Changes in Spending, Risk Selection, and the Response of Frontline Clinicians: Understanding Performance Mechanisms in the Medicare Shared Savings Program

By

Adam A. Markovitz

A dissertation submitted in partial fulfillment
of the requirements for the degree of
Doctor of Philosophy
(Health Services Organization and Policy)
in The University of Michigan
2018

Doctoral Committee:

Associate Professor Andrew M. Ryan, Chair
Professor John Z. Ayanian
Associate Professor John M. Hollingsworth
Professor Edward C. Norton

Adam A. Markovitz

amarkov@umich.edu

ORCID ID: 0000-0002-4660-6978

Acknowledgements

I dedicate this dissertation to all those who have helped me reach this point. First and foremost, to Andrew Ryan, for being everything one could hope for in a mentor and committee chair: energizing my research; never slowing us down;[1] always sharing credit for our work; providing as-needed feedback;[2] modeling how to conduct research rigorously, efficiently, and collaboratively;[3] adapting his mentorship style to my needs over time; and encouraging me to seek other mentors[4] and always happy when those collaborations bore fruit. In short, he gave me the freedom to become an independent researcher along with the guidance to become a good one.

To the additional members of my dissertation committee, for being brilliant, responsive, and constructive. John Hollingsworth provided far more data and technical support than I could have hoped for and encouraged me to do the up-front thinking necessary to ensure this work is relevant to policy and practice. Edward Norton was a precise but pragmatic teacher of econometrics, challenging me to perform rigorous research for physicians and economists alike. John Ayanian pushed me to be clear yet nuanced in my writing, somehow making time to provide rapid, thoughtful feedback while leading UM's Institute for Healthcare Policy and Innovation. As I move forward, my committee members will serve as examples of excellent scholars and physicians who perform elite research while elevating the scholarship of their peers, research fields, and

the University at large. To my honorary committee members,[4] for including me in exciting research collaborations; coaching me how to write, speak, and translate health economics and statistics research for physicians and policy makers; sponsoring me to attend national conferences and providing feedback on practice talks; connecting me to other research colleagues in academic medicine; and reminding me of the dual nature of this MD/PhD training. I would like to thank Eve Kerr in particular, as well as Tim Hofer, Jeremy Sussman, Brahmajee Nallamothu, Chad Ellimoottil, Devraj Sukul, Lena Chen, Lawrence McMahon, and Justin Dimick. I continually drew on lessons, big and small, from these mentors as I completed my dissertation research. As I move forward, these individuals will serve as important models for how to integrate medicine and health policy research in a unified career.

To those who have lent critical technical, administrative, and financial support. Ron Koenig and Ellen Elkin took a risk in allowing me and others to pursue PhDs in non-traditional fields through UM's Medical Scientist Training Program. Tim Peterson lent the organizational and financial support necessary for administering our survey to the Physician Organization of Michigan Accountable Care Organization. Susan Goold and Michael Rozier provided vital guidance on developing this survey. This work would not have been possible or as enjoyable without the support from so many others at UM's Institute for Healthcare Policy: Phyllis Yan and Nick Moloci prepared the Medicare claims data used in our evaluations of accountable care organizations; Phyllis Wright-Slaughter secured these data on our behalf; and Tedi Engler provided invaluable project support[3] to all aspects of this and other work. Michael McWilliams provided critical

feedback on analyses of changes in spending and risk selection. Funding included: the AHRQ individual dissertation award (R36 HS025615); the Horowitz Foundation for Social Policy Irving Louis Horowitz Award for Best Overall Project; the AcademyHealth Alice S. Hersh Student Scholarship; the Society of General Internal Medicine Young Scholar award; UM's MCubed Grant; and several grants from UM's Rackham Graduate School, including travel grants, a pre-candidate grant, a candidate grant, and the Ernest and Martha Hammel Award.

To my friends and family. You have centered me and propelled me through this work. To Michael and Charley, our escapes to the coffee shop have helped me to complete this work with a healthy mind. To Jesse, our escapes to the gym have helped me to complete this work with a healthy body. To Hillary and Tyler, our escapes to Detroit have taught me to carve out time for the experiences that revitalize my spirit. To my sister Rebecca, you were my first academic role model. To my sister Lara, you believed in me as a researcher before even I did. Your work for the disenfranchised inspires me. To my mother, Ruth, your unconditional love and interest in my life have helped me weather so many difficult times. To my father, David, you demonstrate every day that it is possible to be fully dedicated to both research and family, and to find time for fun in-between. Finally, to my wife, Allison, you traveled this road with me – I cannot imagine this life without you by my side.

Table of Contents

Acknowledgements ii

List of Tables vii

List of Figures ix

List of Appendices xi

Abstract xii

Introduction 1

Chapter I. Changes in spending and quality in the Medicare Shared Savings Program after accounting for clinician's non-random exit 9

Abstract 9
Introduction 12
Methods 13
Results 19
Discussion 23
Tables and Figures 28

Chapter II. Changes in coded severity and risk selection in the Medicare Shared Savings Program 34

Abstract 34
Introduction 35
Methods 36
Results 42
Discussion 46
Tables and Figures 52

Chapter III. Low-value care and clinician engagement in the Medicare Shared Savings Program: a survey of frontline clinicians 58

Abstract 58
Introduction 61

Methods 62
Results 68
Discussion 70
Tables and Figures 74

Appendices 83

Appendix A. Changes in spending and quality in the Medicare Shared Savings Program after accounting for non-random exit 84

Appendix B. Changes in coded severity and risk selection in the Medicare Shared Savings Program 125

Appendix C. Low-value care and clinician engagement in the Medicare Shared Savings Program: a survey of frontline clinicians 145

References 166

List of Tables

Table I.1. Characteristics of beneficiaries and clinicians across Medicare Shared Savings Program (MSSP) accountable care organization (ACO) participation and MSSP supply 28

Table II.1. Characteristics of beneficiaries in Medicare Shared Savings 52

Table III.1. Characteristics of respondents in accountable care organization 74

Table III.2. ACO engagement measures and scales 76

Table III.3. Examples of low-value recommendations presented to respondents 77

Table A1. Non-linear estimation of discount factor used in MSSP supply 116

Table A2. Pre-period trends in spending across observed MSSP status 117

Table A3. Formal test of statistical differences between instrumental variable 118

Table A4. Sensitivity analyses of adjusted longitudinal and instrumental 119

Table A5. Estimates of the association between prior spending and probability 120

Table A6. Simulated changes in average MSSP beneficiary spending with attrition of high-cost clinicians from MSSP ACOs (2013-2014) 121

Table A7. Simulated changes in average MSSP beneficiary spending when including clinicians exiting MSSP and excluding clinicians entering the MSSP (2013-2014) 122

Table B1. Characteristics of Medicare beneficiaries in the analytic sample vs. excluded sample (2008-2011). 134

Table B2. Relationship between beneficiary risk score and beneficiary exit or entry in the Medicare Shared Savings Program (MSSP) 135

Table B3. Growth in risk score and beneficiary entry and exit to and from the Medicare Shared Savings Program (MSSP) between 2012 and 2014 141

Table C1. Survey domains and instrument sources 148

Table C2. Awareness of ACO incentives 158

Table C3. Perceived influence of ACO on practice 159

Table C4. Perceived effect of ACO on quality 160

Table C5. List of recommendations against low-value care 161

Table C6. Characteristics of survey respondents vs. non-respondents in the ACO 165

List of Figures

Figure I.1. Changes in Medicare spending for beneficiaries attributed to MSSP ACOs vs. controls 30

Figure I.2. Changes in spending and hospitalization for hip fracture across models for fixed differences across MSSP participants vs. controls 31

Figure I.3. Changes in clinical quality performance for beneficiaries attributed to MSSP ACOs vs. controls 32

Figure I.4. Association between clinician spending performance and probability of exiting or entering an MSSP ACO 33

Figure II.1. Association between beneficiary attribution to the Medicare Shared Savings Program and change in risk score 54

Figure II.2. Association between risk score and change in probability of beneficiary entering or exiting the Medicare Shared Savings Program (MSSP) 56

Figure II.3. Growth in risk score and beneficiary entry and exit to and from the Medicare Shared Savings Program from 2012-2014 57

Figure III.1. Relationship between clinician involvement in decision to join ACO and low-value care 79

Figure III.2 Relationship between clinician ACO awareness and low-value care 80

Figure III.3. Relationship between perceived ACO impact on practice and quality and low-value care 81

Figure A1. Schematic of the MSSP supply instrumental variable 103

Figure A2. Pre-period trends in spending across the instrumental variable 104

Figure A3. Pre-period trends in spending across observed MSSP participation status 105

Figure A4. Changes in total Medicare spending according to ACOs' year of entry into the MSSP 106

Figure A5. Association between county-level MSSP penetration and Medicare Advantage penetration 107

Figure A6. Robustness tests of MSSP supply as an instrumental variable 108

Figure A7. Changes in hospital use for beneficiaries attributed to MSSP ACOs vs. controls 110

Figure A8. Sensitivity analyses of association clinician spending performance and probability of exiting an MSSP ACO 111

Figure A9. Association between provider group spending performance and probability of exiting or entering an MSSP ACO 114

Figure B1. CONSORT diagram 128

Figure B2. Sensitivity analyses of relationship between beneficiary attribution to the Medicare Shared Savings Program (MSSP) and change in risk score 129

Figure B3. Relationship between beneficiary risk score and beneficiary exit or entry in the Medicare Shared Savings Program (MSSP) among beneficiaries attributed via outpatient claims submitted by primary care providers 130

Figure B4. Relationship between clinician’s average patient panel risk score and clinician entry and exit to and from the Medicare Shared Savings Program (MSSP) 131

Figure B5. Trends in risk score across beneficiaries in the Medicare Shared Savings Program (MSSP) and controls (2008-2014) 132

Figure B6. Event study of change in risk score before and after attribution to the Medicare Shared Savings Program (MSSP) 133

Figure B7. Heterogeneity in relationship between beneficiary risk score and exit from Medicare Shared Savings Program (MSSP) 136

Figure B8. Relationship between beneficiary’s average risk score in prior two or three years and entry and exit in the Medicare Shared Savings Program (MSSP) 140

Figure C1. CONSORT diagram 147

List of Appendices

Appendix A. Changes in Spending and Quality in the Medicare Shared Savings Program After Accounting for Non-Random Exit 83

Appendix B. Changes in coded severity and beneficiary selection in the Medicare Shared Savings Program 125

Appendix C. Low-value care and clinician engagement in the Medicare Shared Savings Program: a survey of frontline clinicians 145

Abstract

Confronted by an increasingly expensive and fragmented health care system, public and private payers have established a series of reforms designed to lower costs and improve quality. A leading example is the Medicare Shared Savings Program (MSSP) establishment of Accountable Care Organizations (ACOs), where groups of providers voluntarily assume responsibility for the spending and quality outcomes of a population of Medicare beneficiaries. Studies suggest MSSP ACOs have achieved modest savings. However, prior research may overstate savings if ACOs strategically select for low-cost clinicians or drop high-cost clinicians.

In Chapter I, I evaluated the effect of the MSSP on spending and quality while accounting for clinicians' non-random entry and exit in the program, using the share of nearby clinicians in the MSSP as an instrumental variable. Hip fracture served as a falsification test. In instrumental variable models, the MSSP was not associated with spending ($6 per beneficiary-quarter; 95% confidence interval [CI], -$50, $63), any clinical quality indicator, or hip fracture (0.09 per 1000 beneficiary-quarters; CI: -0.07, 0.26). Adjusted longitudinal models — similar to what other ACO studies have estimated — failed the falsification test (-11% decrease in hip fracture), suggesting the presence of residual confounding. Highest-cost clinicians (99th percentile spending) had a 30.0% chance of exiting the MSSP compared to a 14.7% chance

among median (50th percentile) clinicians. Together, these results suggest that exit of high-cost clinicians may drive estimates of savings.

ACOs may also avoid sick or high-cost beneficiaries. To encourage ACOs to care for high-risk beneficiaries, MSSP savings benchmarks are adjusted by beneficiary risk score. To discourage coding intensification ("upcoding"), benchmarks are not adjusted upward if risk scores rise. In Chapter II, I evaluated the impact of this risk-adjustment policy. I examined whether MSSP exposure was associated with within-beneficiary risk score change and whether risk score was associated with beneficiary entry or exit. MSSP exposure was not associated with consistent changes in within-beneficiary risk score. Conversely, highest-risk beneficiaries (99th percentile of risk score) had a 25.1% chance of exiting the MSSP compared to a 16.0% chance among median-risk beneficiaries. I conclude the decision to not upwardly adjust risk score has successfully deterred coding increases but may have led ACOs to avoid high-risk beneficiaries.

Why have ACOs struggled to improve spending and quality? One possibility is ACOs have not engaged frontline clinicians. In Chapter III, I surveyed 1,620 clinicians in the Physician Organization of Michigan ACO (response rate: 34%). I found limited ACO engagement: few clinicians participated in the decision to join the ACO (3%); few were aware of ACO incentives, including knowing the ACO was accountable for spending and quality (23%) or faced upside risk only (3%); and few agreed (moderately or strongly) the ACO changed compensation (20%) or practice (19%) or reduced inappropriate care (13%). However, increased ACO awareness (one standard

deviation) was associated with decreased difficulty implementing recommendations against low-value care (-2.3 percentage point; CI: -3.8, -0.8). I conclude limited clinician engagement may hamper efforts to reduce low-value care.

Collectively, these results suggest that the MSSP has had little impact on spending and quality. Instead, selection effects – including strategic dropping of high-cost clinicians and high-risk beneficiaries – may drive estimates of improved performance. Future research should investigate how Medicare can structure incentives that motivate ACOs to lower spending, care for high-risk beneficiaries, and engage frontline clinicians.

Introduction

A remarkable consensus is developing – the current approach of separately paying physicians and hospitals for discrete services has inflated spending and fragmented care[5–9] Yet this consensus hides disagreement about what should replace fee-for-service (FFS) payment. At the heart of this disagreement lies an evidence gap: research is only beginning to understand which alternative payment models have the most potential to improve value.[10–13] The Centers for Medicare and Medicaid Services (CMS) have most aggressively initiated payment reforms among FFS Medicare beneficiaries. Yet evaluating these reforms has proven challenging. CMS rarely randomizes, and providers have considerable discretion in whether and how they elect to participate in these reform. As a result, evaluations of these programs are subject to bias from volunteering effects and unintended consequences such as severity upcoding and patient cherry-picking.[14,15]

Of the many reforms proposed to FFS, none has drawn more attention than Accountable Care Organizations (ACOs).[16–20] Under this payment model, groups of providers assume collective responsibility for the spending and quality outcomes of a defined group of patients. A leading example is the Medicare Shared Savings Program (MSSP), where ACOs receive shared saving bonuses if they lower spending below a

financial benchmark for attributed FFS Medicare beneficiaries while maintaining quality standards. Studies suggest that MSSP ACOs are associated with modest savings.[21–24]

However, prior research may overstate savings in this voluntary program if ACOs strategically select for low-cost clinicians or drop high-cost clinicians. Clinicians and provider groups may choose to join, or be recruited to join, ACOs because of their ability to deliver efficient care. Conversely, the ability of MSSP ACOs to drop clinicians with high-cost patients provides ACOs with a straightforward means of lowering measured spending.[25] In Chapter I of my thesis, I evaluated the effect of the MSSP on spending and quality while accounting for clinicians' non-random entry and exit to and from MSSP ACOs. To account for selection effects, I used the share of nearby clinicians participating in the MSSP as an instrumental variable. I also used hip fracture as a falsification test. [26] Because hip fracture incidence should not be affected by medical practice (or ACOs) in the near term and invariably requires hospitalization (i.e., not subject to variation in diagnostic intensity), any observed association between the MSSP and hip fracture would suggest the presence of residual confounding between ACOs and controls not addressed by a given statistical model. [27–29]

In instrumental variable models, the MSSP was not significantly associated with changes in spending ($6 per beneficiary-quarter; 95% CI, -$50 to $63), any clinical quality indicator, or hip fracture (hospitalization rate of 0.09 per 1000 beneficiary-quarters; 95% CI: -0.07 to 0.26). Adjusted longitudinal models — similar to what other

studies have used to estimate the impact of ACOs — failed the falsification test (hip fracture hospitalization, -0.24 per 1000 beneficiary-quarters; 95% CI: -0.33 to -0.17), suggesting residual confounding from unobserved patient severity. Prior to the start of the MSSP, spending trends were also different between beneficiaries who did and did not eventually enter the MSSP. In supplemental analyses, I found that high-cost clinicians were disproportionately likely to exit MSSP ACOs. Together, these results suggest that selection effects – including strategic dropping of high-cost clinicians – may drive estimates of savings in the MSSP.

ACOs may also try to avoid sick beneficiaries or beneficiaries with high anticipated spending.[25] To encourage ACOs to care for high-risk beneficiaries, savings benchmarks in the MSSP are adjusted by beneficiaries' Centers for Medicare and Medicaid Services (CMS) Hierarchical Condition Category (HCC) risk score. To minimize ACOs' incentive to raise benchmarks by increased diagnostic coding,[15,30–32] CMS ruled that the benchmark cannot be adjusted upward if the risk score rises while the beneficiary is in the MSSP. [33–35] If the risk score falls, however, the benchmark is adjusted downward. Because CMS' approach does not capture growth in risk score over time, ACOs face an incentive to drop high-risk beneficiaries with increasing comorbidities or acute illness, i.e., "favorable selection." The fact that beneficiaries are attributed to ACOs at the end of the performance year rather than its beginning (i.e., retrospective attribution) based on a known subset of claims submitted by the ACOs themselves provides ACOs with a means of avoiding beneficiaries who become acutely sick or more expensive during the performance period. At the same time, because this risk adjustment system still

penalizes ACOs whose beneficiaries' risk scores (and spending benchmarks) fall, ACOs must remain vigilant in maintaining coding intensity at or above current levels.

