappa} \cdot (Q_2^* - Q_3^*)$.

Figure 1.4: **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 $S_1^* < S_3^* < S_2^*$.

1.4.2 Computation of Equilibrium and Net Income

In order to compute net income y_j and perform counterfactual exercises, I must solve for equilibrium values of S_j^* and Q_j^* using the estimate of $\boldsymbol{\theta}_2$. The following systems of two equations hold for any equilibrium and can be solved for the two unknowns, $\hat{S}_j^*$ and $\hat{Q}_j^*$:

$$F \cdot W_j^{(3)} B_j \hat{Q}_j^* = \hat{\delta}_S \hat{S}_j^* + \hat{\boldsymbol{\gamma}}_S' X_j^{perf} + \hat{\kappa} \hat{Q}_j^* \tag{1.42}$$

$$F \cdot W_j^{(3)} B_j \hat{S}_j^* = \hat{\delta}_Q \hat{Q}_j^* + \hat{\boldsymbol{\gamma}}_Q' X_j^{perf} + \hat{\kappa} \hat{S}_j^* \tag{1.43}$$

where F is the sharing rate for the ACO, and

$$0 = \hat{\delta}_S \hat{S}_j^* + \hat{\boldsymbol{\gamma}}_S' X_j^{perf} + \hat{\kappa} \hat{Q}_j^* \tag{1.44}$$

$$0 = \hat{\delta}_Q \hat{Q}_j^* + \hat{\boldsymbol{\gamma}}_Q' X_j^{perf} + \hat{\kappa} \hat{S}_j^* \, . \tag{1.45}$$

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 1.4.3.

The equilibrium that's played is the utilitarian equilibrium where net income is larger.

Note that while Assumption 1.4.3 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.

For net income, recall Equations 1.17 and 1.18. Net income has the equivalent expression:

$$y_j = \text{Earned Shared Savings of ACO } j - \sum_{i \in I_j} \left[c\left(s^*_{ij}, q^*_{ij}; x^{perf}_{ij}, \boldsymbol{\theta}_2\right) - c\left(\tilde{s}_{ij}, \tilde{q}_{ij}; x^{perf}_{ij}, \boldsymbol{\theta}_2\right) \right]. \tag{1.46}$$

The first term in Equation 1.46 is observed directly in data. The summation over $i \in I_j$ requires computing values of $\left(\mathbf{s}^*_j, \mathbf{q}^*_j\right)$ and $\left(\tilde{\mathbf{s}}_j, \tilde{\mathbf{q}}_j\right)$. This is not directly possible since only X^{perf}_j and not x^{perf}_{ij} is observed. In the results that follow, I approximate the second term in Equation 1.46 using participant choices coming from a symmetric equilibrium, such that $s^*_{ij} \equiv S^*_j$ and $q^*_{ij} \equiv Q^*_j$ (and similarly for $\tilde{s}_{ij}$ and $\tilde{q}_{ij}$). The estimated value of net income is therefore

$$\hat{y}_j = \text{Earned Shared Savings of ACO } j - n_j \left[c\left(s^*_{ij}, q^*_{ij}; X^{perf}_j, \hat{\boldsymbol{\theta}}_2\right) - c\left(\tilde{s}_{ij}, \tilde{q}_{ij}; X^{perf}_j, \hat{\boldsymbol{\theta}}_2\right) \right]. \tag{1.47}$$

1.4.3 Identification and Estimation of Utility from Participation

Recall the utility specification for participating in ACO j:

$$u_{ij} = \alpha_i y_j + \boldsymbol{\beta}' X^{part}_j + \xi_j + \rho \zeta_{id} + (1 - \rho)\epsilon_{ij} \tag{1.48}$$

Since participant level data on participating providers is unavailable, I follow Berry (1994) and Cardell (1997) to estimate α_i, $\boldsymbol{\beta}$, and ρ with aggregate data and accounting for unobserved heterogeneity ξ_j. For the nested logit specification, I divide ACOs $j \in \mathcal{J}$ into four nests d: the outside option, ACOs that are physician led, ACOs that are hospital led, and

ACOs with mixed leadership. Control variables in X_j^{part} are indicated by a checkmark in the fourth column of Table 1.6. Several variables present in X_j^{perf} omitted from X_j^{part} because they are not realized at the time participation decisions are made.

Provider preference for net income y_j is represented by the parameter α_i. To estimate dispersion in this preference without individual data, I let α_i be a normally distributed random variable with mean and standard deviation to be estimated. Formally, I specify $\alpha_i = \alpha_0 + \alpha_\eta \eta_i$ with $\eta_i \sim N(0,1)$.

To understand the variation in data that drives estimates of parameters in Equation 1.48, let's examine the estimating equation for *mean* participant utility, where variation in taste for net income $\alpha_\eta \equiv 0$. Let a_j be the share of potential ACO participants joining ACO j, let a_0 be the share not joining any ACO, and let a_d be the share of participants choosing any ACO in nest d. Under the nested logit assumptions on ϵ_{ij} and on ζ_{id} we have the estimating equation

$$\ln(a_j/a_0) = \alpha_0 y_j + \boldsymbol{\beta}' X_j^{part} + \rho \ln(a_j/a_d) + \xi_j. \tag{1.49}$$

Variation in net income y_j across ACOs for given characteristics X_j^{part} identifies α_0. Conceptually, controlling for ACO characteristics in X_j^{part} adjusts the estimate of α_0 so that it represents tastes for net income only, and is not biased by ACO characteristics that are correlated with (but do not determine) net income. This is particularly important for counterfactual analysis, where I will examine the change in participation in ACOs in response to changes in net income stemming from alternate payment mechanisms.

The unobserved heterogeneity term ξ_j is likely correlated with net income y_j. In particular, "demand" for ACO participants may confound estimates of α_0. ACOs with high net income y_j may become more selective of the providers they welcome into their ACO, amounting to a negative relationship between y_j and ξ_j. This manifests as a downward biased estimate of α_0. Moreover, as participation in an ACO grows, incentives continue to dilute and the free-rider problem grows, also amounting to a negative relationship between y_j and ξ_j and a downward biased estimate of α_0.

I use an instrumental variables approach to recover unbiased estimates of the preference for net income α_0. Specifically, I instrument for y_j using the elements in X_j^{perf} that are omitted from X_j^{part}. These omitted characteristics are correlated with net income by shifting effort cost, but there is otherwise no explicit preference for a characteristic not determined at the time of participation. Specifically, I use the fractions of expenditures on inpatient and outpatient services, number of primary care services, and fraction of primary care services served by a PCP as instruments for net income.

The parameter ρ acts as a weight on ACO-type specific shocks relative to ACO specific shocks for a given provider. When ρ is high, the utility shock $\rho\zeta_{id}$, which influences all choices in nest d, is high relative to the utility shock $(1-\rho)\epsilon_{ij}$, which influences choices independently. The parameter is identified by variation in shares of participation between ACO type: the correlation of utilities within groups would have a small estimate if there is significant substitution in the share of participation between groups in response to changes in other variables. There is an endogeneity problem here too: increases in the value of ξ_j cause increases in the share of participating providers a_j. If the additional providers participating in ACO j came primarily from not participating in an ACO, then $\ln(a_j/a_d)$ would be low, causing a downward bias of ρ. On the other hand, if ACO j gains participation share from a different group (such that $j \in d'$ with $d' \neq d$ initially), $\ln(a_j/a_d)$ would be high, causing a upward bias of ρ. I instrument for $\ln(a_j/a_d)$ with h_j/h_d, where h_j is the share of HMO enrolled physicians in an ACO's area, and h_d is similarly defined. These objects are correlated through providers' latent preferences for joining group payment systems (Frech et al. (2015) discuss this point), though relative shares of HMO enrollment alone are otherwise unlikely to contribute directly to participation in ACOs.

Finally, some control variables X_j^{part} may be equilibrium objects, and hence endogenous. While recovering unbiased estimates of $\boldsymbol{\beta}'$ is not of particular concern in this analysis (X_j^{part} is held constant in the counterfactuals), they will also be instrumented for so that estimates and standard errors of the parameters of interest α_0, α_η, and ρ are not impacted.[18] I assume the number of states occupied by an ACO's assigned beneficiaries is uncorrelated with unobserved heterogeneity ξ_j, and so it does not need an instrument. I obtain exogenous variation correlated with controls describing ACO assigned beneficiaries (number, average risk score, percent older than 75, percent male, and percent nonwhite) from Medicare county-level public use files. Each of these instruments require the assumption that the characteristics of Medicare beneficiaries in an ACOs area do not impact participation in a particular ACO, except through the ACO's assigned beneficiaries. Moreover, I use elements from X_j^{perf} that are omitted from X_j^{part} to instrument for the total number of individual providers in an ACO and the fraction of individual providers in an ACO that are PCPs.

Table 1.7 gives an outline of the identification strategy for these parameters.

Denote the instruments and exogenous variables in X_j^{part} as the vector Z_j^{part}. The moment

[18]Estimates presented in Table 1.9 ultimately indicate that this is not an issue.

Table 1.7: **Endogenous Variables and Instruments**

Parameter	Endogenous Variable	Instrument
α_0	$\hat{y}_j$ (net income)	Element in X_j^{perf}
$\boldsymbol{\beta}$	# states	Total Medicare beneficiary person-years in ACO area
	average risk score	Average risk score of non-dual beneficiaries in ACO area
	% over 75	Percent of population over 75 in ACO area
	% male	Percent of male population with Medicare in ACO area
	% nonwhite	Percent of population non-white in ACO area
	# providers	Element in X_j^{perf}
	fraction PCP	Element in X_j^{perf}
ρ	$\ln(a_j/a_d)$ (nesting term)	Relative HMO enrollment in ACO area

This table shows endogenous variables in the utility from participation function and the variables' associated instruments used for estimation.

condition for estimation is

$$\mathbb{E}\left[\hat{\xi}_j \middle| Z_j^{part}\right] = 0 \tag{1.50}$$

where $\hat{\xi}_j$ is the same as Equation 1.49, but with $\hat{y}_j$ instead of y_j.[19] I follow the contraction mapping approach from Berry et al. (1995) to estimate $\boldsymbol{\theta}_1 = \{\alpha_0, \alpha_\eta, \boldsymbol{\beta}, \rho\}$.

1.5 Estimation Results

The estimated cost function parameters, $\hat{\boldsymbol{\theta}}_2$, are presented in Table 1.8. The three parameters controlling the shape of the cost function are estimated precisely, and the resulting cost

[19] In order to account for uncertainty introduced by using estimates from the second stage, the standard errors of the parameters estimated in the first stage must be adjusted. I achieve this via bootstrapping. Nonetheless, this issue is small when the estimated component of $\hat{y}_j$ is small and the parameter estimates in $\hat{\boldsymbol{\theta}}_2$ are precise, which are both true in this application. See Table 1.8 and Figure 1.6.

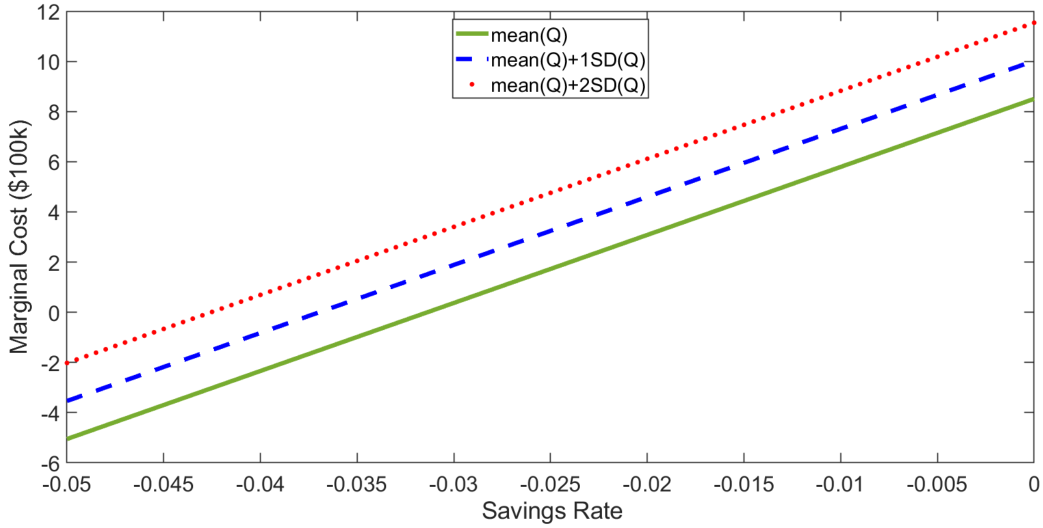
Table 1.8: **Cost Function Parameter Estimates**

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

Coef.	Variable	Estimate	Std. Err.	P-value	95% CI	
	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
$\boldsymbol{\gamma}_S$	# states	4.460	2.067	0.031	0.811	7.545
	# beneficiaries	0.210	0.143	0.142	0.000	0.462
	average risk score	116.170	38.847	0.003	36.326	166.970
	% over 75	0.131	0.242	0.587	-0.283	0.505
	% nonwhite	-0.124	0.061	0.041	-0.225	-0.021
	% male	1.097	1.082	0.310	-0.573	2.850
	# providers	-2.195	2.017	0.276	-5.512	0.780
	fraction PCP	5.292	7.131	0.458	-6.010	16.738
	fraction inpatient	-134.000	60.540	0.027	-232.780	-30.566
	fraction outpatient	-113.930	30.472	0.000	-163.880	-62.070
	# PC services	-4.279	1.335	0.001	-6.241	-1.816
	# admissions	-60.862	27.000	0.024	-101.780	-11.279
	fraction PC served by PCP	6.711	10.569	0.525	-11.044	23.293
	all group	29.339	5.204	0.000	21.085	38.347
$\boldsymbol{\gamma}_Q$	# states	-5.031	3.003	0.094	-9.541	0.024
	# beneficiaries	0.327	0.198	0.099	-0.021	0.641
	average risk score	0.011	0.011	0.318	-0.004	0.032
	% over 75	8.352	3.910	0.033	1.985	14.895
	% nonwhite	0.009	0.017	0.594	-0.020	0.037
	% male	-0.004	0.004	0.294	-0.010	0.003
	# providers	0.055	0.081	0.497	-0.067	0.192
	fraction PCP	-0.152	0.153	0.320	-0.429	0.057
	fraction inpatient	0.403	0.475	0.396	-0.345	1.190
	fraction outpatient	-10.024	4.461	0.025	-17.659	-3.337
	# PC services	-6.613	2.015	0.001	-9.979	-3.533
	# admissions	-0.298	0.104	0.004	-0.465	-0.123
	fraction PC served by PCP	-4.148	2.426	0.087	-8.299	-0.201
	all group	0.598	0.680	0.380	-0.477	1.751
	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.

Figure 1.5: **Marginal Cost vs. Savings Rate**

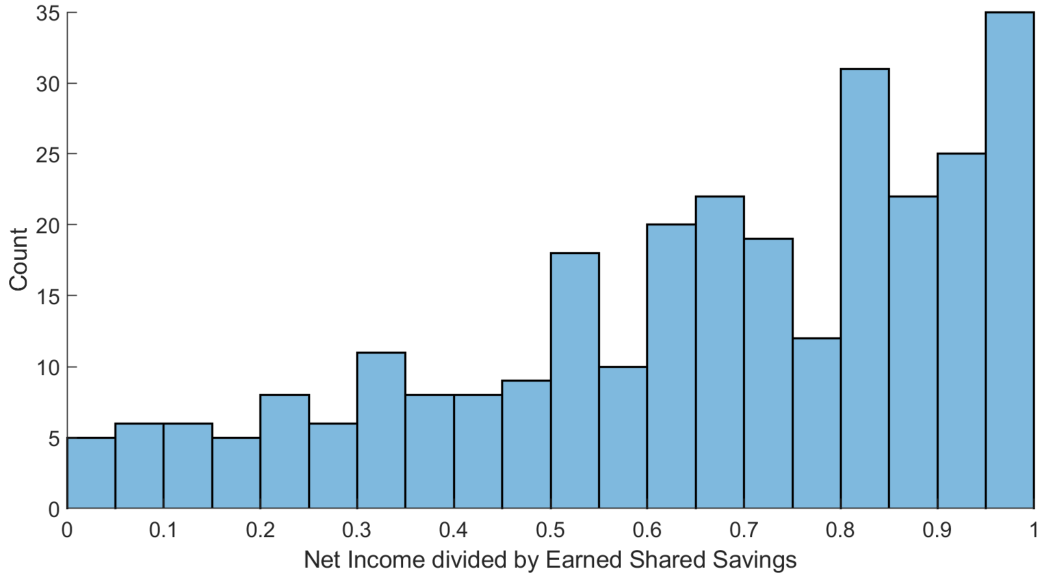


Note: This figure shows the marginal cost of savings rate as a function of savings rate. At higher quality scores, savings becomes more expensive to increase, so the marginal cost of savings shifts upwards.

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. Figure 1.5 plots the marginal cost of savings as a function of savings. Like the top panel in Figure 1.4, it shows that as the quality score increases, in this case by one and two standard deviations, the marginal cost of savings increases. Other parameter estimates in Table 1.8 indicate several determinants of the marginal cost of savings and quality.

Using the estimated cost function, I compute net income $\hat{y}_j$. In Figure 1.6, I show net

Figure 1.6: **Histogram of Net Income divided by Earned Shared Savings**



Note: This figure is a histogram of an ACO's net income $\hat{y}_j$ as a fraction of earned shared savings. Roughly one third of ACO's have net income that's less than half of their earned shared savings.

income $\hat{y}_j$ as a fraction of earned shared savings. The figure shows substantial variation in the margins of ACOs. The average ACO looses 34% of their earned shared savings to increases in effort cost, with some ACOs barely breaking even.

Finally, I present the results to estimation of the participation equation in Table 1.9. The first column of estimates is of the OLS logistic regression without IVs. The Random Coefficients (non-nested) Logit with IVs is in the second column (RC) and the Random Coefficients Nested Logit with IV is in the third column (RCNL). After accounting for endogeneity, both models estimate a significant response of ACO participants to ACO net income. A $100,000 increase in ACO net income increases the amount of participants in an ACO by over 7%, all else constant. This is an increase in two to three participants for the average ACO. The elasticity of participation with respect to net income is 0.5.