In Chapter II of my thesis, I examined the relationship between the MSSP and risk profiles in a cohort of beneficiaries continuously enrolled in FFS Medicare from 2008 through 2014. I found limited evidence that exposure to the MSSP increased within-beneficiary risk score. At the same time, high-risk beneficiaries were disproportionately likely to exit MSSP ACOs: highest-risk beneficiaries (99th percentile of risk score) had a 25.1% chance of exiting the MSSP compared to a 16.0% chance among median-risk beneficiaries (50th percentile). MSSP exit was particularly concentrated among beneficiaries who exhibited high growth in risk score while in the MSSP and following exit from the program. Additional analyses suggested that exit by clinicians with high-risk panels drove exit of high-risk beneficiaries from the MSSP. These findings suggest that CMS' decision to not upwardly adjust risk score in the MSSP has successfully deterred coding increases but may have led ACOs to avoid high-risk beneficiaries with increasing comorbidities or acute illness.

Evidence from Chapters II and III suggest ACOs have had difficulty improving spending and quality in the MSSP. Instead, it appears that ACOs have relied, at least in part, on strategically dropping high-cost clinicians and high-risk beneficiaries to reduce measured spending and achieve shared savings.

Why have MSSP ACOs had so much difficulty achieving meaningful improvements to spending and quality? One possibility is that ACOs have struggled to engage the frontline clinicians charged with achieving ACO objectives. Since the inception of the ACO model, policymakers and physician leaders have commented on the need to increase clinician awareness of ACO goals, provide useful performance data, and foster effective incentives aligning clinician and ACO objectives.[7,9,19,36–44] Whether this has occurred is largely unknown. Most empirical studies have relied on executive surveys and interviews with ACO leadership,[10,20,45–48] with only one survey targeting individual clinicians (i.e., primary care physicians and internal medicine specialists) during the early ACO experience (2014-2015).[48]

In Chapter III of this thesis, I examined the perspective of frontline clinicians, a dimension of ACOs that has been largely missing from policy conversations. I designed and administered a novel survey to individual clinicians in the Physician Organization of Michigan MSSP ACO. With approximately 80,000 attributed beneficiaries, the POM ACO is the largest MSSP ACO in Michigan and among the ten largest in the United States.[49] To evaluate the mechanisms by which ACOs' collective incentives are transmitted to individual clinicians, I integrated previous methodologies from surveys that have separately assessed clinicians' engagement in payment reform and clinicians' ability to provide high-value care.

In the largest survey of ACO clinicians to date (n=1,620 respondents), I found limited ACO engagement of frontline clinicians: few clinicians participated in the decision to join the ACO (3%); few were aware of ACO incentives, including knowing that the ACO was accountable for both spending and quality (23%), successfully lowered spending (9%), or faced upside risk only (3%); and few reported the ACO changed compensation (moderate to strong agreement, 20%), overall practice (19%), or feedback (15%) or that it improved care coordination (17%), inappropriate care (13%), or unnecessary hospitalizations (12%). At the same time, greater clinician engagement – in particular, greater awareness of ACO incentives and perceived ability to improve care quality – was associated with less difficulty implementing recommendations against low-value services. For example, increased ACO awareness (one standard deviation) was associated with decreased difficulty (-2.3 percentage point) implementing recommendations against low-value care (95% CI: -3.8, -0.8), representing an 8% improvement. Taken together, our results suggest that limited engagement of ACO clinicians may hamper ACO efforts to achieve spending and quality.

Based on this work, what can I conclude about ACOs and their potential to improve upon the FFS paradigm? The pathologies of FFS structure are well-known.[50] Yet the ACO model always risked introducing new problems. ACOs, like any global payments holding providers accountable for total costs of care, raise the possibility that providers will avoid sick beneficiaries or drop the clinicians that care for them.[30,51] While CMS sought to lessen this incentive by using beneficiary risk score to adjust MSSP benchmarks, concerns about coding intensification led CMS to not upwardly adjust

benchmarks once beneficiaries are in the MSSP. [33–35] ACOs still face an incentive to avoid patients with increasing comorbidities or acute illness. Evidence from this thesis suggests that ACOs have acted upon this incentive: high-cost clinicians and their patient panels are more likely to exit the MSSP. Collectively, these findings suggest that MSSP ACOs may be strategically dropping high-cost clinicians and avoiding high-risk beneficiaries.

The tension between appropriate risk adjustment and inappropriate changes in coding practices, already evident during early MSSP rulemaking, will increase rapidly as the MSSP transitions to regional benchmarks and two-sided risk contracts. What can be done to reduce favorable selection of beneficiaries and clinicians while minimizing coding changes in the MSSP? The most obvious solution is better risk adjustment. For instance, CMS could incorporate additional social risk factors (e.g., race) into the diagnosis-based risk score methodology.[52,53] However, this would not capture growth in chronic or acute illness, which I found to be strongly associated with MSSP exit independent of race. Instead, CMS could adjust for beneficiaries' pre-existing "risk velocities" (e.g., growth in patient risk score) in the years prior to MSSP attribution.

A second solution is making it more difficult for ACOs to avoid high-risk beneficiaries or drop clinicians with high-risk patient panels. Prospective attribution – used in the Next Generation and MSSP Track 3 models – may help prevent ACOs from dropping beneficiaries during the performance year and prior to attribution. Likewise, the MSSP

could modify or rescind ACOs' ability to modify the roster of ACO provider participants year-to-year over the course of a single ACO contract.[25] Finally, CMS could couple efforts to improve risk adjustment and reduce favorable selection with aggressive auditing of compositional changes and coding practices in the MSSP. Although CMS has repeatedly stated in each MSSP Final Rule that it plans to do this,[33–35] it is uncertain whether this has occurred.

Encouraging clinicians and hospitals to take accountability for spending and quality while minimizing unintended consequences remains a central challenge in health care reform.[30,54] The evidence presented in this thesis suggest that the MSSP has had little success in achieving this objective and may have prompted ACOs to avoid the beneficiaries most likely to benefit from efforts to improve spending and quality. Nonetheless, the ACO experiment remains relatively new. It is unknown if the MSSP or future ACO models can better balance these competing objectives. Only through continued rigorous policy evaluation will we be able to answer these questions and improve the value of US health care.

Chapter I. Changes in spending and quality in the Medicare Shared Savings Program after accounting for clinician's non-random exit

Abstract

Importance: Medicare Shared Savings Program (MSSP) Accountable Care Organizations (ACOs) are associated with modest savings. However, prior research may overstate this effect if high-cost clinicians disproportionately exit MSSP ACOs.

Objective: To evaluate the effect of the MSSP on spending and quality while accounting for clinicians' non-random exit from MSSP ACOs.

Design, settings, participants: A 20% random sample of Medicare fee-for-service beneficiaries was used to compare performance among MSSP ACOs versus controls from 2008 through 2014 (N=97,795,756 beneficiary-quarters). Similar to prior MSSP analyses, adjusted longitudinal models accounted for secular trends, market factors, and beneficiary characteristics. To account further for selection effects, the share of nearby clinicians participating in the MSSP was used as an instrumental variable. We also tested for compositional changes among MSSP participants and whether these

changes affected spending estimates. Changes in rates of hip fracture served as a falsification outcome.

Exposure: Beneficiary attribution to an MSSP ACO.

Main outcomes. Total spending, 4 clinical quality indicators, and hip fracture.

Results: In adjusted longitudinal models, the MSSP was associated with modest reductions in spending (-$117 per beneficiary-quarter; 95% CI, -$150, -$85) and improvements in 2 out of 4 clinical quality indicators. In instrumental variable models, the MSSP was not significantly associated with changes in spending ($6 per beneficiary-quarter; 95% CI, -50, 63) or any clinical quality indicator. In falsification tests, the MSSP was associated with changes in hip fracture in the adjusted model (-0.24 per 1000 beneficiary-quarters; 95% CI: -0.32, -0.17) but not instrumental variable model (0.06 per 1000 beneficiary-quarters; 95% CI: -0.10, 0.20). Estimated savings decreased in models accounting for fixed differences across beneficiaries (-$88 per beneficiary-quarter), ACOs (-$57 per beneficiary-quarter), and clinicians (-$31 per beneficiary-quarter; all $P<.001$). Compositional changes were driven by ACOs dropping high-cost clinicians: clinicians in the 99th percentile of spending had a 30.0% chance of exiting the MSSP, compared to a 14.7% chance among median-cost (50th percentile) clinicians (risk difference, $P<.001$).

Conclusions and relevance: After accounting for clinicians' non-random entry and exit to and from the program, the MSSP was not associated with improvements in spending or quality. Selection effects – including exit of high-cost clinicians – may drive estimates of savings in the MSSP.

Introduction

Accountable care organizations (ACOs) are arguably the most widespread and far-reaching value-based reform in the United States, with over 900 ACO contracts covering more than 32 million lives.[17] A leading example is the Medicare Shared Savings Program (MSSP), in which groups of clinicians, hospitals, and other providers voluntarily assume responsibility for the spending and quality outcomes of a defined population of fee-for-service (FFS) Medicare beneficiaries. Evidence suggests that MSSP ACOs are associated with modest spending and quality improvements.[21–23]

However, evaluations of the MSSP may be subject to confounding from non-random participation or attrition within ACOs. Clinicians and provider groups may choose to join, or be recruited to join, ACOs because of their desire or ability to deliver high-quality and efficient care.[48] These providers may also simultaneously engage in other payment reforms such as Medicare Advantage (MA) that could lower spending.[55,56] Conversely, ACOs may selectively drop high-cost clinicians from their contracts to reduce measured spending and earn shared savings.[25] Despite these selective pressures operating at the clinician, provider group, and MSSP contracting level, research to date has not identified the degree to which selection bias may affect evaluations of the MSSP.

In this study, we used national Medicare data to evaluate changes in spending and quality performance while accounting for selection effects in the MSSP. Similar to previous analyses, we estimated adjusted longitudinal models that account for secular

trends, market-level factors, and observed differences across MSSP participants and local controls. To account for clinicians' non-random entry and exit to and from the MSSP, we used the share of clinicians participating in the MSSP within a 50-mile radius of a given clinician's practice location as an instrumental variable. We also tested for changes in the composition of MSSP participants over time and whether these compositional changes drive estimates of spending.

Methods

Data Sources and Study Population. We analyzed national claims data from 2008 through 2014 for a random 20% sample of Medicare FFS beneficiaries 65 years of age or older. MSSP participation was defined using CMS' ACO Provider- and Beneficiary-level files, which list the beneficiaries, clinicians, and provider groups participating in the MSSP. Following CMS' MSSP specifications [57], ACO beneficiaries were attributed to the provider group and clinician (defined by the Taxpayer Identification Number [TIN]) within their ACO from which they received the plurality of eligible evaluation and management services.. An analogous method was followed for attributing control beneficiaries to non-MSSP provider groups and clinicians. To improve comparability to prior work [21,22,58], we restricted analyses to beneficiaries attributed using claims submitted by a primary care clinician (i.e., step one of the MSSP attribution process) and in the outpatient setting. (Details are provided in the Methods of Appendix A.)

As specified by the MSSP,[57] we excluded beneficiaries who were ineligible for attribution to an MSSP ACO, enrolled in Medicare Advantage, or not continuously enrolled in Medicare Parts A and B. To improve comparability between MSSP beneficiaries and controls, we also excluded beneficiaries residing outside of a hospital referral region (HRR).

Study Variables

Primary and Secondary Outcomes. Our primary study outcome was total price-standardized Medicare spending per beneficiary per quarter. Price-standardization accounted for variation resulting from regional wage indices and payments for indirect medical education, Disproportionate Share Hospitals, and new technologies.[59]

To ensure comparability to previous studies,[21–23,60] we included the following secondary outcomes: (A) component spending for inpatient, outpatient, professional, and skilled nursing facility services; (B) quality indicators, including indicators for diabetes (glycated hemoglobin testing, low-density lipoprotein cholesterol testing, diabetic retinal examinations) and mammography; and (C) hospital utilization, including rates of all-cause hospitalization, preventable hospitalization (hospitalization for 1 of 11 ambulatory care-sensitive conditions),[61] all-cause 30-day readmissions,[62] and emergency department visits.

As a falsification test, we evaluated the effect of the MSSP on the rate of hospitalization for hip fracture.[26] We selected hip fracture as a falsification outcome for two reasons. First, because medical practice is unlikely to affect hip fracture in the near term, recent reforms such as ACOs are unlikely to affect population incidence rates during our study period. Thus, observed changes in hip fracture would suggest the presence of residual confounding between ACOs and controls. Second, because hip fracture requires hospitalization, it is less subject to changes in diagnostic intensity or treatment preferences potentially correlated with ACO participation.[27] Previous research has demonstrated hip fracture hospitalization rates to be a sensitive indicator of true population incidence.[28,29] (See Methods in Appendix A.)

Exposure. We defined beneficiary attribution to the MSSP using a time-varying indicator that equaled 1 if the beneficiary was attributed to an MSSP ACO in a given quarter, 0 otherwise. This captured ACOs' staggered entry into MSSP contracts (April 2012, July 2012, January 2013, January 2014) and the ability of participants to enter and exit ACOs each year. Following reports that MSSP savings are concentrated among early entrants,[21–23] we also evaluated spending changes across ACOs' year of entry into MSSP contracts.

Instrumental Variable – MSSP Supply. We hypothesized that clinicians practicing closer to other MSSP providers would be more likely to join and remain in MSSP ACOs but would but would otherwise be similar to local clinicians. Our instrument is conceptually similar to the canonical "differential distance" instrument,[63–65] which

exploits the fact that beneficiaries residing relatively closer to specialized hospitals versus other hospitals are more likely to receive specialized care. Similar to the use of differential distance to address non-random selection of patients into treatments, we hypothesized that "MSSP supply" could address non-random selection of clinicians and provider groups into and out of MSSP ACOs.

In each quarter, MSSP supply was defined as the ratio of MSSP clinicians to all clinicians within a 50-mile radius of the clinician's practice location, after discounting the supply of clinicians who were relatively further away within the 50-mile radius.[66] The discount factor was estimated by modeling how the likelihood of two clinicians participating in the same MSSP ACO varied by the inverse distance between those two clinicians[66,67] and was allowed to vary across urban, large rural, small rural, or isolated ZIP codes (Methods, Figure A1, and Table A1 of Appendix A).[68–70] We tested the robustness of the instrumental variable using MSSP supply restricted to a 10-, 25-, 50-, 75-, or 150-mile radius, with or without discounting. In contrast to standard longitudinal models, instrumental variable models capture treatment effects among marginal clinicians, i.e., those participating in the MSSP due to greater exposure to MSSP supply.

Statistical Analysis. We estimated 2 sets of linear regression models for each spending and quality outcome. First, similar to previous evaluations, we estimated adjusted longitudinal models comparing concurrent changes in spending and quality between MSSP participants and local controls within the same health care market, as

defined by HRR. [12,21–23,58,71] These models included a quarterly indicator for MSSP attribution, market fixed effects, year fixed effects, seasonal indicators, and beneficiary characteristics. As in prior studies, we adjusted for beneficiary age, sex, race/ethnicity (white, black, Hispanic, or other/unknown), disability, end-stage renal disease, dual-eligibility for Medicaid (months of enrollment), the CMS Hierarchical Condition Category (HCC) risk score, and area-level poverty (proportion below federal poverty level) and education (proportion graduated from high school, college).[72]

Next, to examine whether non-random entry and exit to and from the MSSP affects estimates of spending and quality in the MSSP, we estimated instrumental variable models that were analogous to the adjusted longitudinal model but used MSSP supply as an instrumental variable. By integrating MSSP supply into a longitudinal framework, these models specifically used within-market changes to MSSP supply to identify the effect of the MSSP, thus accounting for secular trends and fixed differences across markets that could confound the relationship between MSSP supply and local performance changes. (Model details are provided in Methods in Appendix A.)

As a falsification test, we estimated the effect of the MSSP on the rate of hospitalization for hip fracture using the adjusted longitudinal model and the instrumental variable model. Because the MSSP should not affect hip fracture rates during our study period, observed associations would suggest the presence of residual confounding not fully addressed by the given statistical model.

Compositional changes. To identify the degree to which changes in the composition of MSSP participants drives performance estimates, we estimated changes in spending and hip fracture (the falsification outcome) using supplemental longitudinal models that included: (1) market-year fixed effects, capturing market-specific spending trends; (2) beneficiary fixed effects; (3) market-year and ACO fixed effects, with ACOs defined by groups of TINs ultimately forming MSSP ACOs;[21,22,60] or (4) clinician fixed effects (Methods in Appendix A). Differences across these models would indicate the influence of compositional changes on standard longitudinal estimates. Models including market-year and ACO fixed effects are most comparable to those used in the principal MSSP evaluations.[12,21–23,58,71] We tested instrumental variable robustness in a similar manner, estimating changes in spending and hip fracture in models that included: (1) market-year fixed effects; (2) market-year and ACO fixed effects; or (3) MSSP supply restricted to a 10-, 25-, 50-, 75-, or 150-mile radius, with or without discounting.

In supplemental analyses, we evaluated two potential mechanisms underlying changes in the composition of MSSP participants or their practice environment. First, we assessed whether MSSP ACOs strategically recruit or prune providers according to spending performance in the program. Specifically, we tested whether a clinician's or provider group's spending performance in the prior year predicted the probability of the clinician or provider group entering or exiting an existing MSSP ACO in the subsequent year. (Details provided in Methods in Appendix A.) Second, we analyzed whether MA penetration, a proxy for managed care, was associated with MSSP penetration at the county level. For these analyses, we estimated linear regression models with year fixed

effects and either market fixed effects (MSSP entry/exit) or county fixed effects (MA penetration analyses). If accounting for county-level MA penetration reduced savings estimates of ACOs, it would suggest that other time-varying changes occurring alongside ACOs may introduce bias into standard estimates.

All analyses specified robust standard errors to account for clustering at the market level. Statistical analyses were performed using Stata version 15.1. Our study was deemed exempt from review by the University of Michigan Institutional Review Board.

Results

MSSP Participation and Spending. Medicare FFS beneficiaries contributed 97,795,756 beneficiary-quarters from 2008 through 2014 (6,105,325 unique beneficiaries). MSSP participants included 852,964 beneficiaries, 30,813 clinicians, 7,086 provider groups, and 337 ACOs from 2012 through 2014. Average Medicare spending per beneficiary-quarter was $2,336 (standard deviation: $7,319).

Instrument Validity. Before estimating instrumental variable models, we confirmed the validity of MSSP supply as an instrumental variable.[73] First, MSSP supply was highly correlated with MSSP participation ($F_{1,306}$ statistic = 444), where instruments with F-statistics above 10 are considered strong.[74] Second, beneficiary covariates were extremely well-balanced across MSSP supply (Table I.1). Third, spending trends in the pre-MSSP period were similar between beneficiaries with high versus low MSSP supply

(differential trend, –$2 per beneficiary-quarter; 95% CI: -$4 to -$1; Figure A2 and Table A2 in Appendix A).

In contrast, across observed MSSP status, MSSP beneficiaries were less likely to be dual-eligible or disabled, have substantial comorbidity, or live in areas with high poverty or low educational attainment (Table I.1). Further, spending trends in the pre-period differed substantially between beneficiaries who did and did not eventually enter the MSSP (differential trend, $18 per beneficiary-quarter; 95% CI: $16 to $19; Figure A3 and Table A2 in Appendix A), a violation of the parallel trends assumption. These findings suggest that unobserved sources of selection bias were more evenly distributed across MSSP supply than observed MSSP status.

Changes in Spending. Figure I.1 shows the association between the MSSP and spending in the adjusted longitudinal and instrumental variable models. In the adjusted longitudinal model, the MSSP was associated with a modest reduction in total spending (-$117 per beneficiary-quarter; 95% confidence interval [CI], -$150 to -$85). Savings in the adjusted model were due to reductions in inpatient (-$58 per beneficiary-quarter; 95% CI: -$73, -$43), outpatient (-$17 per beneficiary-quarter; 95% CI: -$25, -$8), and skilled nursing facility (-$41 per beneficiary-quarter; 95% CI: -$55, -$26) services. In the instrumental variable model, the MSSP was not associated with changes in total spending ($6 per beneficiary-quarter; 95% CI, -$50 to $63). The instrumental variable spending estimate differed significantly from the adjusted estimate (difference-in-Sargan

test, *P*<.001; Table A3 in Appendix A).[75] Estimated savings were smaller in models across ACOs' year of entering the MSSP (Figure A4 in Appendix A).

Falsification Test. Figure I.1 shows the results of the falsification test. In the adjusted longitudinal model, the MSSP was associated with a substantial decrease in the falsification outcome, rate of hospitalization for hip fracture (-0.24 per 1000 beneficiary-quarters; 95% CI: -0.32, -0.17). This represents an 11% decrease in the rate of hip fracture hospitalization (≈ -0.24 / adjusted mean of 2.12). In contrast, the MSSP was not associated with hip fracture in the instrumental variable model (0.05 per 1000 beneficiary-quarters; 95% CI: -0.10, 0.20). The instrumental variable estimate for hip fracture differed significantly from the adjusted estimate (difference-in-Sargan test, *P*=.001; Table A3 in Appendix A).[75]

Effects of Compositional Change. Figure I.1 shows the large influence of compositional changes on estimates of the MSSP. There was a consistent decrement in the association between the MSSP and hip fracture (the falsification outcome) in models that controlled for fixed differences across market-years (-0.29 per 1000 beneficiary-quarters; 95% CI: -0.37, -0.20; Panel A), beneficiaries (-0.20 per 1000 beneficiary-quarters; 95% CI: -0.28, -0.12), ACOs (-0.16 per 1000 beneficiary-quarters; 95% CI: -0.23, -0.09), and clinicians (-0.06 per 1000 beneficiary-quarters; 95% CI: -0.12, -0.01). The association between the MSSP and spending (Panel B) followed a similar pattern: estimates of total savings decreased in models that accounted for fixed differences across market-years (-$124 per beneficiary-quarter), beneficiaries (-$88 per beneficiary-

quarter), ACOs (-$59 per beneficiary-quarter), and clinicians (-$31 per beneficiary-quarter; all *P*<.001).

In contrast, the MSSP was not associated with spending or hip fracture in the instrumental variable model. This finding was robust across a wide range of specifications, including instrumental variable models using market-year fixed effects, market-year and ACO fixed effects, or MSSP supply restricted to a 25-, 50-, 75-, or 150-mile radius, with or without discounting (Figure A6 and Table A4 in Appendix A). Although changes in MSSP and MA penetration were modestly associated at the county level (Figure A5 in Appendix A), adjustment for MA penetration did not affect adjusted longitudinal or instrumental variable estimates (Table A4 in Appendix A).

Changes in Quality and Hospital Use. The MSSP was associated with improvements in two of the four clinical quality indicators in the adjusted longitudinal model but not the instrumental variable model (Figure I.3). In the adjusted model, the MSSP was associated with increased rates of glycated hemoglobin testing among beneficiaries with diabetes (2.4 percentage point [p.p.] per quarter; 95% CI: 0.3, 4.5) and mammography (0.6 p.p. per quarter; 95% CI: 0.3, 1.0). In the instrumental variable model, the MSSP was not associated with changes in any quality indicator, with estimates for all four indicators differing significantly from adjusted estimates (Table A3 in Appendix A). The MSSP was associated with modest decreases in all-cause hospitalizations and preventable hospitalizations in the adjusted model but not the instrumental variable model; conversely, reductions in all-cause 30-day readmissions

and emergency department visits were observed in both adjusted and instrumental variable models (Figure A7 in Appendix A).

Participation and Attrition Within MSSP ACOs. In 2013 and 2014, 4,297 and 4,574 clinicians exited and entered already-formed MSSP ACOs, respectively. Figure I.4 shows the relationship between a clinician's spending performance in the MSSP and subsequent exit or entry from the program. Clinicians in the 95th and 99th percentile of average spending had a 23.0% and 30.0% chance of exiting the MSSP in the next year, as compared to a 14.7% chance among median-spending (50th percentile) clinicians, a risk difference of 8.4% and 15.3%, respectively, both *P*<.001; Table A5 in Appendix A). Results were robust when using clinicians' average spending performance in the prior one or two years or when not adjusting for average beneficiary characteristics (Figure A8 in Appendix A). Entry into already-formed ACOs was also modestly higher among clinicians in the 99th percentile of prior spending as compared to median-spending clinicians (risk difference, 3.6%, *P*<.001; Figure 4 and Appendix Table A5 in Appendix A). Conversely, MSSP entry and exit by provider groups was less common and not associated with groups' prior spending (Figure A9 in Appendix A).

Discussion

Participation in MSSP ACOs was not associated with improvements in spending, quality, or hip fracture (the falsification outcome) in our instrumental variable analysis. In contrast, adjusted longitudinal models — similar to what other studies have used to

estimate the impact of ACOs — failed our falsification test. Prior to the start of the MSSP, spending trends were also different between beneficiaries who did and did not eventually enter the program. Supplemental analyses found that high-cost clinicians were disproportionately likely to exit MSSP ACOs. Together, our results suggest that improved quality and spending performance in this voluntary program may have been driven by clinicians' non-random participation and attrition in the MSSP.

Our conclusion that the MSSP was not associated with improvements in spending, quality, or most measures of hospital use differs from previous Medicare ACO evaluations.[12,21–23,58,71] Our instrumental variable models address selection effects not directly captured in previous evaluations. The influence of compositional changes within the MSSP is highlighted by the consistent reduction in estimates of both savings and hip fracture that occurred with progressively greater adjustment for fixed differences across MSSP participants and controls. Instrumental variable models may address these compositional changes by evaluating spending and quality among clinicians who join and remain in the MSSP due to their proximity to other MSSP clinicians rather than their record of spending performance in the program.

The ability of MSSP ACOs to recruit or drop clinicians provides ACOs with a relatively straightforward means of lowering measured spending.[25] Our finding that high-cost clinicians disproportionally exit the MSSP is consistent with a recent study showing that clinicians with higher predicted spending were more likely to leave Partners HealthCare Pioneer ACO.[25] Conversely, we found no evidence that MSSP ACOs recruit low-cost

clinicians. This may be because ACOs can more readily observe spending of clinicians already in their ACO than the performance of those who have not yet joined.

Pruning high-cost clinicians from ACO contracts could have large effects on spending estimates and may contribute to reported findings that MSSP savings grow over time.[22] For example, dropping clinicians in the top 95th or 99th spending percentiles from the MSSP would lower average ACO spending per beneficiary in our sample by 3.1% or 0.6%, respectively (Table A6 in Appendix A). Conversely, average ACO spending per beneficiary would increase by 1.1% in our sample if we included clinicians that exited the MSSP in our sample and excluded those clinicians that entered (Table A7 in Appendix A). These changes are similar in magnitude to previous estimates of total savings in the MSSP.[12,21–23,71]

Possible limitations of our analysis merit discussion. First, as with voluntary MSSP participation status, some unobserved factor may confound the relationship between MSSP supply and performance. Markets with greater MSSP supply likely contain more clinicians capable of implementing value-based initiatives. However, because our instrumental variable models used only within-market MSSP supply changes over time, the instrumental variable is not subject to this concern. Moreover, in contrast to the adjusted longitudinal model, MSSP supply displayed excellent covariate balance, parallel spending trends, and was not associated with changes in the falsification outcome. Second, instrumental variable models typically generate less precise estimates than those from adjusted models, risking type II error. However, instrumental

variable estimates of spending, quality, and hip fracture all differed significantly from adjusted longitudinal estimates.[75] Third, it is possible that changes in hip fracture hospitalization, our falsification outcome, represent true improvements by ACOs to underlying beneficiary health. However, this is unlikely during our brief post-period and in light of ACOs' relatively modest improvements to other health outcomes.[12,21–23,58,71] Fourth, because we attributed beneficiaries to MSSP ACOs using CMS' official participant lists of MSSP beneficiaries and clinicians, our results may differ from initial MSSP evaluations that used their own attribution methodology [21–23,60]. However, assessing whether MSSP ACOs strategically recruit or drop high-cost clinicians required that we use CMS' official participants lists.

Finally, because instrumental variable models estimate treatment effects among marginal clinicians, our findings may not generalize to clinicians who would either always join the MSSP (e.g., because they are employed by an early adopter health system) or never join the MSSP (e.g., because they lack the desire or capacity to undertake such a reform). At the same time, this population of clinicians — those joining the MSSP due to greater engagement with and access to other MSSP clinicians — is of intrinsic policy relevance. That the MSSP had little effect among these participants is directly relevant for assessing the potential economic and health rewards of governmental policies intended to promote ACO participation, such as the 2015 enactment of the Medicare Access and CHIP Reauthorization Act (MACRA).[8]

Our results challenge the view that MSSP ACOs have lowered spending and improved quality, suggesting that improved performance by MSSP ACOs may be driven by clinicians' non-random entry and exit to and from this voluntary program. These findings suggest caution in extending ACOs to other settings and patients without stronger evidence that the program saves money or improves quality of care. Our study underscores the degree to which selection bias may affect evaluations of voluntary reforms and the challenges inherent in evaluating these programs.

Table I.1. Characteristics of beneficiaries and clinicians across Medicare Shared Savings Program (MSSP) accountable care organization (ACO) participation and MSSP supply[a]

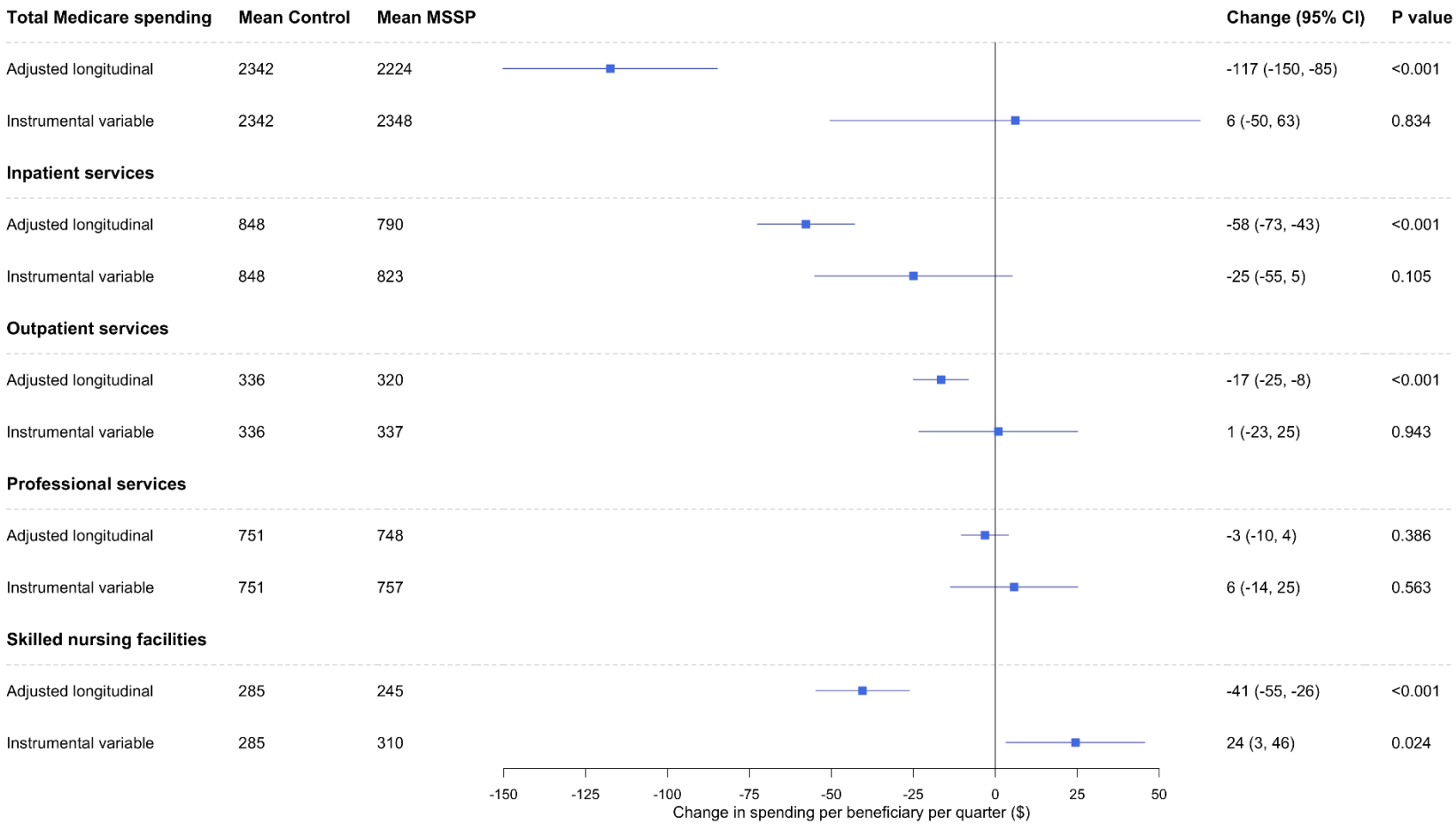
	Covariate balance across participants and non-participants[b]			Covariate balance across the instrumental variable[c]		
Characteristics	**Controls[b] (N=48,121,396 beneficiary-quarters)**	**MSSP[b] (N= 8,547,645 beneficiary-quarters)**		**Low MSSP supply[c] (N=32,782,038 beneficiary-quarters)**	**High MSSP supply[c] (N=23,887,002 beneficiary-quarters)**	
Patient characteristics	**Unadjusted mean**	**Unadjusted mean**	**Standardized difference**	**Unadjusted mean**	**Unadjusted mean**	**Standardized difference**
Age, y	76.8	75.4	-0.20	76.6	76.6	0.00
Female	60.2	60.9	0.01	60.3	60.4	0.00
Race/Ethnicity[d]						
Non-Hispanic white	85.8	86.7	0.03	85.2	86.9	0.05
Non-Hispanic black	6.9	6.2	-0.03	6.9	6.7	-0.01
Hispanic	4.1	3.7	-0.02	4.3	3.7	-0.03
Other	3.2	3.3	0.01	3.5	2.8	-0.04
Dual-eligibility for Medicaid (months per year)[e]	1.5	1.1	-0.12	1.4	1.4	-0.01
HCC Risk Score[f]	1.1	1.0	-0.22	1.1	1.1	0.01
Disability[g]	8.8	7.2	-0.06	8.5	8.6	0.01
End-stage renal disease[h]	0.7	0.3	-0.06	0.7	0.7	0.00
Area-level characteristics						
Below federal poverty level[i]	14.0	12.4	-0.19	13.8	13.6	-0.03

With high school degree[i]	86.4	87.8	0.16	86.5	86.8	0.04
With college degree[i]	28.0	31.2	0.19	28.6	28.3	-0.02
In Medicare Advantage[j]	22.0	21.7	-0.03	21.9	22.0	0.00
Patient outcomes						
Total spending per beneficiary-quarter ($)	2,514	1,725	-0.12	2,381	2,414	0.00
Hip fracture per 1000 beneficiary-quarters[k]	2.6	1.3	-0.03	2.4	2.4	0.00

Abbreviations: MSSP, Medicare Shared Savings Program; HCC, Hierarchical Condition Category.
[a] Descriptive characteristics were based on the pre-MSSP period (January 2008-March 2012) and were calculated for all beneficiaries appearing in at least one post-period quarter, as this was required for determining whether the attributed clinician was exposed to high vs. low MSSP supply (N= 56,669,720 beneficiary-quarters).
[b] For descriptive analyses, MSSP participants were defined as beneficiaries who ever participated in the MSSP over the study period; controls were defined as beneficiaries who never participated in the MSSP over the study period. For regression analyses of the effect of the MSSP on study outcomes, MSSP participation was defined using a time-varying indicator that equaled 1 if the beneficiary was attributed to the MSSP in a given quarter, 0 otherwise.