The parameter α_η, which describes the dispersion of taste for net income, has a precise

Table 1.9: **Participation Equation Estimates**

$$u_{ij} = (\alpha_0 + \alpha_\eta \eta_i)\hat{y}_j + \boldsymbol{\beta}' X_j^{part} + \xi_j + \zeta_{id}(\rho) + (1-\rho)\epsilon_{ij}$$

Coefficient	Variable	OLS	RC	RCNL
α_0	$\hat{y}_j$ (net income)	-0.007*	0.076**	0.072**
		(0.003)	(0.028)	(0.022)
α_η	$\eta_i \hat{y}_j$ (net income×pref. shock)		0.012***	0.014
			(0.002)	(0.009)
ρ	$\ln(a_j/a_d)$ (nesting term)			0.544*
				(0.227)
$\boldsymbol{\beta}$	# states	-0.045	0.021	-0.014
		(0.039)	(0.071)	(0.073)
	# beneficiaries	0.026***	0.026+	0.023+
		(0.003)	(0.014)	(0.013)
	average risk score	2.137***	-0.306	-0.850
		(0.567)	(0.866)	(0.723)
	% over 75	0.027***	0.135***	0.113***
		(0.007)	(0.028)	(0.028)
	% nonwhite	0.028***	0.045***	0.034***
		(0.002)	(0.009)	(0.008)
	% male	0.100***	0.275**	0.160***
		(0.017)	(0.095)	(0.039)
	# providers	0.134*	-0.457	-0.791*
		(0.060)	(0.414)	(0.366)
	fraction PCP	0.249	-2.494	-3.268*
		(0.201)	(1.771)	(1.439)
N		1486		

+ $p < 0.10$; * $p < 0.05$; ** $p < 0.01$; *** $p < 0.001$

Bootstrapped standard errors (1,000 rep.) in parentheses. Estimates include year and Census Division FE. $\hat{y}_j$ is in units of \$100k.

estimate of 0.012 in the RC model and an imprecise estimate of 0.014 in the RCNL model. These estimates imply that the proportion of providers in $\mathcal{I}$ without utility increasing in net income is $1 - \Phi(\alpha_0/\alpha_\eta) \approx 0$, where Φ is the standard normal CDF. In the RCNL estimation, the nesting parameter ρ is estimated with some precision at 0.544. Given the definition of nests d as leadership types of ACOs (hospital, physician, or mixed), this means the correlation of utilities of participants in ACOs under similar leadership is fairly high. Management structure of an ACO plays an important role in a participant's utility. In a related study, McWilliams et al. (2016) discusses the role ACO leadership with regards to ACO performance.

1.6 Counterfactuals

In this section, I use the estimated model of participation and performance to evaluate the available contracts between ACOs and CMS. Each counterfactual follows the following steps:

1. Predict ACO outcomes S_j^{CF} and Q_j^{CF} using the shared savings function $R_j^{CF}(S_j, Q_j)$, where CF denotes the counterfactual policy or behavioral assumptions.

2. With S_j^{CF} and Q_j^{CF}, compute net income $\hat{y}_j^{CF}$ under the counterfactual policy, and use that to compute changes in participation.

3. To account for ACO exit under the the counterfactual, I estimate a logit with dependent variable equal to one if an ACO exits in performance year $t + 1$. I use estimates from this model to predict ACO exit under counterfactual scenarios. Formally, I estimate

ν_0, ν_1, ν_2, ν_3, and $\boldsymbol{\psi}$ in

$$exit_{jt+1} = \mathbf{1}\left\{\nu_0 + \nu_1 \hat{y}_{jt} + \nu_2 \mathbf{1}\left\{\hat{y}_{jt} > 0\right\} + \nu_3 age3_{jt} + \boldsymbol{\psi}' X_{jt}^{perf} + \varepsilon_{jt+1}\right\} \tag{1.51}$$

where $age3_{jt}$ indicates the ACO is three years old and would start a new agreement with CMS in year $t+1$, and $\varepsilon_{jt+1} \sim \text{Logistic}(0,1)$. Results for this are in Table A.1 in Appendix A.4.[20]

While participation change in ACOs and ACO exit are predicted, I do not predict which or how many ACOs would *not enter* (as opposed to exit) in the face of the counterfactual changes. Nonetheless, if several ACOs exit under counterfactual changes, it's likely a portion of those exiting ACOs would have never entered in the first place, so the predictions are still informative.

To put counterfactual changes in ACO overall quality score into meaningful terms, I use regularized regressions from Machine Learning. As discussed in Section 2.4, an ACO's overall quality score is a composite measure of 30 to 40 sub-measures. The actual method of computing overall quality score from these sub-measures is discontinuous, unintuitive, and presents no method to find the marginal effect of a sub-measure on overall quality score. For this reason, I instead estimate a very simple model of overall quality score, where it is merely a linear combination of quality sub-measures. Techniques known as the Least Absolute Shrinkage and Selection Operator (LASSO) and Elastic Net automatically select a subset of sub-measures that explain the most variation in overall quality score. While a simplification, this model still explains over 80% of the variation in overall quality score.

[20] In short, estimates of Equation 1.51 indicate ACOs with non-negative net income are 15 percentage points less likely to exit. No element in X_{jt}^{perf} is a significant predictor of ACO exit.

Table 1.10: **ACO Contract Options**

Proportion of Savings Shared	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.

Appendix A.5 details the process and results.

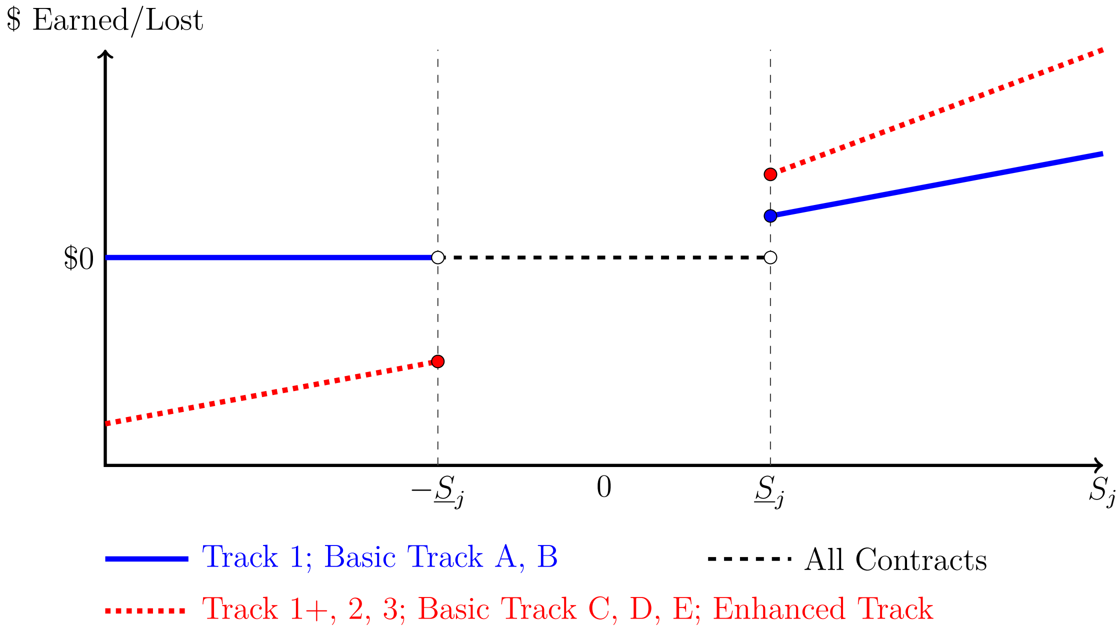
1.6.1 Evaluating Existing Contracts between Medicare and ACOs

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 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 1.10 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 1.7 shows how the various contracts differ in power by graphing an ACO's earned

Figure 1.7: **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 $-\underline{S}_j$.

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

$$R_j(S_j, Q_j) = \begin{cases} F \cdot B_j S_j Q_j & \text{if } S_j \geq \underline{S}_j \text{ and } Q_j \geq \underline{Q} \\ 0 & \text{otherwise} \end{cases} \tag{1.52}$$

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

Table 1.11: **ACO Performance Predictions**

	One-Sided Risk Model					Two-Sided Risk Model					
F	S_j^*	Q_j^*	PQ	PF	PS	S_j^*	Q_j^*	PQ	PF	PS	SL
0.25	-0.0088	0.8843	0.3351	0.3950	-0.1432	0.0117	0.7167	0.3338	0.0168	0.3971	0.0919
0.30	-0.0048	0.8867	0.3526	0.3869	-0.0301	0.0150	0.7211	0.3546	0.0168	0.4883	0.0919
0.40	0.0023	0.8920	0.3863	0.3769	0.1096	0.0213	0.7293	0.3937	0.0168	0.5837	0.0919
0.50	*0.0078*	*0.8953*	*0.4118*	*0.3668*	*0.1182*	0.0317	0.7660	0.4468	0.0249	0.7408	0.1385
0.50^d	*0.0084*	*0.8840*	*0.3201*	*0.2011*	*0.0919*	N/A	N/A	N/A	N/A	N/A	N/A
0.60	0.0137	0.8998	0.4388	0.3580	0.0354	0.0360	0.7987	0.4704	0.0511	0.5340	0.1911
0.75	0.0202	0.9046	0.4623	0.3520	-0.2984	**0.0375**	**0.7620**	**0.4643**	**0.0289**	**0.0403**	**0.1136**

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, Q_j^* is average ACO quality score, PQ is the proportion of ACOs that qualify for shared savings ($S_j^* \geq \underline{S}_j$), and PF is the proportion of ACOs with savings rare below minimum loss rate ($S_j^* \leq -\underline{S}_j$). PS is total program savings in \$ billions, defined in Equation 1.54. SL is total shared losses paid to CMS in \$ billions. 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.

is invariant to contract changes). For the following predictions, two-sided ACOs have the shared savings formula

$$R_j^{TS}(S_j, Q_j) = \begin{cases} FQ_j \cdot B_j S_j & \text{if } S_j \geq \underline{S}_j \text{ and } Q_j \geq \underline{Q} \\ (1 - FQ_j) \cdot B_j S_j & \text{if } S_j \leq -\underline{S}_j \\ 0 & \text{otherwise} \end{cases} \tag{1.53}$$

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.[21]

Table 1.11 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

[21] 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.

$F = 0.50^d$ 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 $\underline{S}_j$.

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 PQ and PF in Table 1.11, 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.[22] For a fixed F, however, ACOs facing one-sided 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. According to the Elastic Net results described in Appendix A.5, these changes in average overall quality score from 0.90 under one-sided risk to between 0.70 and 0.80 under two sided risk amount to one of the following:

1. The percentile of all ACO providers COPD/Asthma emergency admissions increasing by four to eight percentage points (e.g., from 5th percentile to 9th percentile).

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

2. The percentile of all ACO providers Heart Failure emergency admissions increasing by 6.5 to 13 percentage points.

3. The percentile of all ACO providers score on Health Promotion and Education portion of CAHPS decreasing by 25 to 50 percentage points.

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 SL in Table 1.11 is the amount paid to CMS by ACOs that fail to save above the minimum loss rate. Column PS 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%.

Table 1.12 shows changes in ACO net income, participation, and exit under different contracts. When switching from one-sided to two-sided incentives and holding the sharing rate fixed, mean net income decreases by less than \$100,000 on average, and mean participation decreases between 3% and 10%. When the sharing rate increases along with the change to a two-sided risk structure (from Track 1 to Track 3, for example), the effect on participation is a net positive.

Table 1.12: **Two-Sided ACO Net Income and Participation**

	One-Sided Risk Model			Two-Sided Risk Model		
F	Net Income	# of Participants	# of Exiting ACOs	Net Income	# of Participants	# of Exiting ACOs
0.25	1.891	21.017	37.152	1.003	19.644	45.208
0.30	2.579	22.146	34.995	1.685	20.691	41.213
0.40	4.204	25.059	28.817	3.326	23.440	35.004
0.50	*6.117*	*28.982*	*25.818*	5.815	28.323	25.933
0.50^d	*N/A*	*34.077*	*25.355*	N/A	N/A	N/A
0.60	8.313	34.248	17.913	8.067	33.614	21.178
0.75	11.892	44.962	12.564	**10.895**	**41.679**	**21.105**

This table shows model simulations for various ACO contract options. F is the proportion of savings shared with an ACO. Net income is in units of \$100k. The superscript d indicates values observed in data. Italicized numbers are performance statistics under estimation sample. Bold numbers are the models predictions for Track 3/Enhanced Track ACOs.

1.6.2 Solving for Contracts between ACOs and Medicare

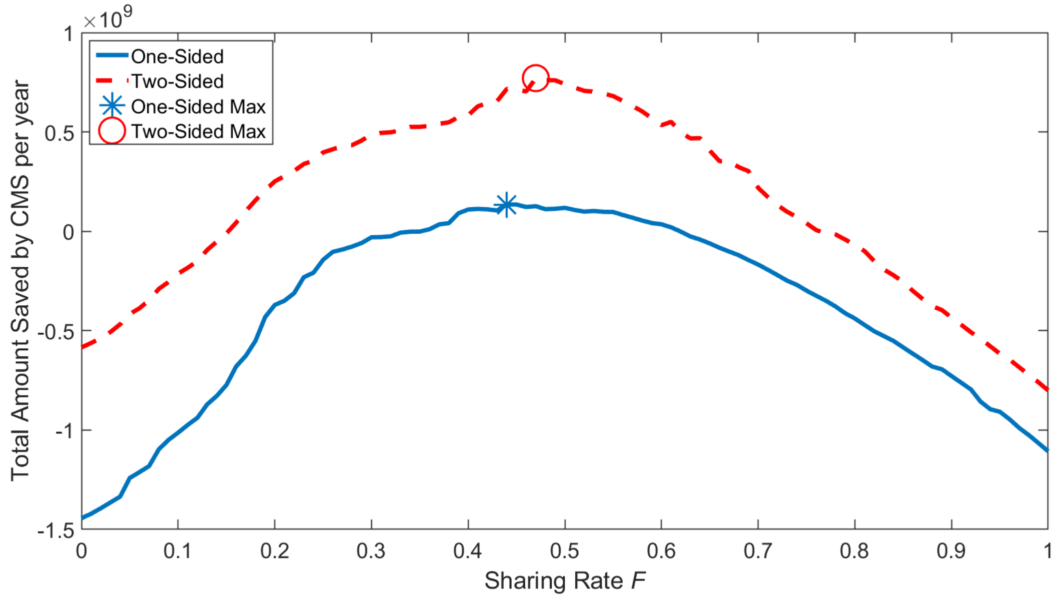
The results in the previous section indicate that the proportion of savings shared with an ACO and the presence of downside risk both play a large role in determining the success of the Medicare Shared Savings Program. To find the contract that maximizes savings of the MSSP, I compute the *savings-optimal*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

$$\max_{F\in[0,1]} \sum_{j\in\mathcal{J}} \left\{ \overbrace{B_j S_j^*(F)}^{\$\text{ saved by ACO } j} - \overbrace{F \cdot B_j S_j^*(F) Q_j^*(F) \mathbf{1}\left\{S_j^*(F) \geq \underline{S}_j\right\} \mathbf{1}\left\{Q_j^*(F) \geq \underline{Q}\right\}}^{\$\text{ paid to ACO } j} \right\} \tag{1.54}$$

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

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

Figure 1.8: **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). When ACOs do not pay shared losses, program savings is maximized at a sharing rate of 0.44 and program savings of $135 million per year. If every ACO pays shared losses, program savings is maximized at a sharing rate of 0.47 and program savings of $770 million per year.

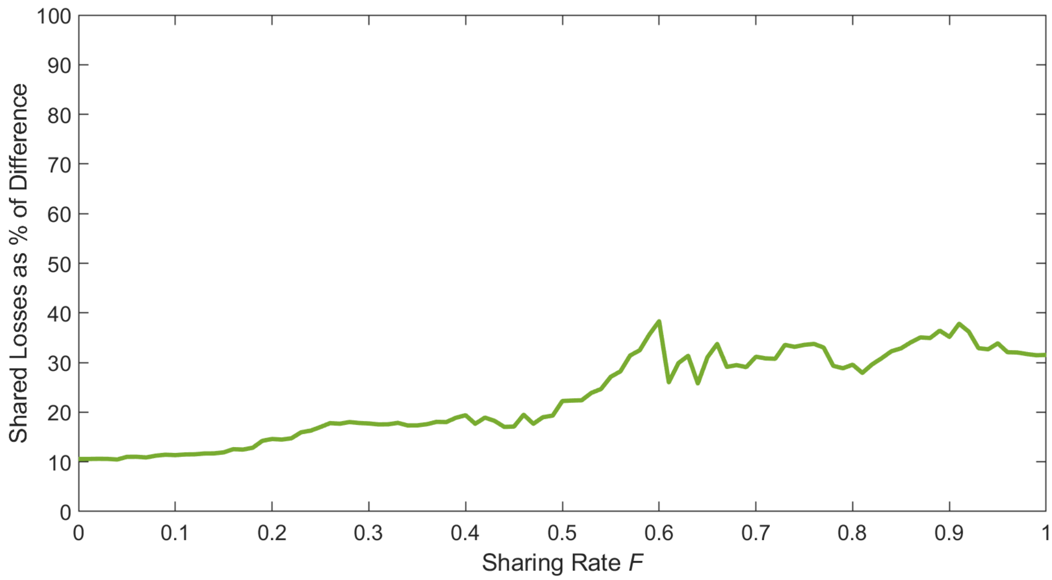
imposed, ACO participants choose savings and quality efforts from a Nash equilibrium.[23]

The solid line of Figure 1.8 plots the objective function of CMS when maximizing total savings with one-sided ACOs (Equation 1.54), and the dashed line plots the objective function of CMS with two-sided ACOs.[24] 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,

[23]Note that I am not imposing participation constraint. Because the optimization is over just one value of F for all ACOs, imposing a participation constraint would yield a solution that is determined only by the most constrained participant, and is therefore uninformative.

[24]CMS's objective for two-sided ACOs is slightly different than Equation 1.54 and includes an extra term for shared losses paid back to CMS.