[c] High vs. low MSSP supply was defined by whether the attributed clinician's median MSSP value was above (1=high) or equal to or below (0=low) the median MSSP supply value for the HRR over the entire post-period. The beginning of the post-period was designated as the first quarter in which the HRR contained at least one clinician participating in the MSSP.
[d] Race/ethnicity was determined from Medicare enrollment files and based on the RTI race/ethnicity designation.
[e] Dual-eligibility for Medicaid was defined by the number of months in the year that the beneficiary's state of residence paid the monthly premium for Part B coverage (range: 0-12).
[f] HCC risk scores were calculated using Medicare demographic and diagnostic data from the prior year's enrollment and claims files. Higher HCC risk scores indicate higher predicted spending in the present year.
[g] Disability was defined by original reason for Medicare entitlement, regardless of current disability status.
[h] End-stage renal disease was defined by original reason for Medicare entitlement, regardless of current end-stage renal disease status.
[i] The proportion living below federal poverty level, with a high school degree, and with a college degree was defined at the zip-code tabulation area using American Community Survey data.
[j] The proportion of beneficiaries enrolled in Medicare Advantage was defined at the county-level using enrollment data for a 20% random sample of Medicare beneficiaries and was equal to 1 if the beneficiary was enrolled for $\geq$ 1 month during the year, 0 otherwise.
[k] Hip fracture was defined as an acute care hospital claim with a primary diagnosis of hip fracture (820.xx) in a given quarter.

Figure I.1. Changes in Medicare spending for beneficiaries attributed to MSSP ACOs vs. controls

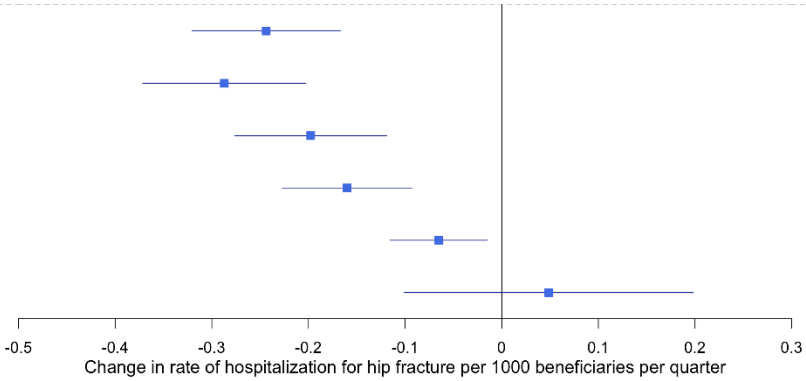


Abbreviations: MSSP, Medicare Shared Savings Program; ACO, accountable care organization CI, confidence interval.
Mean outcomes for control beneficiaries were calculated as the intercept coefficient from adjusted longitudinal models with mean-centered covariates. Mean outcomes for MSSP beneficiaries were calculated as the sum of the control mean and the estimated effect of the MSSP on the outcome. Total spending was the sum of Medicare spending for inpatient, outpatient, professional service, and skill nursing facilities. Component spending was defined by claims for services from the following research identifiable files: Medicare Provider Analysis and Review (inpatient and skilled nursing facility services); Carrier (professional services); and Outpatient Services (outpatient services).

Figure I.2. Changes in spending and hospitalization for hip fracture across models for fixed differences across MSSP participants vs. controls

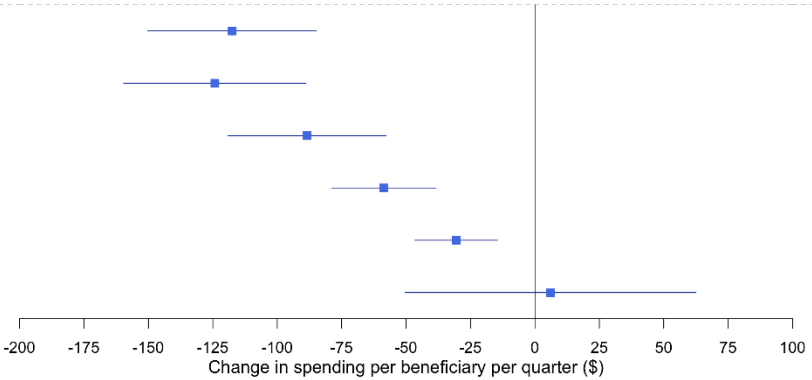
Panel A. Change in rate of hospitalization for hip fracture per 1000 beneficiary-quarters

Hip fracture	Mean Control	Mean MSSP	Change (95% CI)	P value
Adjusted longitudinal	2.12	1.88	-0.24 (-0.32, -0.17)	<0.001
Market-year fixed effects	2.12	1.84	-0.29 (-0.37, -0.20)	<0.001
Beneficiary fixed effects	2.12	1.93	-0.20 (-0.28, -0.12)	<0.001
Market-year + ACO fixed effects	2.12	1.96	-0.16 (-0.23, -0.09)	<0.001
Clinician fixed effects	2.12	2.06	-0.06 (-0.12, -0.01)	0.012
Instrumental variable	2.12	2.17	0.05 (-0.10, 0.20)	0.523



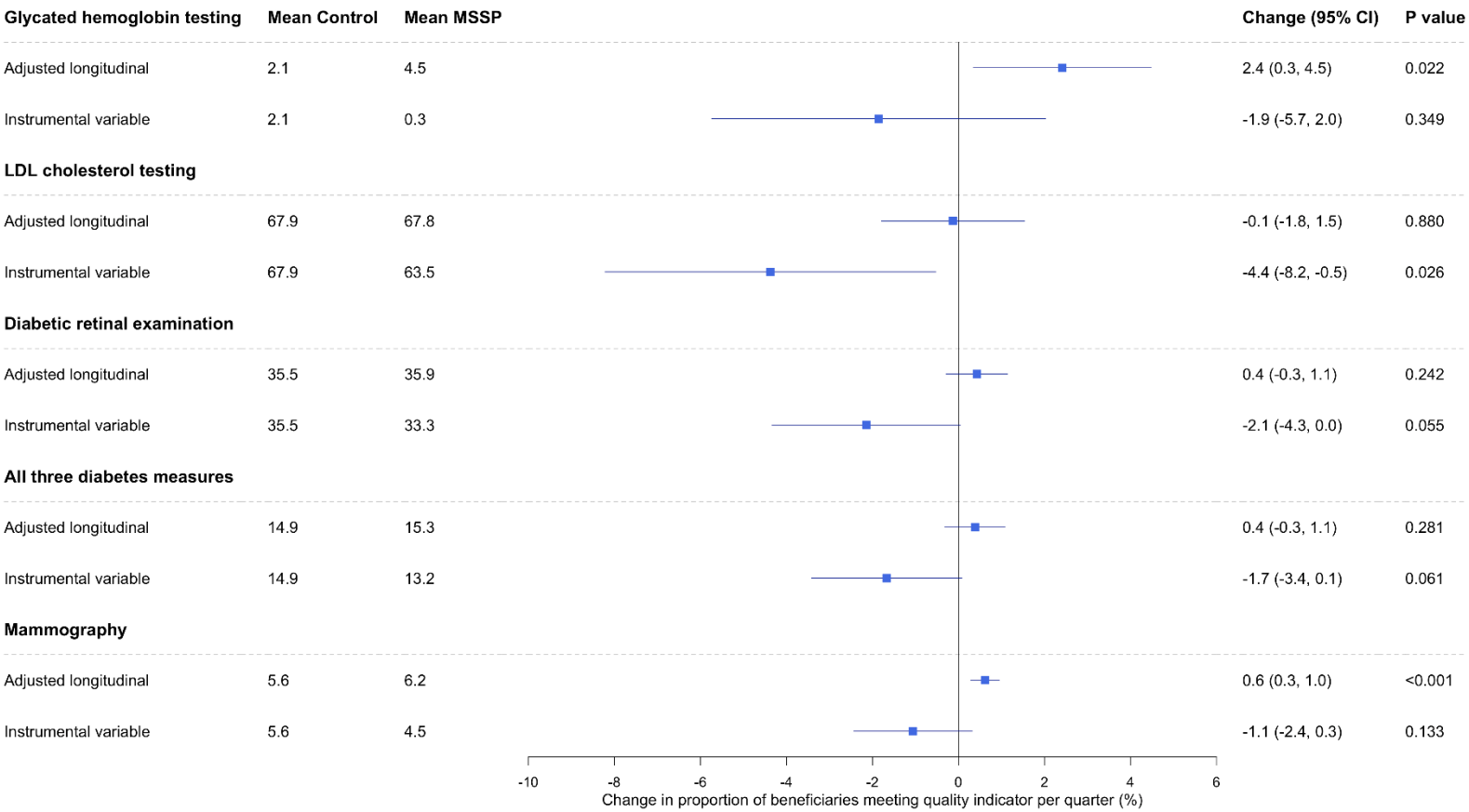
Panel B. Changes in total spending per beneficiary-quarter

Total spending	Mean Control	Mean MSSP	Change (95% CI)	P value
Adjusted longitudinal	2342	2224	-117 (-150, -85)	<0.001
Market-year fixed effects	2342	2218	-124 (-160, -89)	<0.001
Beneficiary fixed effects	2342	2253	-88 (-119, -58)	<0.001
Market-year + ACO fixed effects	2342	2283	-59 (-79, -38)	<0.001
Clinician fixed effects	2342	2311	-31 (-47, -14)	<0.001
Instrumental variable	2342	2348	6 (-50, 63)	0.834



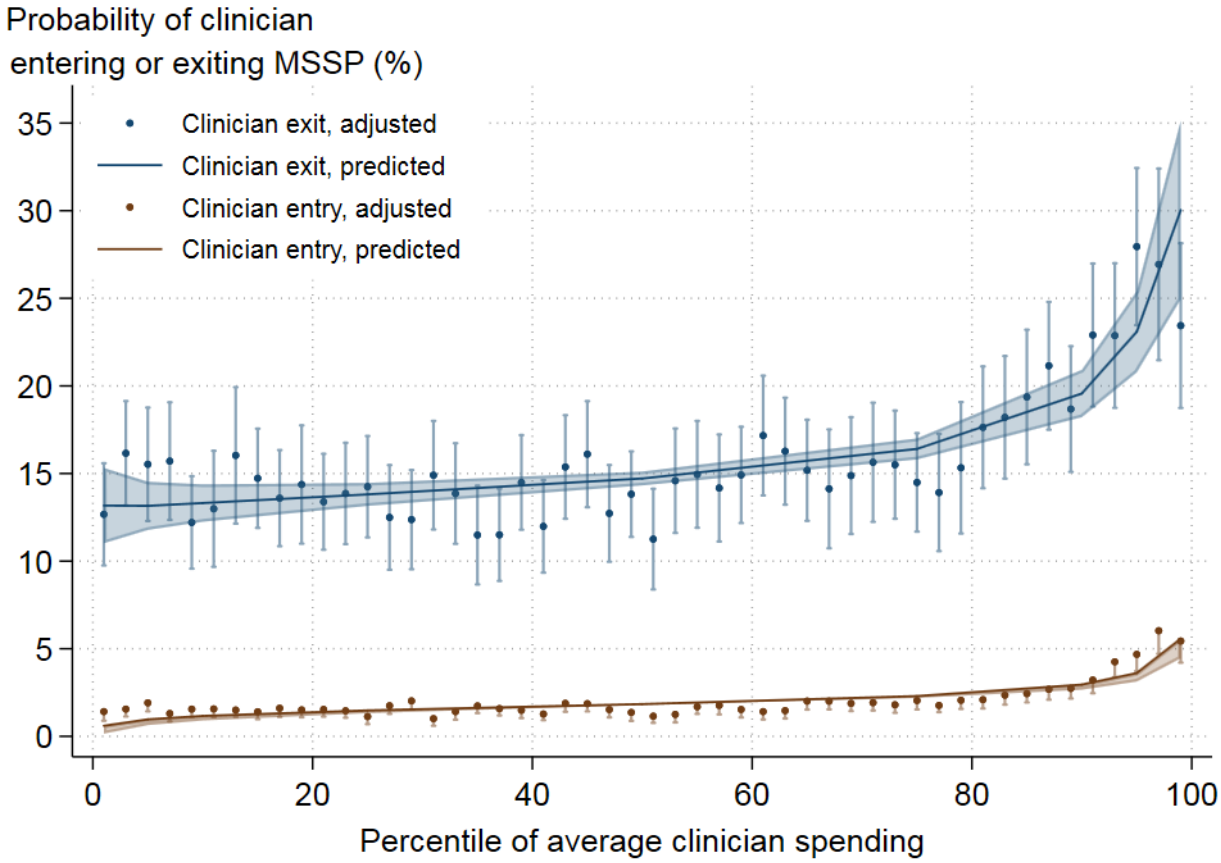
Abbreviations: MSSP, Medicare Shared Savings Program; CI, confidence interval.
Mean outcomes for control beneficiaries were calculated as the intercept coefficient from adjusted longitudinal models with mean-centered covariates. Mean outcomes for MSSP beneficiaries were calculated as the sum of the control mean and the estimated effect of the MSSP on the outcome. Hospitalization for hip fracture was defined as a Medicare acute hospital claim for a primary diagnosis of hip fracture (820.xx). Total spending was the sum of Medicare spending for inpatient, outpatient, professional services, and skill nursing facilities.

Figure I.3. Changes in clinical quality performance for beneficiaries attributed to MSSP ACOs vs. controls



Abbreviations: MSSP, Medicare Shared Savings Program; ACO, accountable care organization; CI, confidence interval.
The proportions of MSSP beneficiaries and controls receiving the preventive service are given in percent. Clinical quality indicators are defined in the Methods of Appendix A. Mean outcomes for control beneficiaries were calculated as the intercept coefficient from adjusted longitudinal models with mean-centered covariates. Mean outcomes for MSSP beneficiaries were calculated as the sum of control mean and the estimated effect of the MSSP on the outcome. Analyses of glycated hemoglobin testing, LDL cholesterol testing, and diabetic retinal examination were limited to beneficiaries with diabetes (N=6,779,900 beneficiary-quarters). Mammography analyses were limited to female beneficiaries with ages 65-69 (N=11,926,849 beneficiary-quarters).

Figure I.4. Association between clinician spending performance and probability of exiting or entering an MSSP ACO



Abbreviations: MSSP, Medicare Shared Savings Program; ACO, accountable care organization. Prior spending performance was defined as each clinician's average Medicare spending per beneficiary per year in the three years prior to MSSP exit or entry determination. The probability of a clinician entering or exiting the MSSP was estimated as a function of the clinician's average spending, average beneficiary characteristics, market fixed effects, and year fixed effects. A quadratic spending term was included to allow for any potential non-linearities in the effect of spending performance on MSSP participation. Analyses of MSSP exit (Panel A; N = 21,970 clinician-years) were restricted to clinicians participating in the MSSP during the year prior to analysis. Analyses of MSSP entry (Panel B; N = 162,336 clinician-years) were restricted to clinicians not participating in the MSSP during the year prior to analysis and did not include ACO formation, i.e., participation in an ACO's first contract year. Both sets of analyses were restricted to ACOs entering MSSP contracts in 2012 or 2013, as 2012-2014 MSSP data could not be used to determine clinician exit or entry for ACOs formed in 2014.

Chapter II. Changes in coded severity and risk selection in the Medicare Shared Savings Program

Abstract

To encourage Medicare Shared Savings Program (MSSP) Accountable Care Organizations (ACOs) to care for high-risk beneficiaries, savings benchmarks are adjusted by beneficiaries' baseline risk scores. To discourage increased coding intensity, benchmarks are not adjusted upward if beneficiary risk score rises while in the MSSP. To evaluate the impact of this risk adjustment policy, we examined whether beneficiary exposure to the MSSP was associated with within-beneficiary changes in risk score and whether risk score was associated with beneficiary entry and exit in the MSSP. We found that the MSSP was not associated with consistent changes in within-beneficiary risk score. Conversely, the highest-risk beneficiaries (99th percentile of risk score) had a 25.1% chance of exiting the MSSP compared to a 16.0% chance among median risk beneficiaries (50th percentile). CMS' decision to not upwardly adjust risk score in the MSSP has successfully deterred coding increases but may have led ACOs to avoid high-risk beneficiaries.

Introduction

Encouraging organizations to care for high-risk beneficiaries while holding them accountable for spending and health outcomes is a central tension of payment reform.[30] In the Medicare Shared Savings Program (MSSP), Accountable Care Organizations (ACOs) are eligible to receive shared savings bonuses if they lower spending below a financial benchmark based on the historical spending of the beneficiaries attributed to the ACO. To avoid penalizing ACOs that care for high-risk beneficiaries (i.e., greater medical complexity and predicted spending), an ACO's financial benchmark is adjusted using each beneficiary's Hierarchical Condition Category (HCC) risk score. To minimize ACOs' incentives to raise benchmarks by increased diagnostic coding, the benchmark is not adjusted upward if the risk score rises while the beneficiary is in the MSSP. [33–35] If the risk score falls, however, the benchmark is adjusted downward.

It is unknown if the Centers for Medicare and Medicaid Services (CMS) approach to risk adjustment has appropriately balanced incentives for ACOs to care for high-risk beneficiaries against not rewarding ACOs for increasing coding intensity in the MSSP. Because CMS' approach does not capture growth in risk score over time, many commenters expressed concern during rulemaking that ACOs still have an incentive to avoid high-risk beneficiaries with increasing comorbidities or acute illness (i.e., "favorable selection").[33–35] For instance, ACOs may deliberately drop clinicians with high-risk beneficiary panels, excluding clinicians from the ACO Participant List submitted to CMS each year.[25] ACOs may also prevent high-risk beneficiaries from

being attributed to their ACO. Although ACOs typically must provide care to beneficiaries with Medicare, ACOs could exploit the MSSP's retrospective attribution methodology by: 1) submitting claims that cannot lead to attribution; 2) submitting claims from a provider ineligible for the MSSP (e.g., urologists) and thus excluded from attribution; or 3) billing under a provider group not included in the ACO Participant List. At the same time, because this risk adjustment system also penalizes ACOs that allow beneficiaries' risk scores and benchmarks to fall, ACOs also face an incentive to remain vigilant in maintaining their current levels of coding intensity.

To evaluate the impact of this risk adjustment policy, we used national Medicare data from 2008 through 2014 to examine the relationship between the MSSP and beneficiary risk profiles. We assessed changes in coding intensity by evaluating whether beneficiary exposure to the MSSP was associated with within-beneficiary changes in risk score over time. We assessed favorable selection by evaluating whether beneficiary risk score and risk growth was associated with entry and exit of beneficiaries and clinicians to and from the MSSP.

Methods

Data Sources and Study Population. We used national claims data from 2008 through 2014 for a random 20 percent sample of beneficiaries in Fee-for-Service (FFS) Medicare. This included the following Research Identifiable Files: Medicare Provider Analysis and Review (inpatient and skilled nursing facility services); Carrier (physician

services); Outpatient (facility services); and Master Beneficiary Summary File (beneficiary enrollment and sociodemographic information). To ensure accurate designation of MSSP participation, we used CMS' Provider- and Beneficiary-level Shared Savings Program files to attribute beneficiaries and clinicians to MSSP ACOs (2012-2014). We then linked ACOs to Leavitt Partners ACO Database and CMS' Shared Savings Program Public-Use File to obtain data on ACOs' organizational characteristics and performance in the MSSP.

We constructed a cohort of Medicare beneficiaries who were continuously enrolled in FFS Medicare to examine longitudinal relationships between the MSSP beneficiary risk profiles. Following MSSP specifications, we excluded beneficiaries who were enrolled in Medicare Advantage or not enrolled in Medicare Parts A and B.[57] In order to ensure consistent identification of within-beneficiary risk score changes over time, we also excluded beneficiaries who died or had missing risk scores in any year of the study (2008-2014). This also ensured that MSSP exit was not an artifact of beneficiary death. To improve comparability between MSSP beneficiaries and controls, we also excluded beneficiaries who resided outside of a hospital referral region or lacked any eligible primary care claims required for MSSP attribution in any year of the study (CONSORT diagram, Figure B1 in Appendix B).

Measures. CMS derives a single HCC risk score for each Medicare beneficiary to predict spending in the subsequent year.[76] This algorithm is a prospective model that incorporates both diagnostic information in claims data (subsequently classified into

HCCs) and demographic information (including age, sex, Medicaid dual eligibility, and disability). For the purpose of risk adjustment, CMS calculates the HCC risk score based on patient and clinical factors in the prior year. In this study, we calculate the HCC risk score based on patient and clinical factors in the concurrent year.

We defined exposure to the MSSP using a time-varying indicator for cumulative time attributed to an MSSP ACO. This captured ACOs' staggered entry into MSSP contracts (April 2012, July 2012, January 2013, January 2014) and the ability of beneficiaries to enter and exit MSSP ACOs each year.

We used the Master Beneficiary Summary File to define beneficiary age, sex, race/ethnicity (white, black, Hispanic, or other/unknown), original reason for Medicare entitlement (aged, disability, end-stage renal disease), dual-eligibility for Medicaid (months of enrollment), and average county-level Medicare Advantage penetration (share of Medicare beneficiaries). Due to consistent reports that risk scores have increased more among beneficiaries in Medicare Advantage than in FFS Medicare,[77–80] adjusting for Medicare Advantage penetration addressed potential confounding between risk score and local changes in MSSP and Medicare Advantage penetration. We used American Community Survey data to define beneficiary area-level poverty (proportion below federal poverty level) and education (proportion graduated from high school, college). We used the Shared Savings Program Public Use File to define MSSP earned shared savings (yes/no) and the Leavitt database to define ACO organizational

structure (physician-led, hospital-led, or physician-hospital partnership) and concurrent entry in a commercial ACO contract (yes/no).

Analyses. We first examined the relationship between the MSSP and changes in coded risk score. We performed a longitudinal, beneficiary-level analysis of the effect of MSSP exposure on within-beneficiary changes in risk score over time. We estimated linear regression models that included a time-varying indicator of cumulative MSSP exposure, beneficiary fixed effects, year fixed effects (to control for secular trends), and the time-varying characteristics. By using beneficiaries as their own controls, these models controlled for fixed unobserved differences across beneficiaries that may confound the relationship between MSSP exposure and risk score. We performed two sensitivity analyses of the relationship between MSSP exposure and coded risk. First, to better isolate the influence of coding on risk score, we also estimated changes in the component of the risk score most plausibly affected by coding practice (i.e., the count of Condition Categories that originate from provider-reported diagnoses) and excluded the components originating from administrative data (e.g., age, sex, disability status). Second, because ACOs also have an incentive to lower spending in the MSSP, we also evaluated changes in risk scores normalized by total price-standardized spending.

We next assessed whether MSSP ACOs engage in favorable selection, e.g., avoiding high-risk beneficiaries or clinicians with high-risk patient panels. First, we examined the relationship between beneficiary risk score and subsequent entry or exit of the beneficiary to or from the MSSP. In these analyses, we estimated the probability of

beneficiary entry and exit as a function of prior-year risk score, market fixed effects (defined by hospital referral region), year fixed effects, beneficiary age, sex, race/ethnicity, disability, end-stage renal disease, dual-eligibility for Medicaid, and area-level poverty, education, and Medicare Advantage penetration. Robustness tests included assessing entry and exit: 1) using beneficiary average risk score in the prior two or three years; 2) among only those beneficiaries attributed to MSSP ACOs vs. controls using only eligible claims submitted by primary care clinicians (vs. specialists) in the outpatient setting (vs. post-acute care facilities); and 3) across ACO and beneficiary characteristics. In sensitivity analyses exploring the mechanisms for entry and exit, we estimated models of clinician entry or exit from the MSSP as a function of the clinician's patient panel's average risk score, other patient characteristics (listed above), year fixed effects, and market fixed effects (see Methods in Appendix B).

Finally, we investigated the impact of CMS' decision to not upwardly adjust spending benchmarks by examining whether growth in risk scores varied across MSSP status. We compared growth in risk score across beneficiaries who were never in the MSSP, always in the MSSP (through 2014), entered the MSSP (in 2014), or exited the MSSP (in 2014). This classification scheme followed a recent report by the Medicare Payment Advisory Commission.[81] Differences in risk score growth were examined using a linear spline model that included MSSP status (always, never, enter, or exit), splines for the two periods when entry and exit could either not occur (2012-2013) or could occur (2013-2014), and an interaction between MSSP status and the two splines. Models also

included market fixed effects, year fixed effects, and the time-varying beneficiary characteristics described above.

All analyses specified robust standard errors to account for clustering at the market level. Statistical analyses were performed using Stata version 15.1. Our study using administrative data was deemed exempt from review by the University of Michigan Institutional Review Board.

Limitations. Our study has several limitations. First, administrative data cannot be used to determine whether risk score changes reflect true changes to health status, health care utilization, or coding practices. It is possible that ACOs have both improved health and lowered utilization (lowering risk score) and increased coding intensity (raising risk score), with uncertain net effects on average risk score. Nonetheless, results were comparable using risk scores normalized by price-standardized spending (Figure B2 in Appendix B). Second, it is difficult to determine whether MSSP exit reflects passive "leakage" of high-risk beneficiaries to non-ACO specialists or strategic behavior by ACOs intending to drop clinicians who care for high-risk beneficiaries. However, passive leakage to non-ACO specialists is unlikely to explain our finding that beneficiary risk was disproportionately associated with exit among beneficiaries attributed via outpatient claims submitted by primary care clinicians (Figure B2 in Appendix B), nor would it explain our finding that exit of clinicians with high-risk panels partially drove this exit (Figure B3 in Appendix B).

Third, although our analysis of risk score changes controlled for fixed differences across MSSP beneficiaries and controls, time-varying confounding is a threat. For instance, ACOs dropping high-risk beneficiaries would create a negative association between beneficiary risk score and MSSP exposure (e.g., reverse causality), inducing a downward bias and underestimation of the true effect of MSSP exposure on risk score. Our use of cumulative (vs. current) MSSP exposure minimizes but may not eliminate this bias. Fourth, historical data ending in 2014 may not generalize to today's ACOs. It is possible that the relationship between beneficiary risk profiles and MSSP ACOs have changed in the intervening years. Our results may also not generalize to younger beneficiaries or those that move in and out of FFS Medicare and the health system more generally.

Results

Beneficiary MSSP attribution and risk profiles. FFS Medicare beneficiaries in our sample contributed 13,864,627 beneficiary-years from 2008 through 2014 (n=1,980,661 beneficiaries). By the end of 2014, 21.4% of beneficiaries (n=425,353) had been attributed at some point to the MSSP, with MSSP beneficiaries exposed to the program for an average of 1.65 years (range: 1 to 2.75 years).

Unadjusted differences in beneficiaries' clinical and sociodemographic baseline characteristics between MSSP beneficiaries and controls were typically small, though MSSP beneficiaries resided in areas with higher education and lower poverty (Table

II.1). The average risk score in the pre-period (2008-2011) was 1.06 (standard deviation, 0.81). Pre-period trends in adjusted risk score were similar across MSSP beneficiaries and controls (Figures B5 and B6 in Appendix B). Because we required enrollment in FFS Medicare for each year of the study, beneficiaries in our sample were more likely to have become entitled to Medicare due to old age (vs. ESRD or disability) and were thus older, on average, than excluded beneficiaries (Table B1 in Appendix B).

Change in within-beneficiary risk score. The MSSP was not associated with change in average risk score in the pooled analysis (percent change in risk score, 0.0%, 95% confidence interval [CI], -0.3% to 0.3%) (Figure II.1). However, this relationship varied significantly across MSSP cohorts. The MSSP was not associated with risk score change among beneficiaries entering the MSSP in 2012 (change, -0.2%, 95% CI, -0.6% to 0.2%) or 2013 (change, 0.1%, 95% CI, -0.3% to 0.5%) but was associated with a 1.1% increase in risk score (95% CI, 0.4% to 1.7%) among beneficiaries entering the MSSP in 2014. The relationship between the MSSP and risk score did not vary across whether the ACO earned shared savings in the MSSP, whether the ACO held a concurrent commercial ACO contract, or the ACOs' organizational structure. We observed variation in risk score according to Medicare Advantage penetration, with the MSSP associated with a 0.6% increase in risk score (95% CI, 0.1% to 1.1%) among beneficiaries residing in areas with high penetration by Medicare Advantage plans vs. those that did not (change, -0.1%, 95% CI, -0.4% to 0.2%).

Overall, we observed a similar pattern of results in sensitivity analyses which used the count of Condition Categories and risk scores normalized by price-adjusted spending as the study outcomes (Figure B3 in Appendix B).

Beneficiary exit and entry in the MSSP. Beneficiary exit and entry in the MSSP was common (Table II.1). Of the 245,239 beneficiaries initially attributed to ACOs in 2012 or 2013, 40,428 (16.4%) exited the MSSP in 2014. Of the 278,672 beneficiaries attributed to ACOs in 2014, 72,861 (26.1%) entered the MSSP in 2014.

Figure II.2 shows the relationship between beneficiary risk score and exit or entry from the program. Beneficiaries at the 95th and 99th percentile of risk score in the prior year had a 21.6% and 25.0% chance of exiting an MSSP ACO in the subsequent year, as compared to a 16.0% chance among beneficiaries at the median (50th percentile) risk score (risk difference, 5.7 percentage point [p.p.] and 9.1 p.p., respectively, both P<.001; Table B2 in Appendix B). Entry into already-formed ACOs was also modestly higher among beneficiaries in the 99th percentile of risk score as compared to the beneficiaries at the median (risk difference, 1.0 p.p., P<.001; Figure II.2 and Table B2). The disproportionate exit by high-risk beneficiaries from the MSSP was consistent across a range of subgroups and when using beneficiaries' average risk score in the prior two or three years (Figures B7 and B8 in Appendix B), suggesting that exit was specifically associated with beneficiary health status vs. other beneficiary or provider characteristics. We observed a similar pattern of MSSP exit and entry among beneficiaries attributed via outpatient claims submitted by primary care clinicians

(Figures B3 in Appendix B), suggesting that exit was not driven by passive leakage of high-risk beneficiaries to non-ACO specialists.

Clinician exit appeared to drive exit of high-risk beneficiaries from the program (Figure B3 in Appendix B). Clinicians at the 95th and 99th percentile of average panel risk score had a 20.6% and 26.3% chance of exiting an MSSP ACO, as compared to a 16.0% chance among clinicians at the median panel risk score (risk difference, 4.5 p.p. and 10.3 p.p., respectively, both P<.001).

Change in MSSP risk profile composition. Figure II.3 demonstrates risk score growth among beneficiaries who were always in the MSSP (through 2014), never in the MSSP, entered the MSSP (in 2014), or exited the MSSP (in 2014). Among the four groups, beneficiaries who ultimately exited the MSSP demonstrated the highest risk score growth in the periods prior to exit (2012-2013) and following exit (2013-2014). Compared to beneficiaries who were always in the MSSP, beneficiaries who exited the program had risk score growth that was 4.8 p.p. higher prior to exit (95% CI, 3.8 to 5.9) and 3.1 p.p. higher following exit (95% CI, 1.9 to 4.3; Table B3 in Appendix B). Beneficiaries who entered the MSSP also demonstrated risk score growth that was 2.2 p.p. higher prior to entry (95% CI, 1.6 to 2.8) and 1.6 p.p. higher following entry (95% CI, 0.7 to 2.4), as compared to beneficiaries always in the MSSP. When directly comparing exiting vs. entering beneficiaries, those who exited the MSSP had risk score growth that 2.7 p.p. (95% CI, 1.5 to 3.9) in 2012-2013 and 1.6 p.p. higher (95% CI, 0.0 to 3.1) in 2013-2014. These risk growth differences between beneficiaries exiting and

entering the MSSP are large, representing a 71% increase relative to overall risk score growth in 2012-2013 (~ difference of 2.7 p.p. / 3.7 p.p. overall risk score growth in the sample) and 27% increase in 2013-2014 (~1.6 p.p. / 5.9 p.p.).

Discussion

In this national cohort study of Medicare beneficiaries, we found limited evidence that exposure to the MSSP increased within-beneficiary risk score. At the same time, we found that high-risk beneficiaries and clinicians with higher-risk patient panels were disproportionately likely to exit MSSP ACOs. MSSP exit was particularly concentrated among beneficiaries with increased risk score growth in the MSSP and following exit from the program. These findings suggest that the current system of risk adjustment in the MSSP has successfully deterred coding increases but may not adequately encourage ACOs to care for high-risk beneficiaries in the MSSP.

To our knowledge, this is the first comprehensive analysis of within-beneficiary changes in coded risk in the MSSP. Much more attention has been paid to coded risk in Medicare Advantage, where use of HCC risk scores to adjust capitated payments has been associated with substantial increases in coded risk and billions of taxpayer dollars in potential overpayment to Medicare Advantage plans.[77–80] There is also precedent in FFS Medicare, where CMS efforts to link global payment reform with appropriate risk adjustment have been associated with increased diagnostic coding following Medicare's introduction of the Inpatient Prospective Payment System, Hospital Readmissions

Reduction Program, and Physician Group Practice demonstration, an early model for ACOs.[15,31,32]

Several features of the MSSP may have contributed to the lack of an observed change in coded risk. Risk coding in the MSSP may be less susceptible to coding practices. ACOs cannot submit supplemental diagnostic codes, while Medicare Advantage plans conduct health risk assessments in enrollee's homes and perform retrospective reviews of medical charts.[79] Retrospective attribution of beneficiaries to MSSP ACOs may hamper organizational efforts to target coding efforts (vs. prospective Medicare Advantage attribution or clearly defined hospital episodes).[35] The high rates of beneficiary exit and entry we uncovered in the MSSP may also impede efforts to systematically intensify coding across a static beneficiary population.

MSSP ACOs may also face weakened incentives to intensify coding. Discouraging coding intensification is the explicit motivation underlying CMS' decision to not upwardly adjust risk scores of beneficiaries remaining in MSSP ACOs.[33–35] Meanwhile, the spread of risk-based contracts through public and private insurance has raised the salience of coding for all providers,[30] perhaps reducing the marginal incentive to increase coding for MSSP beneficiaries relative to those in standard FFS Medicare. Our finding that coded risk increased among MSSP beneficiaries residing in areas with high Medicare Advantage penetration suggests that MSSP incentives to increase coding may depend on organizations' coding expertise and ability to spread fixed costs of coding initiatives.

ACOs and other stakeholders expressed concern during final rulemaking that CMS' failure to upwardly adjust beneficiary risk may cause ACOs to avoid high-risk beneficiaries in the MSSP. [33–35] However, ours is the first study to empirically evaluate this question. We found that high-risk beneficiaries and clinicians with high-risk patient panels were more likely to exit the MSSP. Our results are consistent with a recent study by Hsu and colleagues, who found that clinicians with high-risk patient panels were more likely to exit Partners HealthCare Pioneer ACO.[25]

There are several reasons why chronically- or acutely-ill beneficiaries may disproportionately exit MSSP ACOs. First, MSSP beneficiaries with complex needs may receive care from specialists outside the ACO, causing them be passively attributed ("leakage") to non-MSSP organizations the following year. However, this is unlikely given MSSP's attribution methodology, which only incorporates specialist claims when there are no eligible claims from primary care clinicians. Moreover, our finding that beneficiary risk was associated with MSSP exit among beneficiaries attributed via outpatient claims from primary care clinicians (Figure B3 in Appendix B) and among both hospital- and physician-led ACOs (Figure B7 in Appendix B) suggests that exit by high-risk beneficiaries was unrelated to ACOs' ability to provide specialty care.

An alternative explanation is that ACOs are strategically dropping high-risk beneficiaries or clinicians with high-risk patient panels to decrease measured spending and increase the likelihood of earning shared savings. MSSP ACOs' ability to contract with clinicians each year may enable ACOs to drop clinicians with higher-risk patient panels.[25,33–35]

ACOs may also prevent high-risk beneficiaries from being attributed to their ACO, exploiting the MSSP's retrospective attribution methodology by either submitting claims not included in MSSP attribution rules (because the claims are ineligible or because the claims are submitted by an ineligible provider) or by billing under a provider group not included in its ACO Participant List (e.g., skilled nursing facility). Because the MSSP does not upwardly adjust risk scores following MSSP attribution, beneficiaries with worsening health status are particularly detrimental to ACOs' savings benchmarks. Our finding that risk score growth was highest among exiting beneficiaries (both prior to and following exit) raises this possibility. Regardless of the mechanism, exit from the MSSP threatens policy efforts to improve the efficiency and quality of care for high-risk beneficiaries.