Figure 1.9: **Shared Losses**



Note: This figure shows model simulations for shared losses paid the as a percent of the difference in program savings between one-sided and two-sided contracts for various sharing rates. At each sharing rate, less than 50% of the increase program savings from two-sided incentives comes from penalties for overspending.

the amount saved at $F^* = 0.47$ is \$730 million larger.[25]

Figure 1.9 plots the shared losses paid by ACOs as a percent of the difference in program savings between one-sided and two-sided incentives. As F increases, shared losses increase in its share of the difference between one-sided and two-sided program savings. Shared losses comprise 40% of the difference between one-sided and two-sided incentives at most.

The objective function in Equation 1.54 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.

[25] 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.

Formally, the problem is

$$\max_{F\in[0,1]} \sum_{j\in\mathcal{J}} \left\{ \left[B_j S_j^*(F) - F \cdot B_j S_j^*(F) Q_j^*(F) \mathbf{1}\left\{ S_j^*(F) \geq \underline{S}_j \right\} \mathbf{1}\left\{ Q_j^*(F) \geq \underline{Q} \right\} \right] Q_j^*(F) \right\} \quad (1.55)$$

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

when $S_j^* \geq 0$ and

$$\max_{F\in[0,1]} \sum_{j\in\mathcal{J}} \left\{ \left[B_j S_j^*(F) - F \cdot B_j S_j^*(F) Q_j^*(F) \mathbf{1}\left\{ S_j^*(F) \geq \underline{S}_j \right\} \mathbf{1}\left\{ Q_j^*(F) \geq \underline{Q} \right\} \right] \left[1 - Q_j^*(F) \right] \right\} \quad (1.56)$$

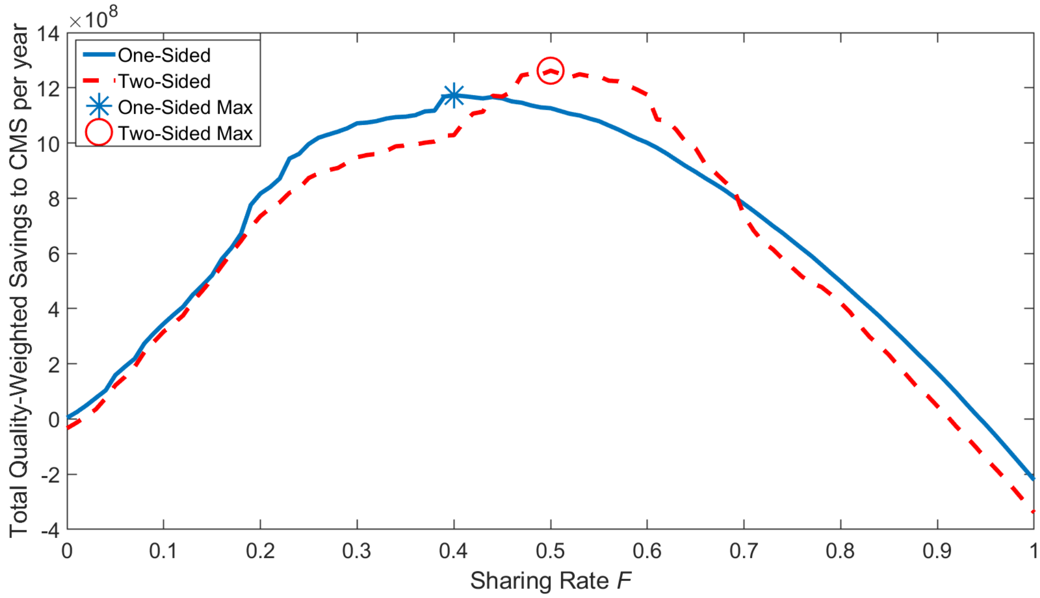
$$\text{s.t. } \left(s_{ij}^*(F), q_{ij}^*(F) \right) = \arg\max_{s_{ij}, q_{ij}} \pi_{ij}\left(s_{ij}, q_{ij} \right) \text{ for all } i \in I_j \text{ and } j \in \mathcal{J}.$$

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

Figure 1.10 plots the objective function of CMS when maximizing total savings weighted by quality score with one-sided ACOs (Equations 1.55 and 1.56) 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

Figure 1.10: **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 ACOs do not pay shared losses, quality-weighted program savings is maximized at a sharing rate of 0.40. If every ACO pays shared losses, quality-weighted program savings is maximized at a sharing rate of 0.50. Neither contract strictly dominates the other for this metric.

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.

1.6.3 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

$$\max_{\mathbf{s}_j,\mathbf{q}_j} R_j\left(S_j,Q_j\right)-\sum_{i\in I_j} c_{ij}\left(s_{ij},q_{ij}\right). \tag{1.57}$$

Table 1.13: **Performance Loss from Non-Cooperative Behavior**

	One-Sided Risk Model w/ Strategic Behavior					One-Sided Risk Model w/Perfect Coordination				
F	S_j^*	Q_j^*	PS	Net Income	# Part.	S_j^*	Q_j^*	PS	Net Income	# Part.
0.25	-0.0088	0.8843	-0.1432	1.891	21.017	-0.0088	0.8843	-0.1432	2.216	21.542
0.30	-0.0048	0.8867	-0.0301	2.579	22.146	0.0299	0.9098	1.6643	3.117	23.070
0.40	0.0023	0.8920	0.1096	4.204	25.059	0.0392	0.9168	1.6687	5.208	27.047
0.50	0.0078	0.8953	0.1182	6.117	28.982	0.0466	0.9239	1.4698	7.605	32.454
0.60	0.0137	0.8998	0.0354	8.313	34.248	0.0540	0.9286	1.2576	10.200	39.533
0.75	0.0202	0.9046	-0.2984	11.892	44.962	0.0623	0.9360	0.6744	14.583	55.172

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. PS is total program savings in $ billions, defined in Equation 1.54. Net Income is in $100k.

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 1.13. 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. This amounts to more than $1 billion per year in additional program savings to CMS.

Table 1.13 also indicates that as coordination increases, the incentives imposed by Medicare should be weakened. While 44% of savings is the optimal amount to share with free-riding, under perfect coordination the optimal amount is 35%.

1.7 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.

I estimate a two-stage structural model of participation and performance in ACOs. I find Medicare providers respond to the income they expect to earn from an ACO, and participation is increasing in the amount an ACO earns. Second, I find that provider 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.

Chapter 2

Spillovers between Medicare and Medicaid: Evidence from the Supply-Side and Payment Parity

2.1 Introduction

Increasing access to health care was a main objective of the Patient Protection and Affordable Care Act of 2010 (ACA). The ACA expanded Medicaid coverage to an additional 17 million people, and recent research is generally very positive towards the impact of the expansion: more people received care, health outcomes improved, and there is little evidence that non-Medicaid populations were disadvantaged as a result of spillover (Carey, Miller, & Wherry, 2018; Alexander & Schnell, 2019; Miller, Altekruse, Johnson, & Wherry, 2019). However, these results come despite research concerning non-ACA institutional changes. Several studies (Garthwaite, 2012; McInerney, Mellor, & Sabik, 2017; Glied & Hong, 2018) find an exogenous increase in access or demand for health care among a certain group causes a decrease in the amount of health care provided to other groups.

Why did the access-increasing provisions of the ACA have a different impact, specifically regarding spillover to other populations, than non-ACA provisions? To answer this question, this paper studies changes in Medicare provider service and patient volumes in response to the payment parity provision in the ACA that increased Medicaid reimbursement significantly. I find that while physicians slightly decreased the total volume of care they provide to Medicare beneficiaries when facing increased Medicaid fees, the particular combination of services physicians provided to Medicare beneficiaries changed drastically.

In 2013 and 2014, all states with fee-for-service Medicaid programs received federal funding to increase payments to physicians for Medicaid services so that they are equal to payments to physicians for Medicare services. This nearly doubled the payment physicians received from Medicaid. Along with the contemporaneous expansion of Medicaid coverage, this so-called Medicaid "fee bump" increased the access to care and the amount of care received by Medicaid beneficiaries, and further improved health and behavioral outcomes (Maclean, McClellan, Pesko, & Polsky, 2018; Alexander & Schnell, 2019). Focusing on Medicare beneficiary utilization, Carey et al. (2018) finds no crowd out and that increasing access to care to Medicaid beneficiaries did not diminish the care received by Medicare beneficiaries.

In this paper, I take a new approach to identify spillover to Medicare by looking at *provider* outcomes. Focusing on providers allows me to exploit exogenous variation in Medicaid payment that varies across providers, along with across states and time. This permits identification of a causal effect by comparing the outcome of the same provider across time, controlling for all provider and state-year-specific heterogeneity. Furthermore, by examining provider outcomes, I can detect changes in the *service-level* provision of care to Medicare beneficiaries that may result from changes in relative service prices.

I find the existence of small overall spillover between Medicaid and Medicare. When providers are in a state that has increased Medicaid fees, they treat 0.315 (± 0.10) percent fewer Medicare beneficiaries, and receive 0.396 (± 0.13) percent less payment from Medicare. After weighting treatment by the Medicaid-to-Medicare fee ratio, the effect is a decrease in number of Medicare beneficiaries treated by 0.244 (± 0.20) percent and a decrease in Medicare payment by 1.536 (± 0.26) percent.

Focusing on the use of specific service codes reveals a moderate to large impact of the parity provision on provider's behavior. Physicians in a state that have increased Medicaid fees treat 7.248 (± 0.35) percent fewer Medicare beneficiaries with established patient services, but 1.077 (± 0.48) percent more Medicare beneficiaries with new patient services. Medicare payment from established patient services decreased by 13.39 (± 0.61) percent, and increased by 2.626 (± 1.01) percent from new patient services.

Related Literature. Recent research has generally been favorable towards both the Medicaid expansion and the accompanying increase in Medicaid payments to physicians. For example, Miller et al. (2019) find a 0.132 percentage point decrease in mortality in states that expanded Medicaid, implying roughly 15 thousand people died between 2014 and 2017 because some states did not expand Medicaid eligibility. Concerning the fee bump specifically, Maclean et al. (2018) finds that higher Medicaid reimbursement rates improved behavioral health outcomes, specifically substance use disorders and tobacco product use. Alexander & Schnell (2019) shows that higher reimbursement for Medicaid services decreased reports of providers turning away patients, increased office visits for Medicaid enrollees, and improved overall health.

An important critique of the Medicaid expansion is that increased service use by Medicaid beneficiaries may spill over and impact the access to and amount of care individuals covered by other insurers receive. Garthwaite (2012) applies the seminal model of Sloan, Mitchell, & Cromwell (1978) to deduce that due to crowd out of private insurance, upon implementation of the State Children's Health Insurance Program, which expanded insurance coverage to low-income Americans below the age of 19, physicians that serve few Medicaid patients under 19 should decrease the quantity of medical services provided. Survey data from physicians confirm this hypothesis. In a similar analysis using a survey of beneficiaries, McInerney et al. (2017) find that a one percent increase in the Medicaid-eligible population causes a decrease in spending among Dual Eligible patients. Finally, Glied & Hong (2018) find that factors increasing demand in the non-Medicare population cause a decrease in Medicare utilization and spending, and the total quantity of services provided did not change.

Contrary to Garthwaite (2012) and McInerney et al. (2017), Carey et al. (2018) finds no evidence of spillover between Medicare and Medicaid. Using a large sample of Medicare claims, the authors find that Medicaid expansions did not reduce utilization among Medicare beneficiaries. Furthermore, Maclean et al. (2018) and Alexander & Schnell (2019) find no change in behavioral health and access to care among non-Medicaid populations due to the Medicaid fee bump in secondary analyses.

This study provides new evidence surrounding physicians' responses to changes in payment for health services and their ability to expand the capacity of their practices. While Medicare beneficiaries did not see a large decrease in care due to the Medicaid fee bump, providers decisions surrounding Medicare beneficiaries were influenced, nonetheless. I justify these findings by extending the mixed-economy model of Sloan et al. (1978). The empirical results of this paper are consistent with a model of physician decision-making where marginal revenue (or equivalently marginal cost) is heterogeneous across patients, and the Medicaid fee bump simultaneously caused a crowed out of low-marginal revenue patients and a decrease in the marginal cost of care.

2.2 Institutional Background

Medicaid has historically reimbursed providers at a far lower rate than Medicare and private insurance. According to Zuckerman & Goin (2012), the Medicaid-to-Medicare fee ratio averaged 59% for primary care services in 2012, and half of all states had a fee ratio below 70%. This was a central concern to policymakers when expanding Medicaid eligibility in the ACA. While Medicaid eligibility would be expanded to an additional 17 million people in 2013, providers may still decide to not treat them, given they earn nearly two times the pay for providing the same service to Medicare or privately insured beneficiaries. Accordingly, the so-called Medicaid "fee bump" was included in the ACA.[1] The provision mandates that in 2013 and 2014, for 146 primary care services and for certain provider taxonomies, the payment for those services from Medicaid will be 100% of the Medicare rate. This provision applied to every state's Medicaid program, regardless of the state's decision to expand Medicaid.[2] The federal government paid states for the full costs of the fee bump until December 2014. From 2015 onwards, states had the decision to continue the fee bump, and as of 2016, 19 states fully or partially continued increased rates, and 30 states plus Washington D.C.

[1] See Section 1202 of the Act.

[2] The exception to this is Tennessee, which does not have a Medicaid FFS program.

Table 2.1: **Procedure Code Examples**

HCPCS Code	Description	Medicaid Fee Bumped?
99202	New patient office or other outpatient visit, typically 20 minutes	Yes
99233	Subsequent hospital inpatient care, typically 35 minutes per day	Yes
99291	Critical care delivery critically ill or injured patient, first 30-74 minutes	Yes
17003	Destruction of 2-14 skin growths	No
36415	Insertion of needle into vein for collection of blood sample	No
73564	X-ray of knee, 4 or more views	No

Note: This table gives examples of services that were and were not affected by the payment parity provision of the ACA.

elected to decrease Medicaid reimbursement rates to pre-bump levels.

Increased fees were given only to certain types of providers providing certain types of services to Medicaid patients. In particular, a service provided to a Medicaid beneficiary received the Medicare payment rate if the physician rendering or supervising the service had specialty designation of family medicine, general internal medicine, pediatric medicine, a subspecialty within these designations, or at least 60% of services provided by the physician in the previous year were among the services qualifying for the fee bump. The services qualifying for the fee bump have Healthcare Common Procedure Coding System (HCPCS) Level I codes 99201-99499, 90460, 90461, and 90471-90474.[3] The first range of HCPCS codes (99201-99499) are for Evaluation and Management Services, and the others are for Vaccine Administration. Table 2.1 contains some examples of services with and without bumped fees. There are over 6000 unique HCPCS codes.

A few states opted to increase rates for all physician types, rather than just those speci-

[3]HCPCS Level I codes are also known as Current Procedural Terminology (CPT) codes.

fied by the ACA: Maryland did so from 2013 through 2016, and Colorado, Idaho, Indiana, Nevada, and Utah did so in 2016.

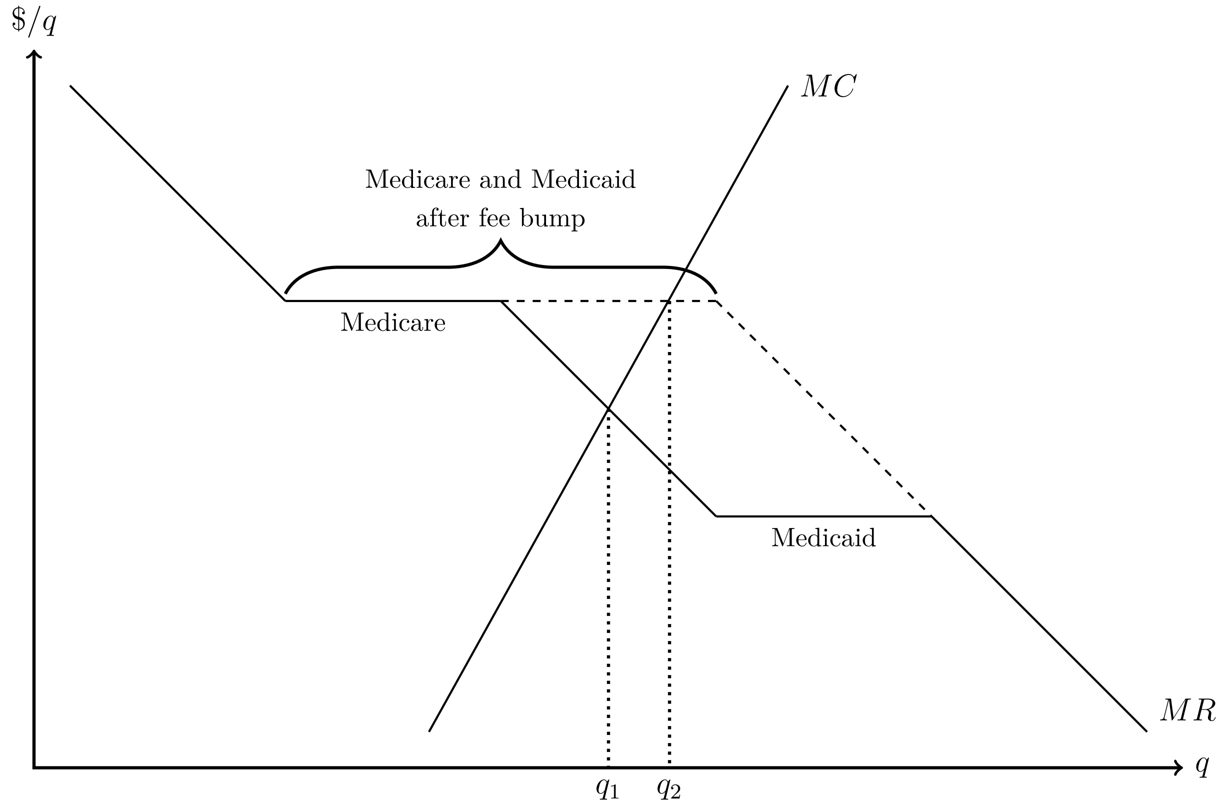
2.3 Conceptual Framework

The seminal work by Sloan et al. (1978) guides my predictions regarding the impact of the fee bump on provider decisions. In this model, a physician's marginal revenue MR is weakly decreasing in service volume q. In Figure 2.1, the flat segments represent the fee-for-service payments made by Medicare and Medicaid to physicians, and the downward sloping segments represent marginal revenue for services provided to privately insured individual and uninsured individuals. The dashed line is a provider's marginal revenue when Medicaid fees are bumped. Marginal costs are depicted by MC, and the fee bump increases service volume from q_1 to q_2. In this setting, physicians are indifferent between Medicare and Medicaid patients after the fee bump, so some Medicare patients may be replaced with Medicaid patients.

Figure 2.2 depicts a slightly more detailed framework, where Medicare and Medicaid patients differ by falling into the "new" and "established" patient categories. Certain evaluation and management services contain designations for "new" patients and "established" patients.[4] Because new patient services pay 30% more, it is more likely that a Medicare patient would be replaced with a Medicaid patient if the Medicare patient is established the Medicaid patient is new. If, for example, a physician has high marginal costs, a physician

[4] According to the American Medical Association, a new patient is "one who has not received any professional services from the physician, or another physician of the same specialty who belongs to the same group practice, within the past three years."

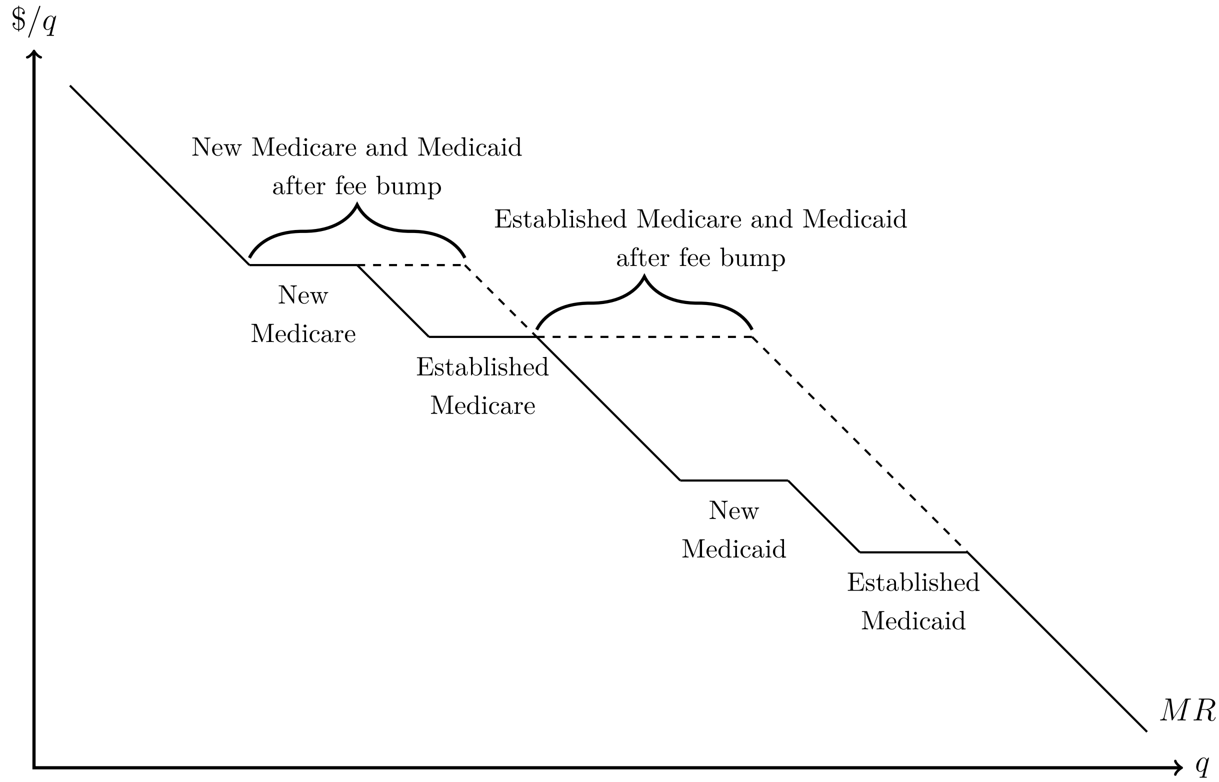
Figure 2.1: **Provider Response to Fee Bump**



Note: This figure shows the marginal revenue (MR) of a physician for units of service q. The Medicaid fee bump increases marginal revenue for a portion of the domain of q, and equilibrium service volume increases from the intersection of MR with marginal cost (MC) at q_1 to q_2.

that initially sees a combination of new and established Medicare patients would replace its established Medicare patients with new Medicaid patients after the fee bump. Figure 2.3 details this. Before the fee bump, q_3 is the optimal level of care for physician with marginal costs MC. A physician sees both new and established Medicare beneficiaries, and does not see Medicaid beneficiaries. After the fee bump, no established patients are treated, physicians see all new patient possible, and total amount of care provided increases to q_4.

Figure 2.2: **Provider Marginal Revenue with New and Established Patient Services**

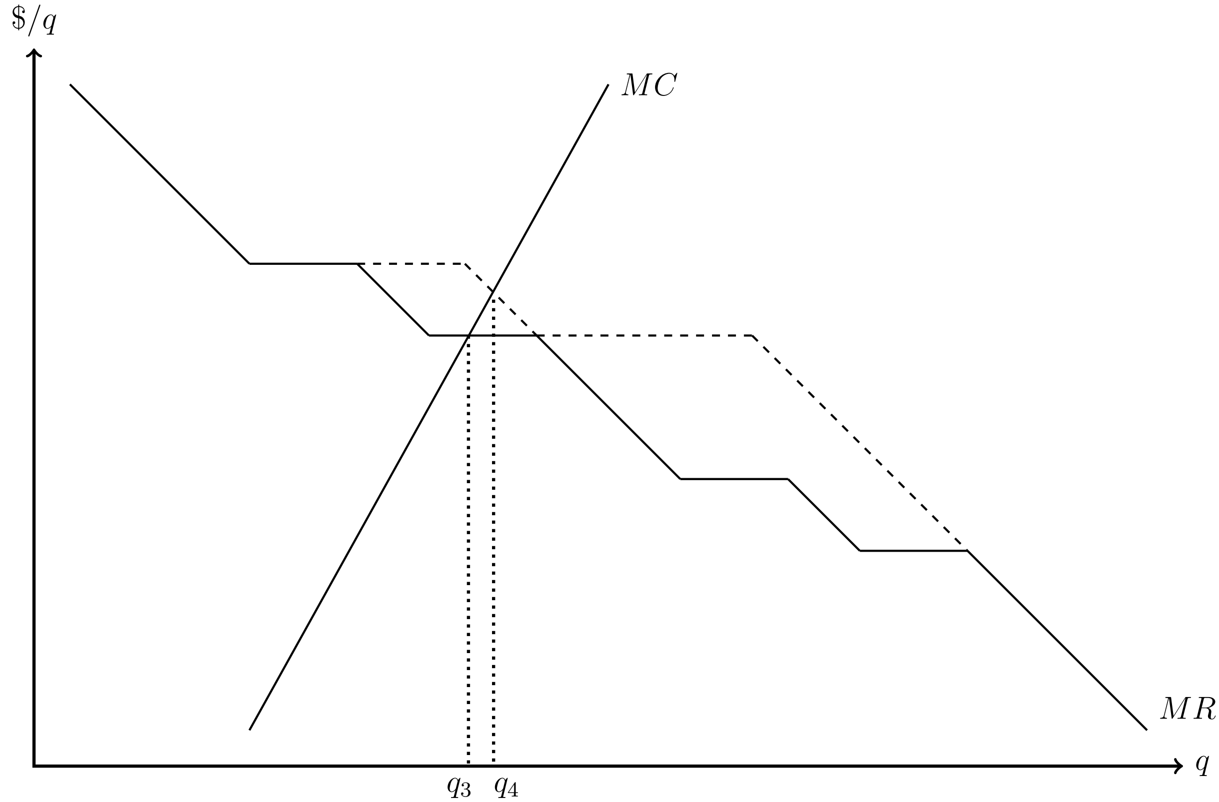


Note: This figure shows the marginal revenue (MR) of a physician for units of service q, where new Medicare and Medicaid patients offer higher marginal revenue than established Medicare and Medicaid patients. The fee bump increases marginal revenue for all Medicaid beneficiaries, and this is represented by the dashed line.

2.4 Data

The purpose of this study is to find the impact of changing Medicaid reimbursement rates on the decisions of Medicare providers. To do so, the main source of data is the Physician and Other Supplier data (henceforth "Service Level data") from the Centers for Medicare and Medicaid Services (CMS). This data provides utilization, charge amount, and actual payment amount for nearly every Medicare Part B provider for the years 2012 to 2016. Providers are identified by a National Provider Identifier (NPI), and for each NPI, the number of unique beneficiaries, total number of services, and average payment amount for every type of service is included in the data. Combined with the Medicare Physician and Other Supplier Aggregate

Figure 2.3: **Provider Response to Fee Bump with New and Established Patient Services**



Note: This figure shows that equilibrium service volume q_3 contains a combination of new and established Medicare patients before the fee bump. After the fee bump, service volume increases to q_4, and all established Medicare patients are replaced by new Medicaid patients.

Table (henceforth "Aggregate Table") also from CMS, the data also contains information on the provider's gender, address, and taxonomy, along with demographic information about the Medicare beneficiaries treated by the provider.

There are numerous advantages to these data. It contains extremely detailed information on specific Medicare providers, and this will allow an identification strategy based on variation in a provider's behavior over time. In particular, taxonomy information allows me to distinguish which providers qualify for the fee bump, exogenously dividing providers into two groups that receive very different payment rates for providing the same service to Medicaid beneficiaries.

I combine the provider-level data with state level data on the Medicaid fee bump. Specifically, I use data on each state's participation in the fee bump in each year and each state's Medicaid-to-Medicare fee ratio in each year.

2.4.1 Dataset Construction

The Service Level data contains observations by year, NPI, and HCPCS code. An important characteristic of this data, and any public Medicare data, is that any observation providing information on 10 or fewer beneficiaries is censored. For example, if a given NPI provided 10 beneficiaries HCPCS code 99212 and 11 beneficiaries HCPCS code 99213, then the former information would be omitted from the dataset and the latter would be included. This presents possible problems when manually aggregating variables from the HCPCS level to the NPI level. Fortunately, by merging with the Aggregate Table, NPI level statistics are observed. This means, for example, the total number of unique beneficiaries treated by a given NPI in a given year is available, as well as the total number of unique beneficiaries receiving a specific service type from a given NPI in a given year. However, the total number of unique beneficiaries receiving a *combination* of service types from a given NPI in a given year isn't necessarily available. Furthermore, if a provider saw 10 or fewer beneficiaries in a given year (for all services), then they are censored from both the Service Level data and Aggregate Table.

To get state-level Medicaid-to-Medicare fee ratios, I follow Maclean et al. (2018) and use values from Zuckerman & Goin (2012), Smith et al. (2015), and Zuckerman, Skopec, & Epstein (2017). These authors first compute the simple average Medicaid and Medicare fees for seven different primary care service types (defined by HCPCS code) for every state (except Tennessee). The Medicaid-to-Medicare ratio in a state is the weighted average of the ratios of average Medicaid fees to average Medicare fees, where weights are the share of US spending on each service.

To combat the censoring issue, to account for the specific details of the fee bump law, and to handle for idiosyncrasies in policy adoption across states, I do the following. First, I keep only observations where services were provided in an office setting (rather than a facility), and only if the NPI is associated with an individual provider (rather than an organization). I drop observations for supplier and non-medical taxonomies.[5] Providers who switch taxonomy at all from 2012 to 2016 are dropped from the data. Importantly, I drop any providers that don't appear in all five years of data, since appearing in one year and not another is indicative of having near the censoring threshold amount of patients. I drop any providers serving more than 3,000 unique beneficiaries in a year (corresponding to the 99th percentile). Finally, I exclude providers in Maryland, Colorado, Idaho, Indiana, Nevada, and Utah from the data, since these states provided increased fees to all physicians at least one year during 2012 to 2016. There is no fee-for-service component in Tennessee's Medicaid program, so it is also omitted from the data.

2.4.2 Summary Statistics

Table 2.2 presents provider statistics in the remaining sample. Of the 286,911 total providers,

[5]The following types are dropped: All Other Suppliers, Clinic or Group Practice, Clinical Laboratory, Independent Diagnostic Testing Facility, Independent Diagnostic Testing Facility (IDTF), Mass Immunization Roster Biller, Mass Immunizer Roster Biller, Multispecialty Clinic/Group Practice, Portable X-Ray Supplier, Portable X-ray, Slide Preparation Facility, Undefined Physician type, Unknown Physician Specialty Code, Unknown Supplier/Provider, and Unknown Supplier/Provider Specialty. These amount to less than one out of every 2000 observations.

Table 2.2: **Summary Statistics: Per-Year Provider Statistics**

	Non-PCPs	PCPs	All Providers
Number of Providers	208,413	78,498	286,911
Number of Unique Beneficiaries	553	413	512
	(526)	(290)	(473)
Number of Unique Ben. (99213 only)	122	162	134
	(182)	(151)	(175)
Number of Unique Ben. (99203 only)	30	9	24
	(63)	(32)	(56)
Total Medicare Payment Amount ($)	154,139	123,459	145,131
	(197,990)	(133,309)	(181,944)
Medicare Payment Amount (99213 only, $)	10,660	17,348	12,624
	(18,450)	(22,263)	(19,881)
Medicare Payment Amount (99203 only, $)	2,228	638	1,761
	(4,732)	(2,247)	(4,222)
Proportion of Medicare Ben. that are White	0.81	0.79	0.80
	(0.16)	(0.19)	(0.17)
Proportion of Medicare Ben. that are Male	0.41	0.41	0.41
	(0.12)	(0.10)	(0.12)
Average HCC Risk Score of Ben.	1.46	1.48	1.47
	(0.58)	(0.89)	(0.68)
N		1,098,302	

Note: This table shows mean provider service volume and patient demographics, with standard deviations in parentheses.

27% are PCPs that qualify to receive the increased Medicaid rates. Note that while non-PCPs don't earn increased rates, nearly two thirds provided at least one service that would receive the increased Medicaid rate had their taxonomy been different.

Table 2.3 shows the number of states by year that had legislated increased fees and had a Medicaid-to-Medicare fee ratio greater than or equal to one. It also shows the average fee bump across states.

Table 2.3: **Summary Statistics: Fee Bump by Year**

Year	# States with Bumped Fees	# States with Fee Ratio$\geq$ 1	Average Fee Ratio
2012	0	2	0.71
2013	50	50	1.01
2014	50	50	1.01
2015	19	14	0.78
2016	19	4	0.72

Note: This table shows the number of states with increased fees and the average Medicaid-to-Medicare fee ratio by year for all states (including Washington, D.C. and excluding Tennessee).

2.5 Empirical Strategy

I use several difference-in-differences specifications to identify and investigate spillover to Medicare in response to the Medicaid fee increase. These specifications differ in treatment measurement, the variation in treatment that identifies the causal impact of treatment on the outcome variable, and the outcome variable.

2.5.1 Treatment Measurement

Treatment in the context of this analysis is exposure to an exogenous increase in Medicaid payment relative to Medicare payment. I measure treatment with the variable z_{st}, defined in the following ways:

1. $\underline{z_{st} = Bump_{st}}$, where $Bump_{st}$ indicates state s had *legislated* Medicaid payment parity in year t. Using this definition, the average treatment effect identified is the average change in the outcome variable to *legislated* bumped fees.

2. $\underline{z_{st} = \mathbf{1}\left\{FeeRatio_{st} \geq 1\right\}}$. This variable indicates a state s that had a Medicaid-to-Medicare fee ratio equal to 1 or more in year t. Using this definition, the average treatment effect identified is the average change in the outcome variable when Medicaid fees are increased to a level *equivalent or greater* than Medicare fees. This variable is usually equal to $Bump_{st}$, though it differs for states such as Alaska, which had large Medicaid payment pre-fee bump, and for years 2015 and 2016, when several states had just partially increased fees (see Table 2.3).

3. $\underline{z_{st} = \mathbf{1}\left\{\Delta FeeRatio_{st} \geq 0.3\right\}}$, where $\Delta FeeRatio_{st} = FeeRatio_{st} - FeeRatio_{s,2012}$. This definition follows Maclean et al. (2018), which defines "large fee bump" states as those that increased their fee ratio by more than 0.3 since 2012.

4. $\underline{z_{st} = \Delta FeeRatio_{st}}$. This treatment variable, taking values other than zero and one, measures the increase in the magnitude of the Medicaid-to-Medicare fee ratio in a given state and year, and thus exploits more variation in Medicaid fees to identify the treatment effect in question.

5. $\underline{z_{st} = FeeRatio_{st}}$. Using the unadjusted Medicaid-to-Medicare fee ratio as treatment simply weights states and years according to the intensity of payment parity in the state. This term exploits variation of treatment differently than $\Delta FeeRatio_{st}$, and will

assist more in identifying spillover if the level of the fee ratio is more important to physicians than the change since 2012.

Along with these definitions of treatment that vary across state and time, since not all physicians qualify for increased Medicaid fees, physicians can be split into treated and untreated groups according to their specialty. Accordingly, let the variable $Qual_i$ indicate physician i is of a specialty that qualifies for increased Medicaid payment.

Drawbacks and Limitations

The variable $FeeRatio_{st}$ is the expenditure share-weighted average Medicaid-to-Medicare fee ratio of seven primary care services in state s and year t (see the discussion in Section 2.4). Because this treatment variable is aggregated to the state level, and is not a fee ratio specific to a physician, there may be a small difference in the actual ratio of Medicaid-to-Medicare fees that a physician faces and what's ultimately used in the empirical specifications. This difference exists because not every physician in a state provides the same mixture of primary care services.

While this is a limitation of this study, I use a couple strategies to rule out that the bias introduced by using a state-specific fee bump drives empirical results. First, treatment definitions under items 1 and 2 in Section 2.5.1 are not impacted at all by this bias, as they are, by definition, state specific.[6] Second, I use physician-specific fixed effects in one difference-in-differences specification, and this would capture any time-invariant differences between a physician's own fee ratio and the statewide fee ratio.

[6]We will see in the coming sections that the empirical results are robust to applying different treatment definitions.

2.5.2 Difference-in-Differences Specifications

I estimate three difference-in-difference models, each differing in the variation in treatment that identifies spillover between Medicare and Medicaid.