Policy Implications. The tension between appropriate risk adjustment and inappropriate coding practices, already evident during early MSSP rulemaking, will increase rapidly as the MSSP transitions to regional benchmarks and two-sided risk contracts. In its 2016 Final Rule, CMS announced that it will continue to not upwardly adjust risk scores even as it implements regional benchmarks incorporating average spending of nearby providers with potentially very different patient compositions.[35]

There are several ways CMS could reduce MSSP incentives and ACOs' ability to avoid beneficiaries with chronic or acute illness. The most obvious solution is to upwardly adjust risk score. This successfully reduced favorable selection in Medicare Advantage[51] but also considerably intensified risk coding.[77–80] A second approach is

allowing for upward risk adjustment while capping risk score changes at plus or minus three percent, as in the Next Generation ACO model. However, this may not alter selection incentives if and when the majority of ACOs reach this threshold. A third approach is incorporating other sociodemographic factors that predict spending but are less subject to coding practices, e.g., race, education, area-level poverty.[52,82,83] However, our study suggests that MSSP exit is more frequent among acutely-ill beneficiaries whose changing health status may not be captured by stable sociodemographic characteristics. To encourage ACOs to care for these beneficiaries, the MSSP could adjust for pre-existing "risk velocities" (e.g., growth in patient risk score) in the years prior to MSSP attribution.

CMS could couple efforts to enhance risk adjustment with aggressive auditing of compositional change and coding practices in the MSSP. Although CMS has repeatedly stated in each MSSP Final Rule that it plans to do this,[33–35] whether or how this has occurred remains unknown. The limits of auditing are demonstrated by the Medicare Advantage experience, where the focus has been on detecting outright fraud rather than on whether the transfer of taxpayer funds to Medicare Advantage plans appropriately reflects true differences in health across beneficiaries.[77] Finally, CMS could make it more difficult for ACOs to drop high-risk beneficiaries. Prospective attribution – used in the Next Generation and MSSP Track 3 models – may help prevent ACOs from dropping beneficiaries during the performance year and prior to attribution.

Conclusion. This study contributes important evidence to ongoing discussions about the role of risk adjustment in alternative payment models. Our results suggest that the current risk adjustment system has successfully deterred coding increases but may not adequately encourage ACOs to care for high-risk beneficiaries in the MSSP. Even with significant reform, it may not be possible to achieve a perfect balance between adequate risk adjustment and favorable selection in the MSSP. As the MSSP and other Medicare ACOs expand, it will be important to monitor the evolution of ACO risk profiles and to weigh the dual policy goals of preventing coding increases against caring for high-risk beneficiaries in the program.

Table II.1. Characteristics of beneficiaries in Medicare Shared Savings Program (MSSP) Accountable Care Organizations (ACOs) vs. controls in 2011

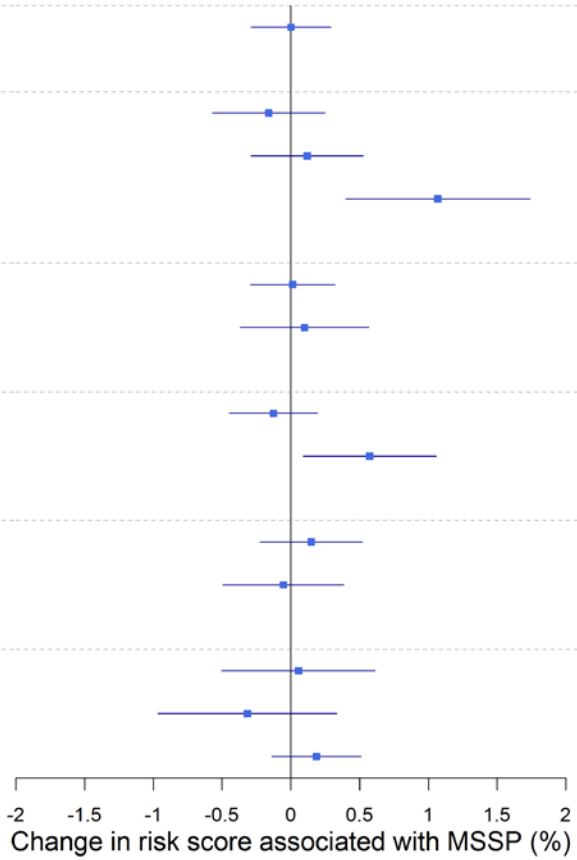
	Covariate balance across all MSSP beneficiaries vs. controls[a]		Covariate balance across beneficiaries who enter or exit the MSSP[b]		
Characteristics	**All controls (N=1,555,308)**	**All MSSP beneficiaries (N=425,353)**	**Always in MSSP (N=205,811)**	**Enter MSSP (N=72,861)**	**Exit MSSP (N=40,428)**
Beneficiary characteristics	**Unadjusted mean**	**Unadjusted mean**	**Unadjusted mean**	**Unadjusted mean**	**Unadjusted mean**
Age, y (SD)	74.2 (10.5)	74.6 (10.2)	74.7 (9.9)	74.5 (10.6)	74.4 (11.1)
Female	60.7%	60.8%	60.7%	61.2%	61.2%
Race/Ethnicity[c]					
Non-Hispanic white	86.1%	86.4%	87.4%	86.4%	86.2%
Non-Hispanic black	7.2%	6.9%	6.3%	6.4%	7.1%
Hispanic	3.7%	3.7%	3.2%	4.2%	4.1%
Other	3.0%	2.9%	3.1%	3.0%	2.7%
Dual-eligibility for Medicaid (months per year)[d]	1.8 (4.3)	1.6 (4.1)	1.6 (4.0)	1.7 (4.2)	1.8 (4.3)
Disability[e]	20.2%	18.1%	17.4%	19.2%	20.2%
End-stage renal disease[f]	0.4%	0.4%	0.4%	0.6%	0.5%
Area-level characteristics					
Medicare Advantage[g]	22.3 (12.7)	22.5 (12.1)	22.8 (11.9)	21.8 (12.7)	22.3 (11.2)
Below federal poverty level[h]	14.0 (8.5)	12.6 (8.3)	12.5 (8.2)	12.6 (8.3)	12.5 (8.3)
With high school degree[h]	27.7 (15.8)	30.6 (16.1)	30.4 (15.8)	30.8 (16.5)	30.9 (16.2)
With college degree[h]	86.3 (8.5)	87.5 (8.2)	87.7 (8.1)	87.6 (8.3)	87.6 (8.2)

Beneficiary outcomes					
HCC risk score[i]	1.160 (0.869)	1.154 (0.865)	1.131 (0.844)	1.202 (0.900)	1.243 (0.939)
Total annual spending[j]	7,485 (14,452)	7,420 (14,016)	7,133 (13,542)	8,017 (14,810)	8,437 (16,327)

[a] MSSP beneficiaries were defined as beneficiaries who were ever attributed to the MSSP between 2012 and 2014. Controls were defined as beneficiaries who were never attributed to the MSSP between 2012 and 2014.
[b] Comparisons of beneficiaries who were always in the MSSP (through 2014), entered the MSSP (in 2014) or exited the MSSP (in 2014) were restricted to ACOs that entered MSSP contracts in 2012 or 2013. We excluded beneficiaries who formed ACOs that entered MSSP contracts in 2014, as we could not observe subsequent exit or entry in ACOs formed in 2014 using 2008-2014 data.
[c] Race/ethnicity was determined from Medicare enrollment files and based on the RTI race/ethnicity designation.
[d] Dual-eligibility for Medicaid was defined by the number of months in the year that the beneficiary's state of residence paid the monthly premium for Part B coverage (range: 0-12).
[e] Disability was defined by original reason for Medicare entitlement, regardless of current disability status.
[f] End-stage renal disease was defined by original reason for Medicare entitlement, regardless of current end-stage renal disease status.
[g] The proportion of beneficiaries enrolled in Medicare Advantage was defined at the county-level using enrollment data for a 20% random sample of Medicare beneficiaries and was equal to 1 if the beneficiary was enrolled for $\geq$ 1 month during the year, 0 otherwise.
[h] The proportion living below federal poverty level, with a high school degree, and with a college degree was defined at the zip-code tabulation area using American Community Survey data.
[i] HCC risk scores were calculated using Medicare demographic and diagnostic data from the prior year's enrollment and claims files. Higher HCC risk scores indicate higher predicted spending in the present year.
[j] Total annual spending was the sum of spending for inpatient, outpatient, professional, and skilled nursing facility services and was price-standardized to account for variation resulting from regional wage indices and payments for indirect medical education, Disproportionate Share Hospitals, and new technologies.

Figure II.1. Association between beneficiary attribution to the Medicare Shared Savings Program and change in risk score

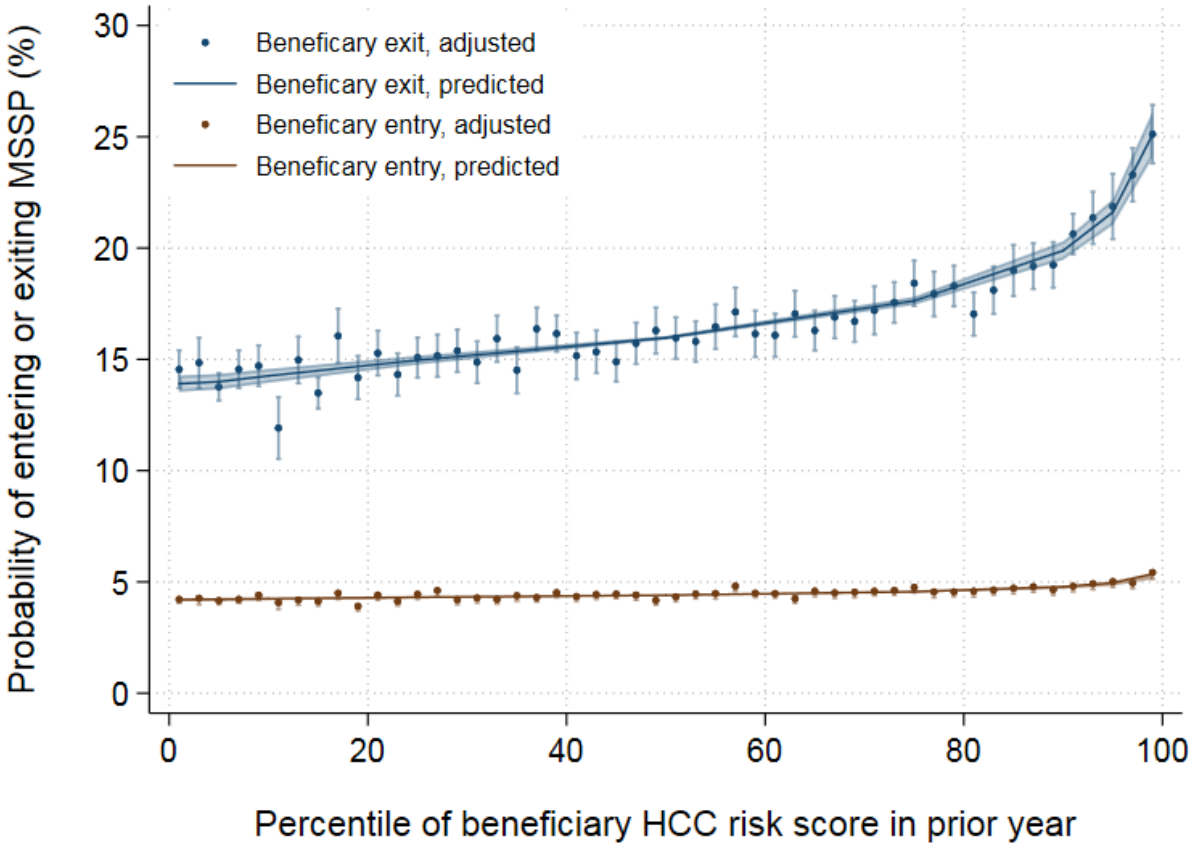
Characteristic	Mean Controls (2011)	Mean MSSP (2011)	Change in risk score (95% CI)	Change relative to base (95% CI)
Average change	1.159	1.159	0.0 (-0.3, 0.3)	N/A
MSSP Cohort				
2012	1.168	1.166	-0.2 (-0.6, 0.2)	--
2013	1.143	1.144	0.1 (-0.3, 0.5)	0.2 (-0.4, 0.8)
2014	1.151	1.164	1.1 (0.4, 1.7)	1.5 (0.8, 2.3)
Earned shared savings				
No	1.148	1.148	0.0 (-0.3, 0.3)	--
Yes	1.181	1.182	0.1 (-0.4, 0.6)	0.0 (-0.5, 0.5)
High Medicare Advantage area				
No	1.159	1.158	-0.1 (-0.4, 0.2)	--
Yes	1.158	1.165	0.6 (0.1, 1.1)	0.8 (0.3, 1.3)
Commercial ACO contract				
No	1.164	1.165	0.1 (-0.2, 0.5)	--
Yes	1.153	1.153	-0.1 (-0.5, 0.4)	-0.3 (-0.9, 0.3)
Organizational structure				
Physician-led	1.171	1.172	0.1 (-0.5, 0.6)	--
Hospital-led	1.151	1.147	-0.3 (-1.0, 0.3)	-0.5 (-1.3, 0.3)
Physician-hospital partnership	1.153	1.155	0.2 (-0.1, 0.5)	0.1 (-0.6, 0.7)



These models used risk score values derived from claims in the same year to capture the relationship between MSSP exposure and contemporaneous changes in risk score. Models included time-varying area-level factors (poverty,

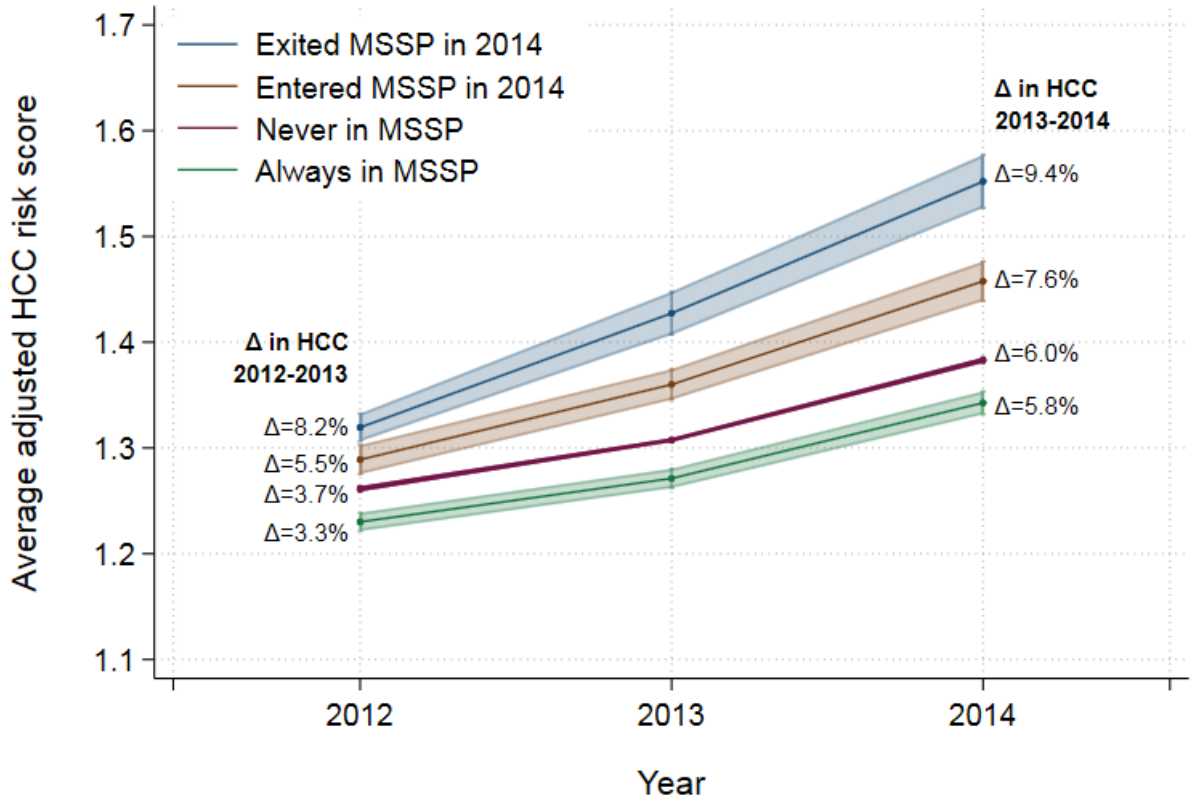
education, Medicare Advantage penetration) and omitted time-varying demographic variables factors into the risk score algorithm (age, sex, Medicaid dual eligibility, disability) to avoid over-adjustment. We measured change in risk scores relative to the base category in the following manner: 1) estimated an interaction model that included an interaction term between the MSSP indicator and the ACO or beneficiary characteristic to the base regression model (explained in the text); 2) estimated risk score changes for each group; 3) formally tested for differential risk score changes across groups. MSSP cohort was defined as the year in which the beneficiary was first attributed to an MSSP ACO. High Medicare Advantage penetration was defined as residing in a county > 80th percentile for the share of fee-for-service beneficiaries enrolled in Medicare Advantage. Percent change in HCC risk score was measured as the estimated change in risk score relative to the average adjust risk score for MSSP beneficiaries in 2011 (prior to the program's start). The error bars indicate 95% confidence intervals. MSSP is Medicare Shared Saving Program. ACO is accountable care organization. HCC is Hierarchical Condition Category.

Figure II.2. Association between risk score and change in probability of beneficiary entering or exiting the Medicare Shared Savings Program (MSSP)



Analyses of MSSP entry were restricted to beneficiaries not attributed to the MSSP in the year prior to analysis and specifically excluded beneficiaries who formed ACOs (i.e., entering in the ACO's first contract year). Analyses of MSSP exit were restricted to beneficiaries attributed to the MSSP in the year prior to analysis. Both sets of analyses excluded beneficiaries who formed ACOs that entered MSSP contracts in 2014, as we could not observe subsequent exit or entry in ACOs formed in 2014 using 2008-2014 data. The probability of entering or exiting was estimated as a linear function of risk score in the prior year, market fixed effects, year fixed effects, time-varying beneficiary characteristics, and a quadratic risk score term (to allow for potential non-linearities). To assess model fit, we also graphed beneficiaries' adjusted probability of entry or exit across the distribution of risk score without assuming any specific functional form.