Variation Across States Over Time

The first specification takes the form of the canonical difference-in-differences model by comparing the change in the outcome variable in states with bumped Medicaid fees with the change in the outcome variable in states without bumped Medicaid fees. Formally, I esimate the model

$$y_{ist} = \gamma_i + \delta_s + \kappa_t + \boldsymbol{\beta}_1' X_{ist} + \beta_2 expanded_{st} + \alpha z_{st} + \epsilon_{ist} \mid Qual_i = 1 \tag{2.1}$$

where y_{ist} is the outcome of provider i in state s in year t. The variables γ_i, δ_s, and κ_t are provider, state, and time fixed effects; X_{ist} is a vector of controls varying across providers, states, and time; and $expanded_{st}$ indicates state s expanded Medicaid in year t. The variable z_{st} represents exogenous treatment, and varies only across states and time. When estimating parameters in this model, I limit the sample to physicians that qualify for increased Medicaid fees.

Variation Across States and Providers Over Time

To leverage more variation in data, I include all providers in the estimation sample and use the interaction of state and time varying treatment with the variable indicating a physician qualifies for treatment, $z_{st} \times Qual_i$, as the treatment variable. This specification takes the form

$$y_{ist} = \gamma_i + \delta_{st} + \boldsymbol{\beta}' X_{ist} + \alpha z_{st} \times Qual_i + \varepsilon_{ist} \tag{2.2}$$

where γ_i is a provider fixed effect, δ_{st} is a state-by-year fixed effect, and X_{ist} is a vector of time varying physician characteristics. The specification in Equation 2.2 has several advantages over the specification in Equation 2.1. First, because all providers are included in the sample, the difference in outcome change between qualifying and non-qualifying physicians contributes to the estimate of the average treatment effect. Furthermore, because treatment varies across states, years, and physicians, a more general state-by-year fixed effect can be added to the specification, absorbing more confounding variation than separate state and time fixed effects.[7]

Variation Across Providers Over Time

To further analyze the impact of the Medicaid fee bump on provider decisions, I conduct state-specific analysis and estimate the regression

$$y_{ist} = \gamma_i + \delta_t + \boldsymbol{\beta}_s' X_{ist} + \alpha_s \left(Bump_{st} \times Qual_i \right) + \epsilon_{ist} \tag{2.3}$$

for all states s in the estimation sample. This specification sacrifices cross-state variation to compute an average treatment effect that varies at the state level. The advantage of this specification is that I can compare estimates of α_s to the intensity of treatment in a given

[7]Including state-by-year fixed renders Medicaid expansion status of a state, $expanded_{st}$, redundant, and so this variable is omitted form the specification in Equation 2.2.

state s.

2.5.3 Outcome Variables

There are six total outcome variables y_{ist} used in this analysis.

1. Total unique Medicare beneficiaries receiving
   (a) any medical service from provider i in state s in year t
   (b) a medical service with HCPCS code 99213 from provider i in state s in year t
   (c) a medical service with HCPCS code 99203 from provider i in state s in year t
2. Total Medicare payment for
   (a) all medical service provided by provided i in state s in year t
   (b) a medical service with HCPCS code 99213 provided by provided i in state s in year t
   (c) a medical service with HCPCS code 99203 provided by provided i in state s in year t

I choose to focus my analysis on two measures of volume: total number of unique beneficiaries treated, and total payment received from CMS. The former measure has a direct relationship to Medicare beneficiaries' access to care, though may not capture all spillover if providers have a hard time adjusting the number of beneficiaries they treat. The latter, payment, addresses a more general notion of spillover, capturing revenue changes as a result of the policy.

For the measures above, I also look at narrower definitions based on HCPCS codes. The code 99213 is by far the most commonly used procedural code treating established patients, and the code 99203 is the most common for treating new patients. Both were in the group of services that received bumped fees, though breaking the outcomes down into these allows an analysis of where treatment changes, if any, were made.

To account for the right skewed distributions of outcome variables and the presence of several observations with value of zero, I apply an inverse hyperbolic sine transformation to all of the dependent variables listed above. While this transformation is unconventional, the resulting coefficients can be interpreted in the same way as if the dependent variables were logged. That is, the proportion change in the outcome variable when a provider qualifies for the fee bump is approximately $\exp(\alpha) - 1$. Furthermore, the approximation is more accurate than if the transformation $\ln(y + 1)$ were applied.[8]

2.5.4 Identification

The key identifying assumption of this study, and all studies that examine the consequences of the payment parity provision of the ACA, is that cross-state and intertemporal variations in the Medicaid-to-Medicare fee ratio are exogenous and unrelated to cross-state and intertemporal variations in the outcome variables. Depending on the definition of treatment z_{st}, different assumptions are required for the average treatment effect α to be identified.

When z_{st} has the definitions described in items 1, 2, and 3 in Section 2.5.1, I assume that a states adoption of bumped fees, whether or not the Medicaid-to-Medicare fee ratio is

[8] The proportion change approximation becomes more accurate with the dependent variable is large. For example, Bellemare & Wichman (2019) suggest using the approximation when the the dependent variable is greater than 10. The dependent variables used in this analysis all have means in the several hundreds (or greater).

larger than 1, and whether or not the change in the fee ratio since 2012 is larger than 0.3 is independent of all unobserved factors contributing to the outcome variables, conditional on fixed effects and controls.

On the other hand, when z_{st} is defined according to items 4 and 5 in Section 2.5.1, I require different identifying assumptions. The usual strict exogeneity assumption is extended: along with the assumption that there is no unobserved factor non-randomly driving adoption of increased fees, it requires that no unobserved factor non-randomly drives the variation in the Medicaid-to-Medicare fee ratio. Specifying treatment in this way, often referred to as "continuous treatment" or "dosage," is common when available (Card, 1992; Weber, 2014), and is typically the treatment of choice in studies concerning the impact of the Medicaid fee bump (Maclean et al., 2018; Alexander & Schnell, 2019).

Time varying control variables in X_{ist} include the average risk score of beneficiaries treated by a provider, the proportion white beneficialness treated by a provider, and the proportion of male beneficiaries treated by the provider. These variables control for patient health and demographics, which may be correlated with a state's adoption of increased Medicaid fees over time.

Ultimately, identification is impeded if physicians in states (and during years) with larger Medicaid-to-Medicare fee ratios also have larger service volume for reasons *other* than the fee bump.

2.6 Results

Tables 2.4 and 2.5 show the results from estimating Equation 2.1. Each cell in the tables

Table 2.4: **Qualifying Physicians' Response to Medicaid Rate Increase: Medicare Beneficiaries**

	Number of Unique Medicare Beneficiaries Treated		
	(1)	(2)	(3)
Treatment	Any Service	Established Patient	New Patient
$Bump_{st}$	0.190	0.324	2.837**
	(0.236)	(0.819)	(1.096)
$\mathbf{1}\{FeeRatio_{st} \geq 1\}$	0.00104	-4.354***	1.067
	(0.282)	(1.140)	(1.470)
$\mathbf{1}\{\Delta FeeRatio_{st} \geq 0.3\}$	-0.111	-2.209***	0.0680
	(0.153)	(0.545)	(0.822)
$\Delta FeeRatio_{st}$	-0.374	-4.456**	3.113
	(0.420)	(1.476)	(2.241)
$FeeRatio_{st}$	-0.374	-4.417**	3.134
	(0.420)	(1.477)	(2.242)
N	315909	315909	315909

Standard errors (clustered by provider) in parentheses. Provider, state, and year fixed effects and controls are used in every regression. All point estimates should be interpreted as percent changes in the dependent variable when a physician qualifies for the fee bump, holding all else constant. "Established" refers to services with HCPCS code 99213; "New" refers to services with HCPCS code 99203. $\Delta FeeRatio_{st} = FeeRatio_{st} - FeeRatio_{s,2012}$.

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table 2.5: **Qualifying Physicians' Response to Medicaid Rate Increase: Medicare Payment**

	Total Medicare Payment for Services		
	(1)	(2)	(3)
Treatment	Any Service	Established Patient	New Patient
$Bump_{st}$	-0.757*	0.432	5.963**
	(0.317)	(1.466)	(2.334)
$\mathbf{1}\{FeeRatio_{st} \geq 1\}$	0.0123	-7.808***	2.560
	(0.379)	(1.990)	(3.153)
$\mathbf{1}\{\Delta FeeRatio_{st} \geq 0.3\}$	0.593**	-2.574**	0.368
	(0.206)	(0.984)	(1.732)
$\Delta FeeRatio_{st}$	1.514**	-5.849*	6.710
	(0.565)	(2.651)	(4.890)
$FeeRatio_{st}$	1.513**	-5.805*	6.761
	(0.565)	(2.652)	(4.892)
N	315909	315909	315909

Standard errors (clustered by provider) in parentheses. Provider, state, and year fixed effects and controls are used in every regression. All point estimates should be interpreted as percent changes in the dependent variable when a physician qualifies for the fee bump, holding all else constant. "Established" refers to services with HCPCS code 99213; "New" refers to services with HCPCS code 99203. $\Delta FeeRatio_{st} = FeeRatio_{st} - FeeRatio_{s,2012}$.

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

contains the estimated average treatment effect, $\hat{\alpha}$. There are 30 estimates total in the two tables: the combination of six independent variables (unique number of beneficiaries given any service, an established patient service, and a new patient service; total Medicare payment for any service, an established patient service, and a new patient service) and 5 independent variables (see Section 2.5.1).

Table 2.4 show small and imprecise changes in the unique number of Medicare beneficiaries treated by providers in states that have legislated an increase in Medicaid fees. For all services, qualifying physicians in states with increased fees treat between -0.29 percent fewer and 0.67 percent more Medicare beneficiaries (95% confidence interval). Using the variation in the Medicaid fee ratio, the average treatment effect is between a 1.21 percent decrease and a 0.47 percent increase (95% confidence interval).

The impact of the fee bump on Medicare beneficiaries is much larger and more precise when looking at services designated for established patients only. When $FeeRatio_{st}$ is treatment, providers offer the most common established patient service to 4% fewer unique beneficiaries, though offer the most common new patient service to 3% *more* new patients.

According to Table 2.5, Physicians qualifying for increased Medicaid fees earn 0.76% less total Medicaid payment in states that have increased fees. When using the Medicaid-to-Medicare fee ratio as treatment, the sign flips, and qualifying providers actually earn 1.5% *more* total Medicare pay in states that adopt the fee bump. Examining columns 2 and 3 of Table 2.5 helps to explain why: payment from the established payment service decreased, though payment from the new patient service increased. Per service payment for the established patient service is roughly 30% less than per service payment for the new patient service, suggesting that while total service volume may have decreased or been unchanged, total pay may still increase.

Tables 2.6 and 2.7 show results from the difference-in-differences specification in Equation 2.2. By including additional variation from non-qualifying providers and more general state-

Table 2.6: **Provider Response to Medicaid Rate Increase: Medicare Beneficiaries**

	Number of Unique Medicare Beneficiaries Treated		
	(1)	(2)	(3)
Treatment	Any Service	Established Patient	New Patient
$Bump_{st} \times Qual_i$	-0.315**	-7.248***	1.077*
	(0.0981)	(0.350)	(0.477)
$\mathbf{1}\{FeeRatio_{st} \geq 1\} \times Qual_i$	0.00482	-4.740***	3.585***
	(0.0874)	(0.325)	(0.454)
$\mathbf{1}\{\Delta FeeRatio_{st} \geq 0.3\} \times Qual_i$	-0.130	-5.105***	3.776***
	(0.100)	(0.375)	(0.518)
$\Delta FeeRatio_{st} \times Qual_i$	-0.240	-9.927***	8.570***
	(0.199)	(0.704)	(1.064)
$FeeRatio_{st} \times Qual_i$	-0.244	-9.933***	8.554***
	(0.199)	(0.703)	(1.064)
N	1061847	1061847	1061847

Standard errors (clustered by provider) in parentheses. Provider and state-year fixed effects and controls are used in every regression. All point estimates should be interpreted as percent changes in the dependent variable when a physician qualifies for the fee bump, holding all else constant. "Established" refers to services with HCPCS code 99213; "New" refers to services with HCPCS code 99203. $\Delta FeeRatio_{st} = FeeRatio_{st} - FeeRatio_{s,2012}$.

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

by-year fixed effects, the impact of the Medicaid fee bump on Medicare providers becomes a bit clearer. Total number of unique beneficiaries decreases slightly when treated, though unique beneficiaries receiving the established patient service decreases by nearly 10% and unique beneficiaries receiving the new patient service increases by 8.5%. Payment follows a similar pattern: once using $FeeRatio_{st} \times Qual_i$ as treatment, we total Medicare payment to treated physicians decreases by 1.5%, and payment for the established and the new patient

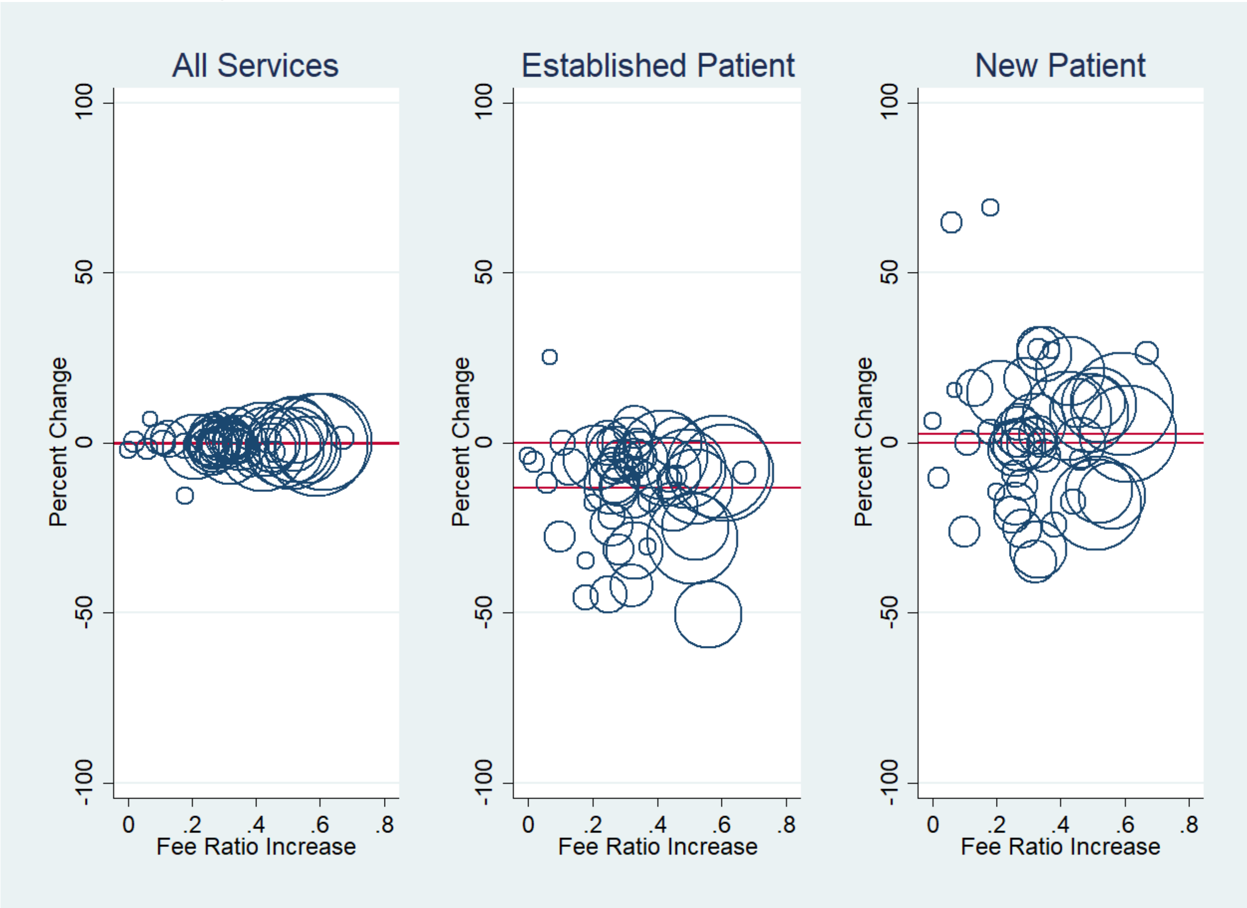
Table 2.7: **Provider Response to Medicaid Rate Increase: Medicare Payment**

	Total Medicare Payment for Services		
	(1)	(2)	(3)
Treatment	Any Service	Established Patient	New Patient
$Bump_{st} \times Qual_i$	-0.396**	-13.39***	2.626**
	(0.128)	(0.606)	(1.013)
$\mathbf{1}\{FeeRatio_{st} \geq 1\} \times Qual_i$	-0.684***	-9.010***	7.844***
	(0.115)	(0.578)	(0.991)
$\mathbf{1}\{\Delta FeeRatio_{st} \geq 0.3\} \times Qual_i$	-0.768***	-9.389***	8.250***
	(0.131)	(0.672)	(1.136)
$\Delta FeeRatio_{st} \times Qual_i$	-1.529***	-18.19***	19.27***
	(0.257)	(1.201)	(2.460)
$FeeRatio_{st} \times Qual_i$	-1.536***	-18.20***	19.24***
	(0.257)	(1.201)	(2.459)
N	1061847	1061847	1061847

Standard errors (clustered by provider) in parentheses. Provider and state-year fixed effects and controls are used in every regression. All point estimates should be interpreted as percent changes in the dependent variable when a physician qualifies for the fee bump, holding all else constant. "Established" refers to services with HCPCS code 99213; "New" refers to services with HCPCS code 99203. $\Delta FeeRatio_{st} = FeeRatio_{st} - FeeRatio_{s,2012}$.

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Figure 2.4: **Impact of Bumped Fees on Payment by State**



Note: This figure shows the average treatment effect (ATE) of the Medicaid fee bump on qualifying physician's payment from Medicare. Each circle corresponds to a U.S. state's treatment effect, where its size is determined by the number of observations used to compute the effect. The left panel is the ATE on payment for all services, the middle panel is the ATE on payment for a service for established Medicare beneficiaries, and the right panel is the ATE on payment for a service for established new beneficiaries. The results indicate, particularly in the middle panel, the ATE of the fee bump is stronger in states where the fee bump is larger.

services decrease and increase by similar amounts, respectively.

Finally, Figure 2.4 plots the coefficients α_s (from Equation 2.3) against the increase in Medicaid-to-Medicare fee-ratio in state s when fees are bumped: $\max_t \{FeeRatio_{st}|bump_{st} = 1\} - \min_t \{FeeRatio_{st}|bump_{st} = 0\}$. The size of each circle is weighted by to the number of observations used to compute its corresponding point estimate. The leftmost panel of Figure 2.4 shows the very slight average decrease of total payment to physicians in states that qualify for the fee bump. Across all states, total payment for all services to physicians qualifying for the fee bump decreased on average by 0.396%, with a 95% confidence interval $(-0.652, -0.140)$. Roughly 57% of states had a decline, and there is a very small positive relationship between the change in payment amount and the change in fee ratio.

The middle panel, however, shows a much stronger impact of the fee bump on states. Only two of the 44 unique values of α_s are positive, implying more than 95% of states in the sample saw a decrease in payment for the established patient service to physicians qualifying for the fee bump. There is a strong negative relationship between the magnitude of the change in payment and the magnitude of the fee ratio increase. The rightmost panel effectively shows the opposite: on average, when physicians qualify for increased fees, payment for the new patient service increases. In the estimation sample, 57% of states showed this pattern.

2.7 Discussion

The results presented in the previous section can be summarized as follows:

1. Physicians exposed to the Medicaid fee bump decreased the unique number of Medicare beneficiaries they treat.

2. Physicians exposed to the Medicaid fee bump decreased the total payment they receive from Medicare.

3. Physicians exposed to the Medicaid fee bump decreased the number of Medicare beneficiaries given and total Medicare payment from the procedural code designated for established patients only. Physicians not exposed behave in the opposite way, increas-

ing volume of this service.

4. Physicians exposed to the Medicaid fee bump increased the number of Medicare beneficiaries given and total Medicare payment from the procedural code designated for new patients only. Physicians not exposed behave in the opposite way, decreasing volume of this service.

Why do we see this pattern? While the mixed economy models of Sloan et al. (1978), Garthwaite (2012), and Glied & Hong (2018) predict decreases in Medicare service volume in response to access and payment increasing provisions for Medicaid, no study explains why volume of one service type could decrease while another increases.

One obvious explanation is that physicians increased the number of Dual Eligible beneficiaries, and so that would account for the large increase in new patient Medicare services. To examine this hypothesis, I estimate the model in Equation 2.2, with dependent variable y_{ist} equal to the percent of unique beneficiaries that are dual eligible in Medicare and Medicaid treated by provider i in state s in year t. Results are in Table 2.8. The point estimates indicate that physicians qualifying for increased Medicaid fees increase the number of Dual Eligible beneficiaries that they treat by about one beneficiary on average. While positive, the magnitudes of these treatment effects are not sufficiently large to account for the size of the increase in beneficiaries that receive the new patient Medicare service.

2.7.1 Conceptual Framework Revisited

The model discussed in Section 2.3 provides a reason for why some Medicare patients are replaced by Medicaid patients as a result of the fee bump, and moreover suggest that es-

Table 2.8: **Provider Response to Medicaid Rate Increase: Dual Eligible Beneficiaries**

Treatment	Percent of Beneficiaries that are Dual Eligible (1)
$Bump_{st} \times Qual_i$	0.147^{***}
	(0.0175)
$\mathbf{1}\{FeeRatio_{st} \geq 1\} \times Qual_i$	0.185^{***}
	(0.0157)
$\mathbf{1}\{\Delta FeeRatio_{st} \geq 0.3\} \times Qual_i$	0.197^{***}
	(0.0186)
$\Delta FeeRatio_{st} \times Qual_i$	0.421^{***}
	(0.0371)
$FeeRatio_{st} \times Qual_i$	0.421^{***}
	(0.0372)
N	954486

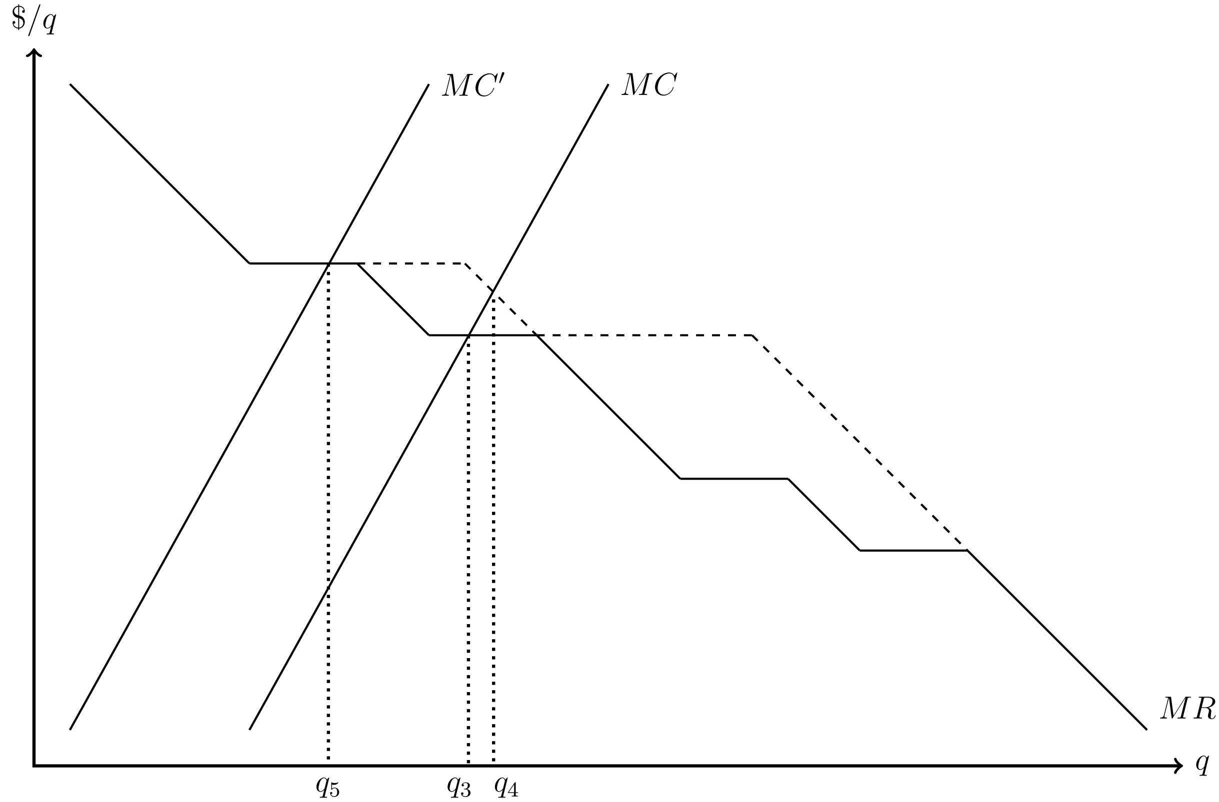
Standard errors (clustered by provider) in parentheses. Provider and state-year fixed effects and controls are used in every regression. All point estimates should be interpreted as percentage point changes. $\Delta FeeRatio_{st} = FeeRatio_{st} - FeeRatio_{s,2012}$.

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

tablished patients are more vulnerable than new patients. However, the empirical results of this paper suggest that provision of care to new Medicare beneficiaries increased, and this simple framework cannot justify that result.

To be specific, consider the model reproduced in Figure 2.5. If a physician has marginal

Figure 2.5: **No Provider Response to Fee Bump**

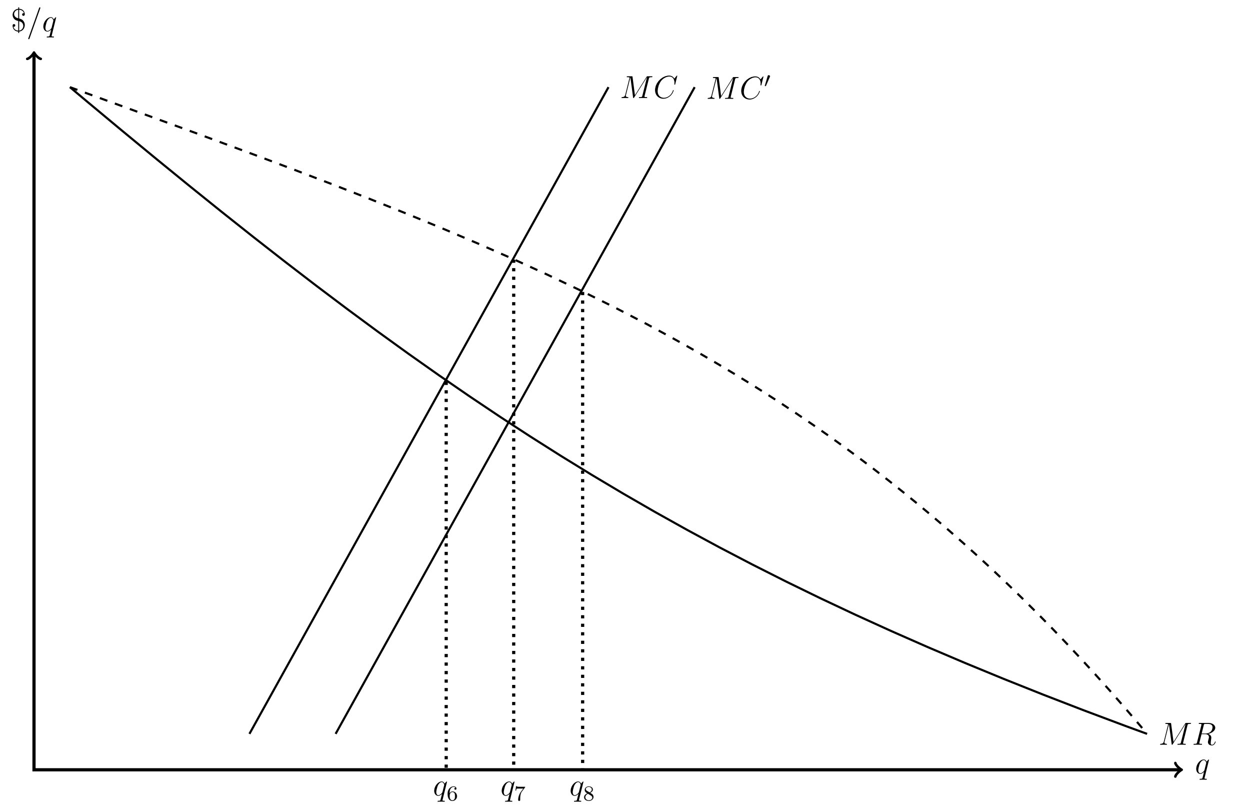


Note: This figure shows that when marginal cost is sufficiently high (MC'), the equilibrium service volume remains unchanged after the fee bump at q_5.

cost curve MC, there's no excess demand for new Medicare patient services, so new Medicare patient services would not increase, and the increase in services from q_3 to q_4 is only to new Medicaid beneficiaries. If a physician has marginal cost MC', there is excess demand, but the fee bump would leave total provision of care unchanged at q_5, and new Medicare beneficiary services would either be unchanged or decrease.

Reconciling this paper's empirical results requires a model with patients that are heterogeneous in dimensions other than Medicare, Medicaid, established, and new. Of course, physicians consistently treat established Medicare patients, and some new Medicare patients remain untreated, despite the differing reimbursement amounts. Including more patient heterogeneity essentially *smooths* the marginal revenue curves in Figures 2.1, 2.2, and 2.3 because the marginal unit of healthcare service goes to the patient with the highest remaining marginal revenue. The fee bump manifests as an outwards bowing of the marginal revenue curve. This is pictured in Figure 2.6. Beneficiaries with all types of payers span the domain

Figure 2.6: **Provider Response to Fee Bump with Heterogeneous Patients**



Note: This figure shows a physician's marginal revenue (MR) of service volume q when patients are heterogeneous. The higher marginal revenue patients correspond to lower values of q, and marginal revenue decreases smoothly. The Medicaid fee bump increases marginal revenue for some patients, effectively bowing out marginal revenue to the dashed line. Payment parity increases equilibrium service volume from q_6 to q_7, where established Medicare patients are replaced by Medicaid patients. Simultaneously, marginal costs decrease (moving the marginal cost curve from MC to MC') and equilibrium service volume increased to q_8, and more new Medicare patients are treated.

of these marginal revenue curves. The fee bump increases marginal revenue for serving the

Medicaid portion of the population, which increases the number of services provided from q_6 to q_7. In this example, crowd out could occur to *any* non-Medicaid beneficiary, new or established, if a Medicaid beneficiary has higher marginal revenue after the fee bump.

Heterogeneous patients alone does not explain why more new Medicare beneficiaries were treated, however. This requires a subsequent decrease in marginal cost *caused* by the fee bump, pictured in Figure 2.6 as MC'. In this case, total quantity of services provided by a physician increases to q_8, and this increase includes both Medicaid patients and new Medicare patients.

I formalize this in the following toy model. A physician chooses services provided to new and established Medicare beneficiaries, x_n and x_e, and services provided to new and established Medicaid beneficiaries, y_n and y_e. Medicare services have prices p_n and p_e, and Medicaid services have prices $r \cdot p_n$ and $r \cdot p_e$, where r is the Medicaid-to-Medicare fee ratio. Physicians maximize utility, which is a linear combination of profit and patient utility. That is, physicians maximize

$$U(x_n, x_e, y_n, y_e) = p_n x_n + p_e x_e + r p_n y_n + r p_e y_e - C(x_n, x_e, y_n, y_e) + V(x_n, x_e, y_n, y_e) \quad (2.4)$$

where C is a cost function, and V is a function representing the utility a physician derives from the total utility of their beneficiaries. To guarantee an interior solution, I assume that the function $C - V$ is strictly convex over the entire domain of x_n, x_e, y_n, and y_e, and that p_n, p_e, and r are all strictly greater than zero.

Proposition 2.7.1 (Comparative Statics)**.** *There exists an upper bound* $\underline{W} < 0$ *such that if* $\frac{\partial^2 C}{\partial x_n \partial y_n} - \frac{\partial^2 V}{\partial x_n \partial y_n} < \underline{W}$*, then the optimal service volumes are such that* $\frac{dx_n^*}{dr} > 0, \frac{dx_e^*}{dr} < 0, \frac{dy_n^*}{dr} > 0,$

and $\frac{dy_n^*}{dr} > 0$.

Proof. See Appendix B.1. □

Proposition 2.7.1 states that if the marginal cost (inclusive of altruistic preferences) of care to new Medicare beneficiaries is decreasing in the amount of care to new Medicaid beneficiaries, then the fee bump increases the optimal amount of care to all Medicaid beneficiaries. A likely interpretation is that a blanket "accepting *all* new patients" policy was implemented by physicians that qualified for increased fees. Moreover, increased profit from higher Medicaid reimbursements may have been reinvested by physicians to facilitate increasing service volume, hence expanding access to more than just Medicaid patients.

Outside of this model, it's also possible that practice-level changes caused the observed differential change in service volume between new and established Medicare beneficiaries. Because specialists do not qualify for increased fees for providing service codes indicated for new patients, the optimal response is to forward new Medicare patients to qualifying physicians, and less-profitable established patients to non-qualifying physicians, thus explaining the observed pattern in the data.

2.8 Conclusion

This paper shows physicians change their delivery of health services to one population in response to changes in payment for services delivered to another population. I find that physicians qualifying for increased fees under the Medicaid fee bump increased service volume to new Medicare patients, but decreased service volume to established Medicare patients.

This pattern remains after looking at several different measures of service volume, as well as several different treatment definitions and using different identifying variation.

I extend the mixed-economy model of Sloan et al. (1978) to interpret the empirical results. Under the condition that patients offer heterogeneous marginal revenue to physicians, an increase in service volume to new Medicare beneficiaries and a decrease in service volume to established Medicare beneficiaries occurs if the fee bump also caused a simultaneous decrease in marginal cost. It's not clear, however, what drives this decrease in marginal cost.

References

Abadie, A., & Kasy, M. 2017, April. *The risk of machine learning* (Working Paper). Retrieved from `https://arxiv.org/pdf/1703.10935.pdf`

Alexander, D., & Schnell, M. 2019, July. *The impacts of physician payments on patient access, use, and health* (Working Paper No. 26095). National Bureau of Economic Research. Retrieved from `http://www.nber.org/papers/w26095` doi: 10.3386/w26095

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. Retrieved from `https://doi.org/10.1287/opre.2018.1821` doi: 10.1287/opre.2018.1821

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. Retrieved from `http://www.sciencedirect.com/science/article/pii/S0147596705000247` (Symposium: Monetary and Exchange Rate Policies in the CIS Countries: Between the EU and Russia) doi: https://doi.org/10.1016/j.jce.2005.03.013

Bellemare, M. F., & Wichman, C. J. 2019. Elasticities and the inverse hyperbolic sine transformation. *Oxford Bulletin of Economics and Statistics*, *0*(0). Retrieved from https://onlinelibrary.wiley.com/doi/abs/10.1111/obes.12325 doi: 10.1111/obes.12325

Berry, S. 1994. Estimating discrete-choice models of product differentiation. *The RAND Journal of Economics*, *25*(2), 242–262. Retrieved from http://www.jstor.org/stable/2555829

Berry, S., Levinsohn, J., & Pakes, A. 1995. Automobile prices in market equilibrium. *Econometrica*, *63*(4), 841–890. Retrieved from http://www.jstor.org/stable/2171802

Bulow, J. I., Geanakoplos, J. D., & Klemperer, P. 1985. Multimarket oligopoly: Strategic substitutes and complements. *Journal of Political Economy*, *93*(3), 488-511.

Burlig, F., Knittel, C., Rapson, D., Reguant, M., & Wolfram, C. 2019, January. *Machine learning from schools about energy efficiency* (Working Paper).