Figure II.3. Growth in risk score and beneficiary entry and exit to and from the Medicare Shared Savings Program from 2012-2014



Comparisons of beneficiaries who were always in the MSSP (through 2014), entered the MSSP (in 2014) or exited the MSSP (in 2014) were restricted to ACOs that entered MSSP contracts in 2012 or 2013. We excluded beneficiaries who formed ACOs that entered MSSP contracts in 2014, as we could not observe subsequent exit or entry in ACOs formed in 2014 using 2008-2014 data. Differences in risk score growth were examined using a linear spline model that included market fixed effects, year fixed effects, the previously described beneficiary characteristics, beneficiary MSSP status (always vs. never vs. enter vs. exit), splines for the years 2012-2013 (when no entry or exit occurred) and 2013-2014 (when entry or exit could occur), and an interaction between MSSP status and the two splines. We then tested for differences in risk score growth estimated from this fully-interacted spline model

Chapter III. Low-value care and clinician engagement in the Medicare Shared Savings Program: a survey of frontline clinicians

Abstract

Background: Although the Medicare Shared Savings Program (MSSP) created organizational incentives to improve healthcare value, Accountable Care Organizations (ACOs) have achieved only modest reductions in the use of low-value care. It is unknown whether this reflects limited engagement of frontline clinicians charged with implementing ACO objectives.

Objective: To assess ACO engagement of clinicians and whether engagement was associated with clinicians' reported difficulty implementing recommendations against low-value care.

Design: Cross-sectional survey of clinicians in an MSSP ACO between February 2018 and August 2018.

Participants: 1,289 clinicians in the Physician Organization of Michigan ACO, including physician specialists (27%), primary care physicians (18%), advanced practice nurses (18%), internal medicine specialists (16%), physician assistants (11%), and surgeons (10%). Response rate was 34%.

Main Measures: Primary exposures included clinicians' participation in ACO decision-making, awareness of ACO incentives, perceived influence on practice, and perceived effect on quality improvement. Our primary outcome was clinicians' reported difficulty implementing recommendations against low-value care.

Results: Few clinicians participated in the decision to join the ACO (3%). Few clinicians were aware of ACO incentives, including knowing that the ACO was accountable for both spending and quality (23%), successfully lowered spending (9%), or faced upside risk only (3%). Few agreed (moderately or strongly) that the ACO changed compensation (20%), overall practice (19%), or feedback (15%) or that it improved care coordination (17%), inappropriate care (13%), or unnecessary hospitalizations (12%). Clinicians reported difficulty following recommendations against low-value care 18% of the time; clinicians reported patients had difficulty accepting recommendations 37% of the time. Increased ACO awareness (one standard deviation [SD]) was associated with decreased difficulty (-2.3 percentage point [pp]) implementing recommendations (95% confidence interval [CI]: -3.8, -0.8), as was perceived effect on quality improvement (change in 1 SD, -1.9 p.p., 95% CI, -3.2, -0.5). Participation in ACO decision-making

and perceived influence on practice were not associated with reported implementation of recommendations.

Conclusions: Clinicians participating in ACOs were broadly unaware of and unengaged with ACO objectives and activities. Limited engagement of ACO clinicians may hamper ACO efforts to reduce low-value care.

Introduction

Encouraging clinicians to decrease inappropriate or low-value care is a central goal of payment reform. In the Medicare Shared Savings Program (MSSP), Accountable Care Organizations (ACOs) assume responsibility for the spending and quality outcomes of a defined group of patients. Despite the potential of shared savings, Medicare ACOs have achieved only modest improvements to spending and quality.[11,84] It is unclear what has hampered ACO success. One potential source is limited organizational engagement of frontline clinicians charged with implementing ACO objectives. Since the inception of the ACO model, many policymakers have commented on the importance of fostering clinician awareness of ACO goals[7,43] and payment structures that align clinicians' and organizations' incentives and norms.[9,38–42] Others have questioned whether exposing clinicians to financial incentives is necessary or desirable, with some practice leaders using non-financial incentives (e.g., performance feedback) or forgoing savings incentives altogether (e.g., retaining productivity-based bonuses).[9,19,36,37,44]

The perspective of individual clinicians has been largely absent from these policy conversations. Research has relied largely on data gathered from executive surveys and interviews.[10,20,45–48] One national survey has focused on primary care physicians (PCPs) and internal medicine specialists during the early ACO experience (2014-2015), finding that Medicare ACOs had limited success engaging physicians in decision-making, awareness of ACO incentives, or changing care delivery.[48] There are no data on MSSP ACO engagement of other physician specialists (e.g., surgeons,

anesthesiologists, dermatologists) or advanced practice providers (e.g., nurse practitioners), despite their substantial contribution to total spending[85] and the fact that MSSP ACOs are held responsible for the cost and quality performance of all participating clinicians. It is also unknown whether ACO engagement of individual clinicians is necessary to improve ACO spending and quality performance.

In this study, we asked two research questions: 1) Have ACOs successfully engaged frontline clinicians?; and 2) Is ACO engagement of frontline clinicians associated clinicians' reported ability to implement recommendations against low-value care? We designed and administered a survey of individual clinicians in the Physician Organization of Michigan (POM) ACO, the largest MSSP ACO in Michigan and among the ten largest in the county.[49] We hypothesized that clinicians would report limited ACO engagement but that ACO engagement would be positively associated with clinicians' reported ability to implement recommendations against low-value care.

Methods

Study Design. We conducted a cross-sectional survey between February and July 2018. Our survey assessed four dimensions of ACO engagement of clinicians: 1) involvement in the decision to join ACO; 2) awareness of ACO incentives and initiatives; 3) perceived influence of the ACO on practice; and 4) perceived effect of the ACO on quality improvement. Our survey also assessed clinicians' reported ability to follow

recommendations against low-value care. This study was deemed exempt from review by the University of Michigan Health Sciences and Behavioral Sciences Institutional Review Board.

Setting and participants. The POM ACO is the largest MSSP ACO in Michigan and among the ten largest in the county,[49] comprising 5,128 clinicians and approximately 80,000 attributed beneficiaries. The POM ACO includes ten clinician organizations: Michigan Medicine; Integrated Health Associates; Huron Valley Physicians Association; MidMichigan Health; St Mary's of Michigan; Answer Health; Wexford-Crawford PHO; Oakland Southfield Physician; Oakland Southfield Physician; and Olympia Medical Services (Table III.1). The survey's target population included all clinicians listed as participants in the POM ACO administrative roster. This includes physicians, physician assistants, advanced practice nurses (nurse practitioners, certified nurse anesthetists, certified nurse midwives), and other clinicians (clinical social workers, psychologists, audiologists, podiatrists, optometrists, chiropractors, and physical therapists).

Survey development. We selected survey domains from existing literature and based on four semi-structured interviews with ACO leaders. We then created or adapted instruments that mapped to those domains. The instrument was refined based on two cognitive interviews and pretests with a purposive sample of ten clinicians representing diverse specialties (e.g., urologists, NPs, PCPs, interventional cardiologists). The survey domains, instruments, and adapted instrument sources[20,48,86–91] are described in the Supplemental Methods and Table C1 in Appendix C.

Exposures. Our main exposures encompassed the following four dimensions of ACO engagement (Table III.2, Tables C2-C4 in Appendix C). First, we asked respondents to indicate their level of involvement in the decision to participate in the ACO (involved in the decision-making process, not involved but aware, not involved or aware). Second, we assessed respondents' awareness of ACO incentives and initiatives, such as whether ACO was held accountable for both spending and quality. Third, we assessed respondents' perception of the ACOs' influence on their practice, for example whether joining an ACO had changed how the respondent practices medicine. Fourth, we assessed respondents' perception of ACOs' effect on quality improvement, for example whether joining an ACO had had a positive impact on care coordination.

Outcome. Our main outcome was clinician's reported difficulty implementing recommendations against low-value care. We presented respondents with four recommendations drawn from the *Choosing Wisely*® campaign.[92] All respondents were presented with the following recommendation, "Don't recommend cancer screening in adults with life expectancy of less than 10 years." The other three recommendations based on the respondent's specialty (Table III.3, Table C5 in Appendix C).[92] We asked two questions for each recommendation: "Do you find this recommendation easy or difficult to follow most of the time?" (easy to follow, difficult to follow, does not apply to my practice); and "Do most patients find this recommendation easy or difficult to accept?" (easy to accept, difficult to accept, does not apply to my practice).

Survey administration. We administered the pilot (n=100) and full (n=5,028) survey in February and May 2018, respectively, including responses from each survey in the cohort of eligible respondents. We mailed survey invitations to clinicians' practice addresses containing the survey description, a token $2 incentive, a stylish coaster, a unique access code, and a link to the online survey (hosted by Qualtrics).[93] Non-respondents received up to three additional follow-up reminders at 1, 2, and 5 weeks by email if possible (74% of roster) or postcard.

Analysis. We estimated linear probability fixed-effects models to assess the relationship between ACO engagement and the probability of a clinician reporting difficulty implementing a given recommendation against low-value care. Our analysis was conducted at the clinician-recommendation-response level. For each dimension of ACO engagement (e.g., the ACO Awareness Scale), we estimated three models.

First, we estimated an unadjusted model that did not account for clinician or organizational characteristics but did account for the fact that different clinical specialties were shown different recommendations. To do so, this model included fixed effects for each unique recommendation displayed in the survey across all respondents (n=63), each specialty-specific block of recommendations (n=27), and whether the prompt pertained to the clinician following the recommendation or the patient accepting the recommendation. By evaluating only within-specialty variation, this model captured potential confounding introduced by variation across specialties in ACO engagement

and *Choosing Wisely*® recommendations' strength of evidence[94] and clinical and financial relevance.[95–97]

Second, we estimated a model that also adjusted for unobserved differences across the ten clinician organizations in the ACO by adding fixed effects for the clinician's organization. Finally, we estimated a model that further adjusted for clinician gender, age, clinician type/specialty (primary care physician, physician with internal medicine specialty, physician with other specialty, surgeon, physician assistant, advanced practice nurse), and professional activity (direct patient care, teaching, research, administration/management, other). In this model, we compared differences in ACO engagement and recommendation implementation among clinicians in the same specialty, of the same clinician type, and practicing within the same organization.

We used the American Association for Public Research RR1 response rate for the overall survey.[98] After survey administration, we restricted our sample to clinician types most frequently represented in the *Choosing Wisely*® campaign (physicians, physician assistants, advanced practice nurses), excluding clinical social workers, psychologists, audiologists, optometrists, podiatrists, chiropractors, physical therapists, and other/unknown (PhD, MBA, MHSA). We excluded clinicians who responded "does not apply to my practice" for all recommendations against low-value care. We also excluded pediatricians, as they are not typically included in Medicare initiatives targeting adult beneficiaries, as well as any clinician who was a member of the research team or

included in survey pre-testing (see Figure C1 and Supplemental Methods of Appendix C for CONSORT diagram and response rate calculation).

To reduce potential bias from survey nonresponse and generalize estimates to the target population (the POM ACO), we applied post-stratification survey weights incorporating characteristics associated with non-response (in this case, clinician organization). We used iterative proportional fitting, or raking, to calibrate survey weights.[99] To reduce bias from missing data among respondents, we used multiple imputation for all models and implemented a recently-developed quadratic-rule procedure to select the number of imputations needed to achieve estimate and standard error replicability.[100]

Results

Respondent characteristics. Of the 4,701 eligible respondents, 1,620 completed the survey (response rate of 34%; Figure C1 in Appendix C). Response rates differed across the clinician organizations but were otherwise comparable across clinician specialty and sex (Table C6 in Appendix C). The analytic sample for the present analysis included 1,289 respondents (Table III.1). Respondents represented a wide range of clinician types and specialties, including physician specialists (27%), PCPs (18%), internal medicine specialists (16%), and surgeons (10%), as well as advanced practice nurses (18%) and physician assistants (11%; Table III.1). Direct patient care was the most common professional activity (85%). Most respondents were either

employed by a medical school (59%) or hospital (25%), and a majority served on the clinical staff of Michigan Medicine (78%).

ACO engagement. ACO engagement of clinicians was low (Table III.2). Most respondents were not aware or involved of the decision to join the ACO (69%); 28% were aware but not involved, and 3% were involved in the decision-making process. Respondents reported limited awareness of ACO incentives and initiatives (Table III.2, Table C2 in Appendix C). For example, 23% knew that the ACO was accountable for both spending and quality, 9% knew the ACO had successfully lowered spending, 6% knew the ACO was accountable for only Medicare patients, and 3% knew the ACO only faced upside risk (i.e., could not lose money). Across the six ACO Awareness Scale items, the mean respondent knew 0.6 items and the median respondent knew 0 items.

Respondents perceived that the ACO had minimal influence on their practice (Table III.3, Table C3 in Appendix C). Few respondents agreed (moderately or strongly) that joining an ACO has “made me more aware of controlling treatment costs” (26%), “changed how I am compensated” (20%) or “changed how I practice medicine” (18%). Only 15% felt they received “useful feedback on ACO cost and quality performance.”

Respondents perceived that the ACO had a minimally positive effect on quality improvement (Table III.3, C4 in Appendix C). Few respondents, for example, felt the ACO had a positive effect on their ability to coordinate care (18%), reduce inappropriate care (14%), or reduce unnecessary hospitalizations (13%).

Implementation of recommendations against low-value care. Respondents provided 8,448 responses for 27 specialty-specific blocks containing 63 unique recommendations against low-value care (Table III.2, Table C5 in Appendix C). On average, respondents found that recommendations were less difficult for clinicians to follow than for patients to accept. Recommendations were “difficult to follow” for clinicians 18% of the time (standard deviation, 38%). Recommendations were “difficult to accept” for patients 37% of the time (standard deviation, 48%).

Relationship between ACO engagement and recommendations against low-value care. Some dimensions of ACO engagement were associated with implementation of recommendations against low-value care (Figures III.1-3.3). In the unadjusted model, respondents who aware of the decision to join the ACO were 3.9 percentage points (p.p.) less likely to report difficulty implementing recommendations against low-value care (95% confidence interval [CI]: -6.2 to -1.6), as compared to respondents who were not aware or involved in the decision (Figure III.1). This relationship was not statistically significant after adjusting for organization and clinician characteristics (-1.9 p.p., 95% CI: -4.3 to 0.5).

Increased awareness of ACO incentives and initiatives was associated with greater reported ability to implement recommendations against low-value care (Figure III.2). After adjusting for organization and clinician characteristics, a one standard deviation (SD) increase in ACO awareness was associated with 2.3 p.p. less reported difficulty

implementing recommendations against low-value care (95% CI: -3.8, -0.8), representing an 9% improvement (~-2.3 p.p./base likelihood of 27.0%). Perceived influence of ACO on practice was not associated with respondents' reported difficulty implementing recommendations against low-value care (Figure III.3, Panel A; 1.1 p.p., 95% CI: -0.3 to 2.6). Conversely, a 1 SD increase in perceived quality improvement was associated with 1.9 p.p. less reported difficulty implementing recommendations (Figure III.3, Panel B; 95% CI: -3.2 to -0.5).

Discussion

In a survey of one of the largest MSSP ACOs in the country, we found limited engagement of the frontline clinicians charged with implementing ACO value-based initiatives. Few clinicians participated in the decision to join the ACO, fewer yet were aware of new organizational financial incentives created by the MSSP, and most reported the ACO had limited effect on practice or quality improvement. At the same time, greater clinician engagement – in particular, improved awareness of ACO incentives and ability to improve care quality – was associated with a modest improvement in clinicians' reported ability to implement recommendations against low-value care. Taken together, our results suggest that limited engagement of ACO clinicians may hamper ACO efforts to reduce low-value care.

There are few data on the degree to which ACOs have engaged individual clinicians in efforts to improve health care value. A national survey of PCPs (~78%), internal

medicine specialists (~18%) during the early MSSP experience (2014-2015) found MSSP ACOs had a modest perceived effect on practice change (e.g., half agreed ACOs had influenced care).[48] Our study extends these results, finding little ACO engagement among physician specialists (e.g., anesthesiologists, dermatologists), surgeons, physician assistants, and advanced practices nurses. The positive association between ACO incentive awareness and reported ability to implement recommendations against low-value care suggests that that improved communication of ACO rules and structures to clinicians may help ACO leaders to achieve MSSP objectives. Our results are also consistent with the finding that Pioneer ACOs have achieved modest reductions in low-value care,[84] suggesting that inconsistent engagement of ACO clinicians may underlie modest results overall.

The ACO model comprises a wide diversity of ACOs. Studies suggests that ACOs based in physician practice networks have had greater success lowering spending than hospital-integrated ACOs.[21] Our finding of low clinician engagement in the POM ACO (which includes multiple hospitals) is consistent with the possibility that hospital-integrated ACOs' inability to lower spending may be partially due to greater difficulty engaging frontline clinicians.

At the same time, the diversity of ACOs also belies an essential commonality – all ACOs rely on frontline clinicians to improve quality, eliminate low-value care, and achieve spending and quality objectives.[48] ACOs' limited engagement of individual clinicians observed in this and previous studies may help to explain ACOs' relatively

modest impact overall.[14,21–23,48,84] The disjunction between clinician and organization incentives extends well beyond the ACO model. Even as organizations increasingly take on risk-based contracts, the most common and substantial financial incentive continuing to face clinicians is "productivity,"[101] even in practices participating in ACOs[36] or other global payment reforms.[101] Our results suggest that transmitting organizational incentives to individual clinicians is a challenge but one that may help to achieve organizational goals.

Our results shed additional light on a recent survey of PCPs' perceived barriers to implementing recommendations against low-value care, which included perceived demands from patients and specialists for treatment and tests, lack of time for shared decision-making, and performance systems that reward ordering more services.[86] We found that perceived improvements in related dimensions of ACO engagement (in the Quality Improvement scale, e.g., shared decision-making, managing care between visits; and in the ACO Awareness scale, e.g., knowledge of changes to payment incentives) were associated with greater reported ability to implement recommendations.