Card, D. 1992. Using regional variation in wages to measure the effects of the federal minimum wage. *Industrial and Labor Relations Review*, *46*(1), 22–37. Retrieved from http://www.jstor.org/stable/2524736

Cardell, N. S. 1997. Variance components structures for the extreme-value and logistic distributions with application to models of heterogeneity. *Econometric Theory*, *13*(2), 185–213. Retrieved from http://www.jstor.org/stable/3532724

Carey, C. M., Miller, S., & Wherry, L. R. 2018, October. *The impact of insurance expansions on the already insured: The affordable care act and medicare* (Working Paper No. 25153). National Bureau of Economic Research. Retrieved from http://www.nber.org/papers/w25153 doi: 10.3386/w25153

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. Retrieved from `http://www.sciencedirect.com/science/article/pii/B9780444535924000062` doi: https://doi.org/10.1016/B978-0-444-53592-4.00006-2

Choné, P., & Ma, A. 2011. Optimal health care contract under physician agency. *Annals of Economics and Statistics*(101/102), 229–256. Retrieved from `http://www.jstor.org/stable/41615481`

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. Retrieved from `http://www.sciencedirect.com/science/article/pii/S016762960800091X` doi: https://doi.org/10.1016/j.jhealeco.2008.07.010

Einav, L., Finkelstein, A., & Mahoney, N. 2018. Provider incentives and healthcare costs: Evidence from long-term care hospitals. *Econometrica*, *86*(6), 2161-2219. Retrieved from `https://onlinelibrary.wiley.com/doi/abs/10.3982/ECTA15022` doi: 10.3982/ECTA15022

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. Retrieved from http://www.aeaweb.org/articles?id=10.1257/aer.20170092 doi: 10.1257/aer.20170092

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. Retrieved from http://www.sciencedirect.com/science/article/pii/S0167268106000072 doi: https://doi.org/10.1016/j.jebo.2006.01.001

Feng Lu, S. 2012. Multitasking, information disclosure, and product quality: Evidence from nursing homes. *Journal of Economics & Management Strategy*, *21*(3), 673-705. Retrieved from https://onlinelibrary.wiley.com/doi/abs/10.1111/j.1530-9134.2012.00341.x doi: 10.1111/j.1530-9134.2012.00341.x

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. doi: 10.1162/ajhe_a_00083

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. Retrieved from https://doi.org/10.1007/s11151-015-9467-y doi: 10.1007/s11151-015-9467-y

Garthwaite, C. L. 2012, April. The doctor might see you now: The supply side effects of public health insurance expansions. *American Economic Journal: Economic Policy*, *4*(3), 190-215. Retrieved from http://www.aeaweb.org/articles?id=10.1257/pol.4.3.190 doi: 10.1257/pol.4.3.190

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. Retrieved from http://www.nber.org/papers/w12301 doi: 10.3386/w12301

Gaynor, M., Ho, K., & Town, R. J. 2015, June. The Industrial Organization of Health-Care Markets. *Journal of Economic Literature*, *53*(2), 235-284. Retrieved from https://ideas.repec.org/a/aea/jeclit/v53y2015i2p235-84.html

Gaynor, M., Rebitzer, J., & Taylor, L. 2004. Physician incentives in health maintenance organizations. *Journal of Political Economy*, *112*(4), 915–931. Retrieved from http://www.jstor.org/stable/10.1086/421172

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. Retrieved from http://www.sciencedirect.com/science/article/pii/S0167629617310123 doi: https://doi.org/10.1016/j.jhealeco.2018.05.001

Grassi, S., & Ma, A. 2016. Information acquisition, referral, and organization. *The RAND Journal of Economics*, *47*(4), 935-960. Retrieved from `https://onlinelibrary.wiley.com/doi/abs/10.1111/1756-2171.12160` doi: 10.1111/1756-2171.12160

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. Retrieved from `http://www.aeaweb.org/articles?id=10.1257/aer.20151057` doi: 10.1257/aer.20151057

Hansen, L. P. 1982. Large sample properties of generalized method of moments estimators. *Econometrica*, *50*(4), 1029–1054. Retrieved from `http://www.jstor.org/stable/1912775`

Ho, K., & Lee, R. S. 2017, August. *Equilibrium provider networks: Bargaining and exclusion in health care markets* (Working Paper No. 23742). National Bureau of Economic Research. Retrieved from `http://www.nber.org/papers/w23742` doi: 10.3386/w23742

Ho, K., & Pakes, A. 2014. Hospital choices, hospital prices, and financial incentives to physicians. *The American Economic Review*, *104*(12), 3841–3884. Retrieved from `http://www.jstor.org/stable/43495359`

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. Retrieved from https://onlinelibrary.wiley.com/doi/abs/10.1111/iere.12278 doi: 10.1111/iere.12278

Jacob, B. A. 2005. Accountability, incentives and behavior: the impact of high-stakes testing in the chicago public schools. *Journal of Public Economics*, *89*(5), 761 - 796. Retrieved from http://www.sciencedirect.com/science/article/pii/S0047272704001549 doi: https://doi.org/10.1016/j.jpubeco.2004.08.004

Kim, M., Sudhir, K., & Uetake, K. 2019, September. *A structural model of a multi-tasking salesforce: Job task allocation and incentive plan design* (Discussion Paper No. 2199). Cowles Foundation. Retrieved from https://ssrn.com/abstract=3452977 doi: 10.2139/ssrn.3452977

Maclean, J. C., McClellan, C., Pesko, M. F., & Polsky, D. 2018, July. *Reimbursement rates for primary care services: Evidence of spillover effects to behavioral health* (Working Paper No. 24805). National Bureau of Economic Research. Retrieved from http://www.nber.org/papers/w24805 doi: 10.3386/w24805

Mansour, M., Favini, N., Carome, M. A., Wolfe, S., Woolhandler, S., & Himmelstein, D. U. 2017. Accountable care organizations' payments to physicians: Patients should have the right to know. *American Journal of Public Health*, *107*(8), 1251-1252. Retrieved from https://doi.org/10.2105/AJPH.2017.303918 (PMID: 28700294) doi: 10.2105/AJPH.2017.303918

McInerney, M., Mellor, J. M., & Sabik, L. M. 2017, August. The effects of state medicaid expansions for working-age adults on senior medicare beneficiaries. *American Economic Journal: Economic Policy*, *9*(3), 408-38. Retrieved from http://www.aeaweb.org/articles?id=10.1257/pol.20150402 doi: 10.1257/pol.20150402

McWilliams, J. M. 2014, Oct 01. Aco payment models and the path to accountability. *Journal of General Internal Medicine*, *29*(10), 1328–1330. Retrieved from https://doi.org/10.1007/s11606-014-2868-z doi: 10.1007/s11606-014-2868-z

McWilliams, J. M., Hatfield, L. A., Chernew, M. E., Landon, B. E., & Schwartz, A. L. 2016. Early performance of accountable care organizations in medicare. *New England Journal of Medicine*, *374*(24), 2357-2366. Retrieved from https://doi.org/10.1056/NEJMsa1600142 (PMID: 27075832) doi: 10.1056/NEJMsa1600142

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. Retrieved from https://doi.org/10.1056/NEJMsa1803388 (PMID: 30183495) doi: 10.1056/NEJMsa1803388

Milgrom, P., & Roberts, J. 1990. Rationalizability, learning, and equilibrium in games with strategic complementarities. *Econometrica*, *58*(6), 1255–1277. Retrieved from http://www.jstor.org/stable/2938316

Miller, S., Altekruse, S., Johnson, N., & Wherry, L. R. 2019, July. *Medicaid and mortality: New evidence from linked survey and administrative data* (Working Paper No. 26081). National Bureau of Economic Research. Retrieved from http://www.nber.org/papers/w26081 doi: 10.3386/w26081

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. Retrieved from https://onlinelibrary.wiley.com/doi/abs/10.1111/j.1756-2171.2009.00090.x doi: 10.1111/j.1756-2171.2009.00090.x

Nevo, A. 2000. A practitioner's guide to estimation of random coefficients logit models of demand. *Journal of Economics & Management Strategy*, *9*(4), 513-548. Retrieved from https://onlinelibrary.wiley.com/doi/abs/10.1111/j.1430-9134.2000.00513.x doi: 10.1111/j.1430-9134.2000.00513.x

Rebitzer, J. B., & Votruba, M. E. 2011, October. *Organizational economics and physician practices* (Working Paper No. 17535). National Bureau of Economic Research. Retrieved from http://www.nber.org/papers/w17535 doi: 10.3386/w17535

Ryan, A. M., Shortell, S. M., Ramsay, P. P., & Casalino, L. P. 2015. Salary and quality compensation for physician practices participating in accountable care organizations. *The Annals of Family Medicine*, *13*(4), 321-324. Retrieved from http://www.annfammed.org/content/13/4/321.abstract doi: 10.1370/afm.1805

Slade, M. E. 1996. Multitask agency and contract choice: An empirical exploration. *International Economic Review*, *37*(2), 465–486. Retrieved from http://www.jstor.org/stable/2527333

Sloan, F., Mitchell, J., & Cromwell, J. 1978. Physician participation in state medicaid programs. *The Journal of Human Resources*, *13*, 211–245. Retrieved from `http://www.jstor.org/stable/145253`

Smith, V., Gifford, K., Ellis, E., Rudowitz, R., Snyder, L., & Hinton, E. 2015. *Medicaid reforms to expand coverage, control costs and improve care: Results from a 50-state medicaid budget survey for state fiscal years 2015 and 2016* (Tech. Rep.). Kaiser Family Foundation.

Weber, C. E. 2014. Toward obtaining a consistent estimate of the elasticity of taxable income using difference-in-differences. *Journal of Public Economics*, *117*, 90 - 103. Retrieved from `http://www.sciencedirect.com/science/article/pii/S004727271400111X` doi: https://doi.org/10.1016/j.jpubeco.2014.05.004

Yasaitis, L. C., Pajerowski, W., Polsky, D., & Werner, R. M. 2016. Physicians' participation in acos is lower in places with vulnerable populations than in more affluent communities. *Health Affairs*, *35*(8), 1382-1390. Retrieved from `https://www.healthaffairs.org/doi/abs/10.1377/hlthaff.2015.1635` doi: 10.1377/hlthaff.2015.1635

Zuckerman, S., & Goin, D. 2012. *How much will medicaid physician fees for primary care rise in 2013? evidence from a 2012 survey of medicaid physician fees* (Tech. Rep.). The Kaiser Commission on Medicaid and the Uninsured.

Zuckerman, S., Skopec, L., & Epstein, M. 2017. *Medicaid physician fees after the aca primary care fee bump* (Tech. Rep.). Urban Institute.

Appendices

Appendix A

Appendix to Chapter 1

A.1 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.[1] 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).[2] 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

[1]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).

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

employed by) an ACO participant, that beneficiary is assigned to that participant's ACO.[3]

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 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").[4]

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.[5] 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.1.1 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

[3]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.

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

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

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

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

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

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

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.1.2 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., 2016, 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.

A.2 Strategic Complementarity and Existence of Equilibrium

Usual definitions of supermodular games (e.g., Bulow et al. (1985); Milgrom & Roberts (1990)) require that 1) the strategy space of every agent is compact, 2) the payoff function of every agent is upper semicontinuous in their own actions, 3) the payoff function of every agent is continuous in other agent's actions, and 4) the payoff function of every agent has increasing differences. Conditions 1) and 2) of this definition are easy to confirm for the game played by ACO participants. The strategy space of each participant is $[-1, 1] \times [0, 1]$, which is compact. Upper semicontinuity in own savings effort s_{ij} is established because ACOs qualify for shared savings when the savings rate S_j is greater than *or equal to* the minimum savings rate $\underline{S}_j$ (and similarly for quality score). Condition 3) fails since participant payoff in other's efforts is only upper semicontinuous, but not fully continuous, due to the minimum savings rate. Finally, condition 4) does not typically hold: increasing differences requires the assumption that $\frac{w_{ij}^3 B_j}{2} \geq \frac{\partial^2 c_{ij}}{\partial s_{ij} \partial q_{ij}}$. I estimate the parameter $\kappa \equiv \frac{\partial^2 c_{ij}}{\partial s_{ij} \partial q_{ij}}$, which measures the savings-quality tradeoff, in Section 1.4. The value is very large (Table 1.8). Since several ACO participants have very small influences weights w_{ij}, it's impossible for this condition to be satisfied for all ACOs.[6]

In the following two propositions, I show that when these conditions hold, the game played by ACO participants exhibits strategic complementarity.

Proposition A.2.1. *Consider the simultaneous move game played by participants in ACO j, and let $i, i' \in j$ with $i \neq i'$.*

1. $\frac{\partial R_{ij}}{\partial s_{ij}}$ *is weakly increasing in $q_{i'j}$ and constant in $s_{i'j}$.*

2. $\frac{\partial R_{ij}}{\partial q_{ij}}$ *is weakly increasing in $s_{i'j}$ and constant in $q_{i'j}$.*

Proof. Note that if $S_j < \underline{S}_j$ or $Q_j < \underline{Q}$, $\frac{\partial R}{\partial s_{ij}}$ is identically zero, so the proof is trivial. Otherwise, we have

$$\frac{\partial^2 R_{ij}}{\partial s_{ij} \partial s_{i'j}} = 0 \tag{A.3}$$

$$\frac{\partial^2 R_{ij}}{\partial s_{ij} \partial q_{i'j}} = 0.5 \cdot B_j w_{ij}^2 w_{i'j} \geq 0 \tag{A.4}$$

which proves item 1 of the proposition. Item 2 has a nearly identical proof. □

Proposition A.2.2. *Consider the simultaneous game played by participants in ACO j, and let $i, i' \in j$ with $i \neq i'$. Let BR_s^i and BR_q^i be the best response functions of the savings and quality efforts, respectively, of participant i. Then, for all $i \in I_j$,*

1. *BR_s^i and BR_q^i are weakly increasing in $q_{i'j}$ and $s_{i'j}$, respectively, for all $i' \neq i$.*

2. *If $\frac{\partial^2 c_{ij}}{\partial s_{ij} \partial q_{ij}} \leq \frac{w_{ij}^3}{2} B_j$, then BR_s^i and BR_q^i are also increasing in $s_{i'j}$ and $q_{i'j}$, respectively, for all $i' \neq i$.*

[6] With data on individual providers, a parameter κ_j specific to ACOs could be estimated. With this parameter, I could confirm exactly how many ACOs are playing in a supermodular game.

Proof. Item 1 of Proposition A.2.2 follows trivially from Items 1 and 2 of Proposition A.2.2.

To prove Item 2, let $\frac{\partial^2 c}{\partial s_{ij} q_{ij}} \leq \frac{w_{ij}^3}{2} B_j$. Suppose $s_{i'j}$ increases to $s'_{i'j}$. From Item 1, q_{ij} increases to $q'_{ij} = BR_q(s'_{-ij}, q_{-ij})$ as well. The first order condition for s_{ij} maintains

$$\frac{\partial R}{\partial s_{ij}}(q'_{ij}) = \frac{\partial c}{\partial s_{ij}} \left(s'_{ij}, q'_{ij}\right) \tag{A.5}$$

The left hand side of the above is marginal revenue, which is increasing under Proposition A.2.1. Thus, either $s'_{ij} \geq s_{ij}$ or $s'_{ij} < s_{ij}$ and $\frac{\partial^2 R}{\partial s_{ij} q_{ij}} < \frac{\partial^2 c}{\partial s_{ij} q_{ij}}$. The latter violates the assumption of this proposition, and so $s'_{ij} > s_{ij}$. An similar argument applies when increasing $q_{i'j}$. □

The intuition behind Proposition A.2.2 is as follows. First, since i's marginal revenue of savings (quality) is increasing in the quality (savings) effort of i', i will always choose a higher savings (quality) effort when i' chooses a higher quality (savings) effort. Second, since i chooses a higher savings (quality) effort in response to a higher quality (savings) effort of i', i's marginal revenue of quality (savings) *also* increases, since $\frac{\partial R_{ij}}{\partial q_{ij}}$ ($\frac{\partial R_{ij}}{\partial s_{ij}}$) is increasing in s_{ij} (q_{ij}). Since i's marginal revenue of quality (savings) is higher, i chooses a higher quality (savings) effort.

The presence of strict strategic complementarity comes only when the ACO's savings rate and overall quality score meet or exceed the minimum savings rate and quality reporting standard. Otherwise, all participants have best response functions that are constant in the strategies of their peers. In essence, ACOs benefit from strategic complementarity when participants are all operating at a high-level of savings and quality, and when there is a relatively small tradeoff between savings and quality for the individual provider. Ultimately, the shared savings formula (defined by law) has the property that ACOs with underachieving participants obtain no advantage from strategic complementarity, but those with participants with high efforts do.

The following proposition establishes existence of equilibrium.

Proposition A.2.3. *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 $\mathbf{q}_{-ij}$)

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

for all $i \in I_j$. 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. To see this, consider the first order conditions to Problem A for all agents:

$$\frac{w_{ij}^2 B_j Q_j^A}{2} = c_{ij,1}\left(s_{ij}^A, q_{ij}^A\right) \tag{A.6}$$

$$\frac{w_{ij}^2 B_j S_j^A}{2} = c_{ij,2}\left(s_{ij}^A, q_{ij}^A\right) \tag{A.7}$$

where $c_{ij,1}$ reflects differentiation with respect to the first element. Note that since c_{ij} is strictly convex, $c_{ij,1}$ and $c_{ij,2}$ are strictly increasing and so inverse functions in a given argument exist:

$$c_{ij,1}^{-1}\left(\frac{w_{ij}^2 B_j Q_j^A}{2}, q_{ij}^A\right) = s_{ij}^A \tag{A.8}$$

$$c_{ij,2}^{-1}\left(\frac{w_{ij}^2 B_j S_j^A}{2}, s_{ij}^A\right) = q_{ij}^A \tag{A.9}$$

First, note that if c_{ij} is quadratic, $c_{ij,1}^{-1}$ and $c_{ij,2}^{-1}$ are linear, and so a unique equilibrium exists. If $c_{ij,1}^{-1}$ and $c_{ij,2}^{-1}$ are otherwise non-linear, consider the mapping $\Psi_j : [-1,1]^{n_j} \times [0,1]^{n_j} \to \mathbb{R}^{2n_j}$

$$\Psi_j\left(\mathbf{s}_j^A, \mathbf{q}_j^A\right) = \begin{bmatrix} c_{ij,1}^{-1}\left(\frac{w_{ij}^2 B_j Q_j^A}{2}, q_{ij}^A\right) - s_{ij}^A \\ c_{ij,2}^{-1}\left(\frac{w_{ij}^2 B_j S_j^A}{2}, s_{ij}^A\right) - q_{ij}^A \end{bmatrix}_{i \in I_j} \tag{A.10}$$

Clearly, zeros to the function Ψ_j are equilibria. To show that a unique zero exists, I'll use the inverse function theorem and show the Jacobian of Ψ_j has full rank at $\left(\mathbf{s}_j^A, \mathbf{q}_j^A\right)$. First, note in the diagonal entires of $D\Psi_j$ are all -1. Next, if the ith row of $D\Psi_j$ is odd, then any odd column's element in that row is zero. If the ith row is even, then any even column's element in that row is zero. Therefore, no row is a linear combination of the others, and $D\Psi_j$ has full rank. □

A.3 Influence weights w_{ij}

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

$$\sum_{i \in I_j} w_{ij} s_{ij} = S_j \qquad\qquad \sum_{i \in I_j} w_{ij} q_{ij} = Q_j \tag{A.11}$$

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

$$S_j = \frac{BE_j - E_j}{BE_j} = \frac{\sum_{i \in I_j} BE_{ij} - \sum_{i \in I_j} E_{ij}}{\sum_{i \in I_j} BE_{ij}} = \sum_{i \in I_j} w_{ij} \frac{BE_{ij} - E_{ij}}{BE_{ij}} = \sum_{i \in I_j} w_{ij} s_{ij} \tag{A.12}$$

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,

$$w_{ij} = \frac{\text{Total Spending by type } k}{(\text{Total \# of } i \text{ with type } k) \times (\sum_{\ell} \text{Total Spending by type } \ell)} \tag{A.13}$$

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 A.13 is the latter ratio.

A.4 Exit Logit

Columns (3) and (4) of Table A.1 show raw coefficient estimates and marginal effects for the logit model

$$exit_{jt+1} = \mathbf{1}\left\{\nu_0 + \nu_1 \hat{y}_{jt} + \nu_2 \mathbf{1}\left\{\hat{y}_{jt} > 0\right\} + \nu_3 age3_{jt} + \boldsymbol{\psi}' X_{jt}^{perf} + \varepsilon_{jt+1}\right\} \tag{A.14}$$

None of the elements in $\boldsymbol{\psi}$ are significant, so they are suppressed from output. I also show the result when net income $\hat{y}_j$ is replaced with an ACO's Earned Shared Savings, which is just income as opposed to net income. Note that when ACOs fail to earn shared savings, they have a 0.15 higher probability of exiting. Otherwise, dollar increases in earnings do not significantly impact exiting decisions. ACOs in there final agreement period have a 0.13 higher probability of exiting.

Table A.1: **Logit of ACO Exit**

	(1) Raw	(2) M.E.	(3) Raw	(4) M.E.
Earned Shared Savings	-0.00334	-0.000288		
	(0.0115)	(0.000986)		
$\mathbf{1}\{$Earned Shared Savings $> 0\}$	-1.430**	-0.123**		
	(0.501)	(0.0441)		
$\hat{y}_{jt}$			-0.00703	-0.000614
			(0.0141)	(0.00123)
$\mathbf{1}\{\hat{y}_{jt} > 0\}$			-1.712**	-0.150**
			(0.565)	(0.0498)
$age3$	1.557***	0.134***	1.475***	0.129***
	(0.215)	(0.0181)	(0.214)	(0.0182)
N	1063	1063	1063	1063

Robust standard errors in parentheses
$\hat{y}_j$ and Earned Shared Savings are in units of $100,000.
$^{+}$ $p < 0.10$, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

A.5 LASSO and Elastic Net for Overall Quality Score

I use both the Least Absolute Shrinkage and Selection Operator (LASSO) and Elastic Net method to compute which sub-measures of the overall quality score explain changes in the overall quality score. Formally, this takes the following steps:

1. Let ACO quality score take the form: $Q_j = \chi_0 + \sum_{m=1}^{M} \chi_m Q_{jm}$, where Q_{jm} is the mth sub-measure and $\{\chi_m\}_{m=0}^{M}$ are parameters to be estimated.

2. Elastic Net coefficients are given by

$$\{\hat{\chi}_m\}_{m=0}^{M} = \arg \min_{\{\chi_m\}_{m=0}^{M}} \sum_{j=1}^{J} \left(Q_j - \chi_0 - \sum_{m=1}^{M} \chi_m Q_{jm} \right)^2 + \lambda \left[\frac{1-\alpha}{2} \cdot \sum_{m=0}^{M} \chi_m^2 + \alpha \sum_{m=0}^{M} |\chi_m| \right] \tag{A.15}$$

where $\lambda > 0$ is a regularization parameter and $\alpha \in [0, 1]$ weights regularization on the L^1-norm of coefficients (relative to the L^2-norm). LASSO coefficients are given in the special case when this problem is solved with $\alpha = 1$. Selection of α and λ are done by *cross-validation*. This is a method where for a given λ (or α), coefficients are computed for a subsample of the data, and then out-of-sample fit is computed for the complement of the subsample. See Abadie & Kasy (2017) and Burlig et al. (2019) for more details. Elastic Net is favorable to LASSO when regressors are highly correlated. Since this is clearly the case here, Elastic Net will be my specification of choice, but I present the results to both.

Table A.2 presents the results of both specifications.

A.6 Robustness Checks

A.6.1 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.

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

Table A.2: **Regularized Regressions of Overall Quality Score on Quality Sub-Measures**

Sub-measure (Percentile 0-100)	(1) Elastic Net	(2) LASSO
ACO-2. CAHPS: How Well Your Providers Communicate	0.0310	
ACO-5. CAHPS: Health Promotion and Education	0.314	0.337
ACO-6. CAHPS: Shared Decision Making	0.391	0.382
ACO-9. Ambulatory Sensitive Conditions Admissions: Chronic Obstructive Pulmonary Disease or Asthma in Older Adults (AHRQ Prevention Quality Indicator (PQI) #5)	-1.259	-1.142
ACO-10. Ambulatory Sensitive Conditions Admissions: Heart Failure (AHRQ Prevention Quality Indicator (PQI) #8)	-2.422	-2.560
ACO-11. Percent of PCPs who Successfully Meet Meaningful Use Requirements 0.134	0.141	
ACO-13. Falls: Screening for Future Fall Risk	0.0256	0.0243
ACO-14. Preventative Care and Screening: Influenza Immunization	0.0312	0.0288
ACO-15. Pneumonia Vaccination Status for Older Adults	0.0280	0.0244
ACO-16. Preventive Care and Screening: Body Mass Index (BMI) Screening and Follow Up	0.0852	0.0883
ACO-17. Preventive Care and Screening: Tobacco Use: Screening and Cessation Intervention	0.0236	0.0195
ACO-18. Preventive Care and Screening: Screening for Clinical Depression and Follow-up Plan	0.0476	0.0487
ACO-27. Diabetes Mellitus: Hemoglobin A1c Poor Control	-0.171	-0.184
ACO-30. Ischemic Vascular Disease (IVD): Use of Aspirin or Another Antithrombotic	0.103	0.113
ACO-33. Angiotensin-Converting Enzyme (ACE) Inhibitor or Angiotensin Receptor Blocker (ARB) Therapy – for patients with CAD and Diabetes or Left Ventricular Systolic Dysfunction (LVEF<40[1em] ACO-12. Medication Reconciliation Post-Discharge	0.0401	0.0409
Constant	6.160	8.456
R^2	0.840	0.841
α	0.6842	1
λ	0.1517	0.1159

The parameter λ is found via cross validation in (1) and (2).

The parameter α is found via cross validation in (1) and set equal to 1 in (2).

random variables

$$\hat{s}_{ij} \sim N\left(s_{ij}, \sigma_S^2\right) \qquad \hat{q}_{ij} \sim N\left(q_j, \sigma_Q^2\right) \tag{A.16}$$

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

$$\max_{s_{ij}, q_{ij}} \mathbb{E}\left[R_{ij}\left(\hat{S}_j, \hat{Q}_j\right)\right] - c\left(s_{ij}, q_{ij}; x_{ij}, \boldsymbol{\theta}_2\right) \tag{A.17}$$

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 1.3.2). The objective function in Equation A.17 becomes

$$E_{\Pi}^{j}(s_{ij}, q_{ij}, S_j, Q_j) = 0.5 \cdot w_{ij} B_j \cdot E_S(S_j) \cdot E_Q(Q_j) - c_{ij}\left(s_{ij}, q_{ij}\right) \tag{A.18}$$

where

$$E_S(S_j) = \mathbb{E}\left[\hat{S}_j \mathbf{1}\left\{\hat{S}_j \geq \underline{S}_j\right\}\right] = S_j \Phi\left(\frac{S_j - \underline{S}_j}{\sqrt{W_j^{(2)}}\sigma_S}\right) + \sqrt{W_j^{(2)}}\sigma_S \phi\left(\frac{S_j - \underline{S}_j}{\sqrt{W_j^{(2)}}\sigma_S}\right) \tag{A.19}$$

$$E_Q(Q_j) = \mathbb{E}\left[\hat{Q}_j \mathbf{1}\left\{\hat{Q}_j \geq \underline{Q}\right\}\right] = Q_j \Phi\left(\frac{Q_j - \underline{Q}}{\sqrt{W_j^{(2)}}\sigma_Q}\right) + \sqrt{W_j^{(2)}}\sigma_Q \phi\left(\frac{Q_j - \underline{Q}}{\sqrt{W_j^{(2)}}\sigma_Q}\right). \tag{A.20}$$

and $W_j^{(2)} = \sum_{i \in I_j} w_{ij}^2$ (see Appendix A.3). 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.

Strategic Complementarity and Existence of Equilibrium

First define the expected revenue function.

$$E_R^j(S_j, Q_j) = 0.5 \cdot B_j \cdot E_S(S_j) \cdot E_Q(Q_j) \; . \tag{A.21}$$

Proposition A.6.1. *Let $i' \neq i$. Marginal expected revenue $\frac{\partial E_R^j}{\partial s_{ij}}(S_j, Q_j)$ is increasing in s_{ij} and $s_{i'j}$ when $\underline{S}_j \left(S_j - \underline{S}_j\right) < \sigma_S \sqrt{W_j^{(2)}}$ and is always increasing in $q_{i'j}$. Marginal expected revenue $\frac{\partial E_R^j}{\partial q_{ij}}(S_j, Q_j)$ is increasing in q_{ij} and $q_{i'j}$ when $\underline{Q}_j \left(Q_j - \underline{Q}_j\right) < \sigma_Q \sqrt{W_j^{(2)}}$ and is always increasing in $s_{i'j}$.*

Proof. First, consider the second order derivative of E_R,

$$\frac{\partial^2 E_R}{\partial s_{ij} s_{i'j}}(S_j, Q_j) = -w_{ij}^2 w_{i'j} B_j E_Q(Q_j) \left[\frac{1}{\sqrt{W_j^{(2)}}\sigma_S} + \frac{\underline{S}_j \left(S_j - \underline{S}_j\right)}{W_j^{(2)} \sigma_S^2} \right] \cdot \phi \left(\frac{S_j - \underline{S}_j}{\sqrt{W_j^{(2)}}\sigma_S} \right) . \tag{A.22}$$

The sign of this equation depends entirely on the term in the square brackets. Rearranging terms, we have

$$\underline{S}_j \left(S_j - \underline{S}_j\right) < \sigma_S \sqrt{W_j^{(2)}} \implies \frac{\partial^2 E_R}{\partial s_{ij} s_{i'j}}(S_j, Q_j) > 0$$

Since $\underline{S}_j > 0$, this condition implies that expected revenue has increasing differences in savings efforts always when average savings effort is less than the benchmark. When average savings effort is larger than the benchmark, there is still increasing differences when the difference is less than $\sigma_S \sqrt{W_j^{(2)}} / \underline{S}_j$. A similar argument applies for Q_j. □

Proposition A.6.1 states that the marginal payoff to a participant in an ACO is strictly increasing in the savings and quality of other participants for large regions of the domains of savings and quality.

Note that satisfying these properties alone do not imply that the game played by ACO participants is necessarily supermodular. That requires the additional condition

$$\frac{\partial E_R^j}{\partial s_{ij} q_{ij}}(S_j, Q_j) > \frac{\partial c_{ij}}{\partial s_{ij} q_{ij}}(s_{ij}, q_{ij}) \tag{A.23}$$

so that the best response of savings is increasing in own quality and visa versa.

As in Section 1.3, since the game played by ACO participants is generally not supermodular, I cannot use that property to prove existence of a pure strategy Nash equilibrium. Instead, I impose a restriction on the expected profit function E_Π^j to achieve existence in the following proposition.

Proposition A.6.2. *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. □

Table A.3: **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.

Identification and Estimation

Identification and estimation of $\boldsymbol{\theta}_2$ and $\boldsymbol{\theta}_1$ in this model (with uncertainty) is nearly identical to their identification and estimation outlined in Section 1.4 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.

Results

Table A.3 shows the estimates of parameters in $\boldsymbol{\theta}_2$ that describe the shape of 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.

Appendix B

Appendix to Chapter 2

B.1 Proof

Proof. Physicians maximize

$$U(x_n, x_e, y_n, y_e) = p_n x_n + p_e x_e + r p_n y_n + r p_e y_e - C(x_n, x_e, y_n, y_e) + V(x_n, x_e, y_n, y_e) \quad \text{(B.1)}$$

Let $W = C - V$ represent "cost net of altruism." Therefore, a physician solves

$$\max_{x_n, x_e, y_n, y_e} p_n x_n + p_e x_e + r p_n y_n + r p_e y_e - W(x_n, x_e, y_n, y_e) \quad \text{(B.2)}$$

which has first order conditions

$$p_n = W_1(x_n^*, x_e^*, y_n^*, y_e^*) \tag{B.3}$$

$$p_e = W_2(x_n^*, x_e^*, y_n^*, y_e^*) \tag{B.4}$$

$$rp_n = W_3(x_n^*, x_e^*, y_n^*, y_e^*) \tag{B.5}$$

$$rp_e = W_4(x_n^*, x_e^*, y_n^*, y_e^*) \tag{B.6}$$

where W_k is the partial derivative of W with respect to the kth argument. Totally differentiating the above with respect to r:

$$0 = W_{11}(\cdot)\frac{dx_n^*}{dr} + W_{12}(\cdot)\frac{dx_e^*}{dr} + W_{13}(\cdot)\frac{dy_n^*}{dr} + W_{14}(\cdot)\frac{dy_e^*}{dr} \tag{B.7}$$

$$0 = W_{12}(\cdot)\frac{dx_n^*}{dr} + W_{22}(\cdot)\frac{dx_e^*}{dr} + W_{23}(\cdot)\frac{dy_n^*}{dr} + W_{24}(\cdot)\frac{dy_e^*}{dr} \tag{B.8}$$

$$p_n = W_{13}(\cdot)\frac{dx_n^*}{dr} + W_{23}(\cdot)\frac{dx_e^*}{dr} + W_{33}(\cdot)\frac{dy_n^*}{dr} + W_{34}(\cdot)\frac{dy_e^*}{dr} \tag{B.9}$$

$$p_e = W_{14}(\cdot)\frac{dx_n^*}{dr} + W_{24}(\cdot)\frac{dx_e^*}{dr} + W_{34}(\cdot)\frac{dy_n^*}{dr} + W_{44}(\cdot)\frac{dy_e^*}{dr} \tag{B.10}$$

Because W is convex, it is positive definite, and $W_{kk} > 0$ for $k = 1, 2, 3, 4$. First, note that if $W_{kk'} = 0$ for $k \neq k'$, we trivially have that $\frac{dx_n^*}{dr} = \frac{dx_e^*}{dr} = 0$ and $\frac{dy_n^*}{dr}, \frac{dy_e^*}{dr} > 0$. This makes sense: when the provision of services is unrelated (that is, if an additional service provided to a new Medicare beneficiary doesn't in any way impact a physician's ability to provide services to other beneficiaries), the fee bump increases the Medicaid service volume and leaves Medicare service volume unchanged. On the other hand, if $W_{kk'} > 0$ for some distinct $k = 1, 2$ and $k' = 3, 4$, then $\frac{dx_n^*}{dr}$ or $\frac{dx_e^*}{dr}$ are negative.

To obtain mixed signs, for example $\frac{dx_n^*}{dr} > 0$ and $\frac{dx_e^*}{dr} < 0$ as the empirical results indicate, we need $W_{13} \ll 0$—that is, marginal cost of providing care to new Medicare beneficiaries is decreasing in the amount of care provided to new Medicaid beneficiaries. Because W is strictly convex, it is W is positive definite, so

$$p_n \frac{dy_n^*}{dr} + p_e \frac{dy_e^*}{dr} > 0 \tag{B.11}$$

Now, suppose there's no $\underline{W} < 0$ small enough such that $\frac{dx_n^*}{dr} > 0$ while $\frac{dx_e^*}{dr} < 0$, $\frac{dy_n^*}{dr} > 0$, and $\frac{dy_e^*}{dr} > 0$. Subtracting the first equation from the third yields

$$p_n = (W_{13} - W_{11}) \frac{dx_n^*}{dr} + (W_{23} - W_{12}) \frac{dx_e^*}{dr} + (W_{33} - W_{13}) \frac{dy_n^*}{dr} + (W_{34} - W_{14}) \frac{dy_e^*}{dr} \tag{B.12}$$

$$\frac{dx_n^*}{dr} = \frac{p_n + (W_{12} - W_{23}) \frac{dx_e^*}{dr} + (W_{13} - W_{33}) \frac{dy_n^*}{dr} + (W_{14} - W_{34}) \frac{dy_e^*}{dr}}{W_{13} - W_{11}} \tag{B.13}$$

Note that

$$\underbrace{\frac{dx_n^*}{dr}}_{>0} = \frac{p_n + (W_{12} - W_{23}) \frac{dx_e^*}{dr} + (W_{13} - W_{33}) \frac{dy_n^*}{dr} + (W_{14} - W_{34}) \frac{dy_e^*}{dr}}{\underbrace{W_{13} - W_{11}}_{<0}}. \tag{B.14}$$

This means the numerator is negative, and we reach a contradiction, if

$$W_{13} < -\frac{p_n + (W_{12} - W_{23}) \frac{dx_e^*}{dr} - W_{33} \frac{dy_n^*}{dr} + (W_{14} - W_{34}) \frac{dy_e^*}{dr}}{\frac{dy_n^*}{dr}}, \tag{B.15}$$

so $\underline{W}$ exists. □