Our study must be interpreted in the context of several limitations. First, data from a single, large, Medicare ACO in Michigan may not generalize to other ACOs. Second, the moderate response rate (34%) raises the possibility of response bias. Although it is possible that clinicians with strong opinions about ACOs might have disproportionately

chosen to take our survey, this seems unlikely given respondents' limited ACO awareness and relatively tepid perceptions of change in the ACO. Our response rate may reflect increasing difficulty conducting clinician surveys in recent years, particularly without large financial incentives[102] Third, our study measured clinicians' reported ability to implement recommendations against low-value care and not actual practice behavior. Finally, causal inference is limited by the cross-sectional nature of this study. Although our analytic approach controlled for fixed differences across organizations, clinician types and specialties, and other observed clinician characteristics, it is also possible that clinicians with greater reported ability to implement recommendations choose to engage ACOs in a more effective manner.

Conclusions. Systematic improvements in the value of care provided by ACOs will likely require greater engagement of individual clinicians tasked with preserving quality and reducing unnecessary care. Our study suggests that current ACO efforts may be hampered by individual clinicians' uncertainty regarding ACO incentives and initiatives. Moving forward, research in this area should focus on evaluating the relationship between ACO engagement and changes in observed clinician behavior. In addition, it will be important to establish methods by which ACOs can more effectively engage individual clinicians in efforts to redesign care and improve health care value.

Table III.1. Characteristics of respondents in accountable care organization

Respondent characteristics (N=1,277)	
Specialty	
Primary care physician	18%
Physician, medicine specialty	16%
Physician, other specialty	10%
Surgeon	27%
Physician assistant	11%
Advanced practice nurse	18%
Gender	
Male	49%
Female	51%
Age	
Under 35	13%
35-44	34%
45-54	22%
55-64	22%
65 or older	8%
Professional activity	
Direct patient care	85%
Teaching	3%
Research	7%
Administration/management	3%
Other	2%
Practice arrangement	
Employed by a medical school	59%
Employed by a hospital	25%
Employed by a managed care organization	1%
Solo practice	2%
Single specialty group practice	4%
Multi-specialty group practice	7%
Other	2%
Professional activity	
Direct patient care	78%
Teaching	9%
Research	4%
Administration/management	3%
Other	1%
Clinician organization	1%
Michigan Medicine	2%
IHA	2%
Huron Valley Physicians Association	0%
MidMichigan	0%
St Mary's of Michigan	18%

Advance Health	16%
Wexford PHO Crawford PHO	10%
Oakland Southfield Physician	27%
Oakland Southfield Physician	11%
Olympia Medical Services	18%

Notes: Physicians included MDs, DOs, and international degrees such as MBBS. Advanced practice nurses included certified nurse midwives, certified nurse anesthetists, and nurse practitioners. Primary care physicians included the following specialties: Family Medicine; General Internal Medicine; Geriatrics; Palliative medicine; Hospitalist; Primary care; General Practice; and Preventive Medicine.

Table III.2. ACO engagement measures and scales

ACO engagement	
Participation in decision to join ACO	**% Respondents**
I was involved in the decision-making process	3%
I was not involved but was aware of the decision-making process	28%
I was not involved or aware of the decision-making process	69%
ACO Awareness (1=correct,0=incorrect or don't know)	**% Respondents**
Held accountable for both spending and quality	23%
Lowered spending in most recent performance year	9%
Emphasized conducting Welcome to Medicare and Annual Wellness visits	17%
Did not receive a financial bonus from Medicare in most recent performance year	3%
Does not face downside financial risk, i.e., cannot lose money	3%
Held accountable for only Medicare patients	6%
Six-item ACO Awareness Scale (possible range: 0 to 6, Cronbach α score, 0.72)	Total=0.6
ACO Practice Change (1=strongly disagree, 2=moderately agree, 3=moderately agree, 4=strongly agree)	**Mean (SD)**
Joining an ACO has changed how I practice medicine	1.7 (0.8)
Joining an ACO has improved the quality of care my patients receive	1.9 (0.8)
Joining an ACO has made me more aware of controlling treatment costs	1.9 (0.8)
Joining an ACO has changed how I am compensated	1.7 (0.8)
I receive useful feedback on ACO cost and quality performance	1.6 (0.8)
ACO financial bonuses are large enough to influence my behavior	1.4 (0.6)
Six-item ACO Practice Change Scale (possible range: 6 to 24, Cronbach α score, 0.88)	Total=10.3
ACO Quality Improvement (-1=negative impact,0=no impact or don't know,1=positive impact)	**Mean (SD)**
Coordinate care across care settings	0.1 (0.4)
Decrease unnecessary hospitalizations	0.1 (0.3)
Help patients manage care between visits	0.1 (0.4)
Engage in shared decision-making	0.1 (0.4)
Reduce inappropriate or harmful care	0.1 (0.4)
Improve health of low-income patients	0.1 (0.3)
Improve health of medically complex patients	0.1 (0.4)
Seven-item ACO Quality Improvement Scale (possible range: -7 to 7, Cronbach α score, 0.88)	Total=0.9

Scale development is described in the main text. Survey weights were applied to generalize to the Physician Organization of Michigan ACO. Multiple imputation was used for missing data. ACO is accountable care organization. SD is standard deviation.

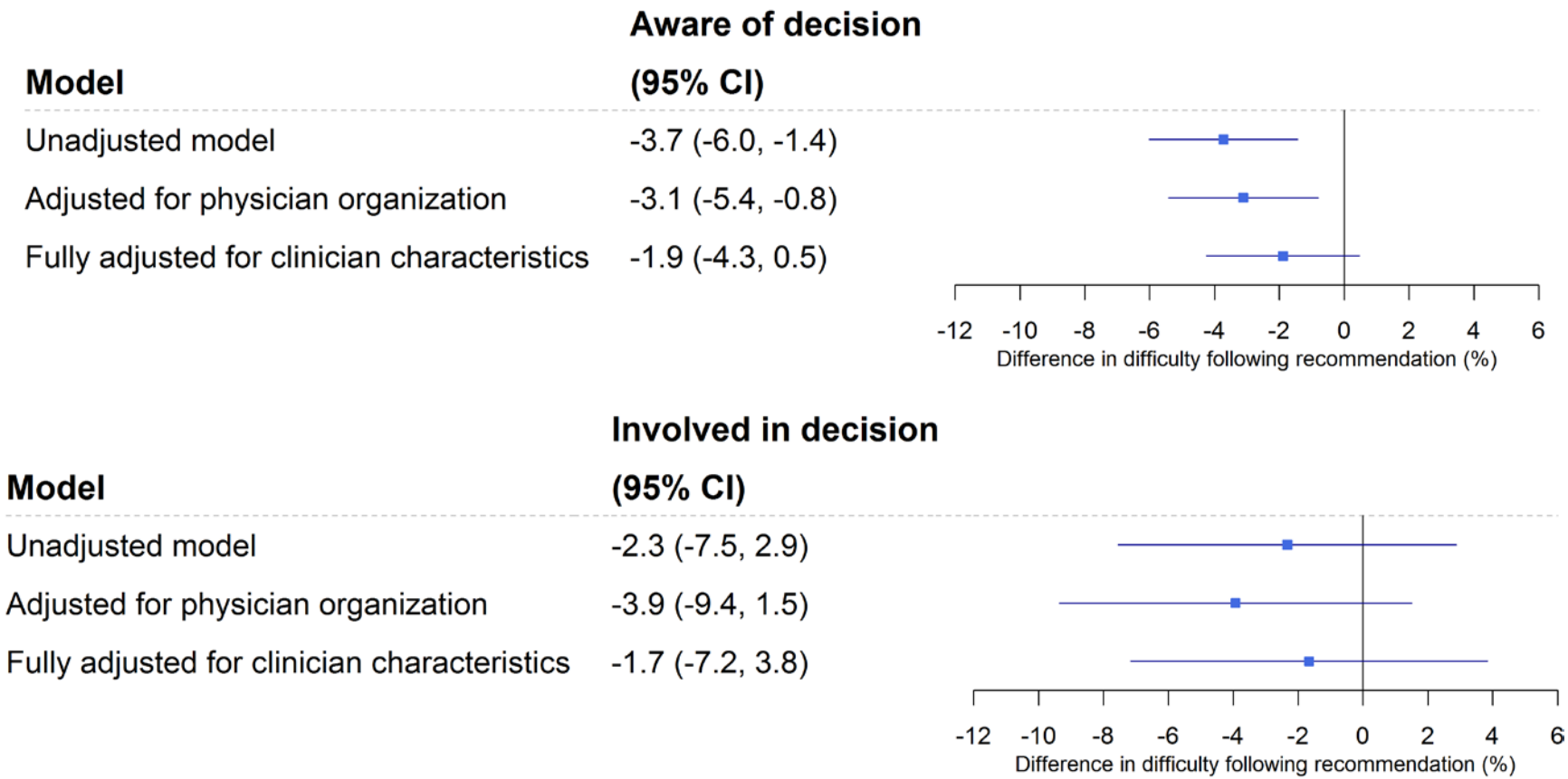
Table III.3. Examples of recommendations against low-value care presented to respondents

Specialty	Examples of recommendations against low-value care	Difficult for clinician to follow most of the time	Difficult for most patients to accept	No. responses
Overall sample		17.7%	36.4%	8470
All specialties	Don’t recommend cancer screening in adults with life expectancy of less than 10 years.	36.5%	58.1%	1649
Anesthesiology	Don’t administer packed red blood cells (PRBCs) in a young healthy patient without ongoing blood loss and hemoglobin of ≥ 6 g/dL unless symptomatic or hemodynamically unstable.	10.1%	17.6%	200
Cardiology	Don’t perform annual stress cardiac imaging or advanced non-invasive imaging as part of routine follow-up in asymptomatic patients.	2.9%	23.2%	139
Dermatology	Don’t routinely use topical antibiotics on a surgical wound.	0.0%	10.0%	40
Emergency Medicine	Don’t image for suspected pulmonary embolism (PE) without moderate or high pre-test probability of PE.	23.2%	21.4%	112
Endocrinology	Don’t prescribe testosterone therapy unless there is biochemical evidence of testosterone deficiency.	0.0%	64.7%	34
Gastroenterology	For a patient with functional abdominal pain syndrome (as per ROME IV criteria) computed tomography (CT) scans should not be repeated unless there is a major change in clinical findings or symptoms.	6.3%	50.0%	32
Hematology	Don’t administer packed red blood cells (PRBCs) in a young healthy patient without ongoing blood loss and hemoglobin of ≥ 6 g/dL unless symptomatic or hemodynamically unstable.	25.0%	65.0%	40
Neurology	Don’t perform imaging of the carotid arteries for simple syncope without other neurologic symptoms.	6.1%	6.3%	65
Obstetrics and Gynecology	Don’t treat patients who have mild dysplasia of less than two years in duration.	12.3%	35.9%	129
Oncology	Don’t perform surveillance testing (biomarkers) or imaging (PET, CT, and radionuclide bone scans) for asymptomatic individuals who have been treated for breast cancer with curative intent.	11.4%	36.4%	68

Pathology	Don’t test vitamin K levels unless the patient has an abnormal international normalized ratio (INR) and does not respond to vitamin K therapy.	0.0%	25.0%	14
Physical Medicine and Rehabilitation	Don’t do imaging for low back pain within the first six weeks, unless red flags are present.	11.5%	50.0%	52
Primary Care	Don’t routinely prescribe antibiotics for acute mild-to-moderate sinusitis unless symptoms last for seven or more days, or symptoms worsen after initial clinical improvement.	8.7%	69.8%	575
Psychiatry	Don’t routinely prescribe two or more antipsychotic medications concurrently.	13.5%	5.3%	75
Rheumatology	Don’t prescribe biologics for rheumatoid arthritis before a trial of methotrexate (or other conventional non-biologic DMARDs).	7.1%	28.6%	28
Surgery	Don’t image for suspected pulmonary embolism (PE) without moderate or high pre-test probability of PE.	34.9%	36.6%	336
Urology	Don’t routinely perform PSA-based screening for prostate cancer.	68.2%	72.7%	44

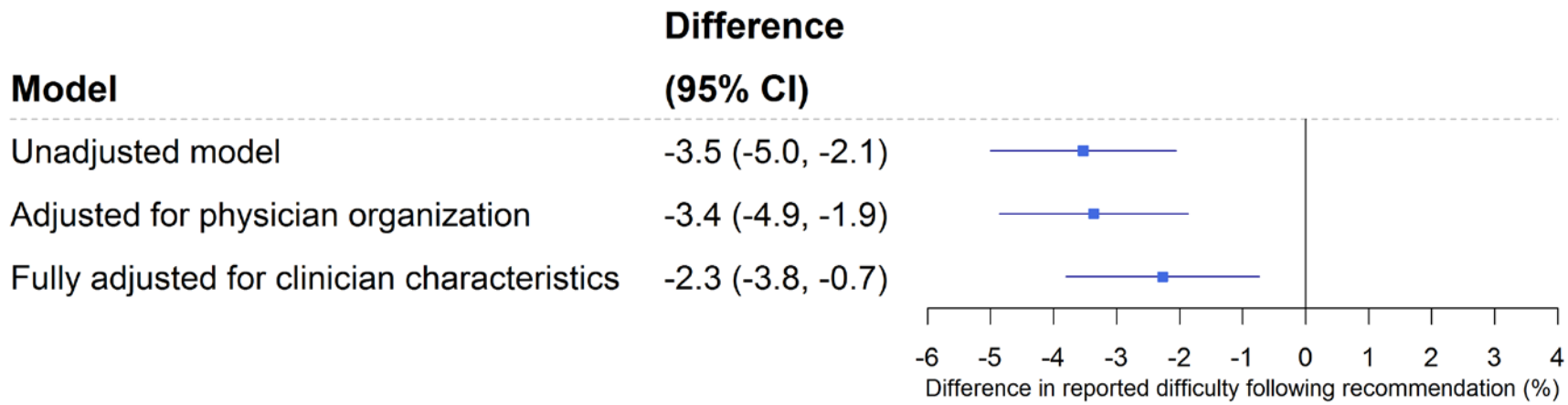
Respondents were shown four recommendations against low-value care. Recommendations were drawn from the *Choosing Wisely*® campaign. For reach recommendation, respondents were asked two questions: “Do you find this recommendation easy or difficult to follow most of the time?” (easy to follow, difficult to follow, does not apply to my practice); and “Do most patients find this recommendation easy or difficult to accept?” (easy to accept, difficult to accept, does not apply to my practice). All respondent were given the following recommendation: “Don’t recommend cancer screening in adults with life expectancy of less than 10 year,” as well as three specialty-specific recommendations. One specialty (Physical Medicine and Rehabilitation) was not shown the cancer screening recommendation due to survey administration error. Regression analyses of the association between ACO engagement and recommendations against low-value care excluded responses where clinicians indicated the recommendation “Does not apply to my practice” (n=1517, 17.8% of total responses). Survey weights were applied to generalize to the Physician Organization of Michigan ACO.

Figure III.1. Relationship between clinician involvement in decision to join ACO and low-value care



The reference group is clinicians who were not involved in or aware of the decision to join the ACO. Models are described in the main text. Survey weights were applied to generalize to the Physician Organization of Michigan ACO. Multiple imputation was used for missing data. ACO is accountable care organization. CI is confidence interval.

Figure III.2 Relationship between clinician ACO awareness and low-value care

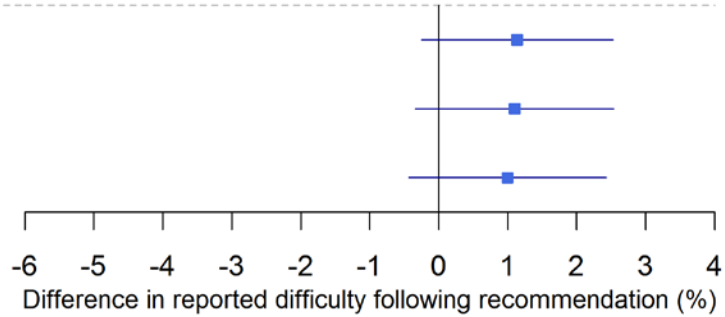


Estimated change is for a one standard deviation increase in the ACO Awareness scale. The scale and models are described in the main text. Survey weights were applied to generalize to the Physician Organization of Michigan ACO. Multiple imputation was used for missing data. ACO is accountable care organization. CI is confidence interval.

Figure III.3. Relationship between perceived ACO impact on practice and quality and low-value care

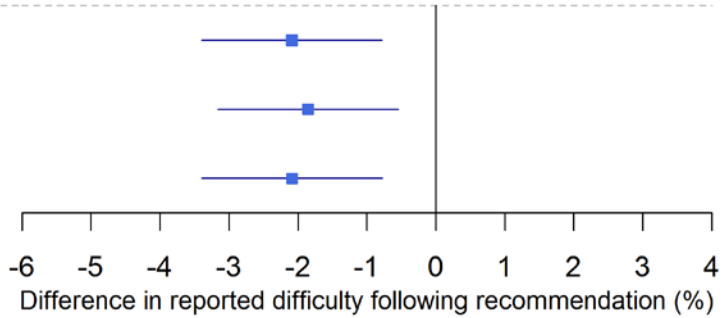
Panel A. ACO Practice Change scale

Model	Difference (95% CI)
Unadjusted model	1.1 (-0.3, 2.5)
Adjusted for physician organization	1.1 (-0.3, 2.5)
Fully adjusted for clinician characteristics	1.0 (-0.4, 2.4)



Panel B. ACO Quality Improvement scale

Model	Difference (95% CI)
Unadjusted model	-2.1 (-3.4, -0.8)
Adjusted for physician organization	-1.9 (-3.2, -0.5)
Fully adjusted for clinician characteristics	-2.1 (-3.4, -0.8)



Estimated change is for a one standard deviation increase in either the ACO Practice Change scale (Panel A) or the ACO Quality Improvement scale (Panel B). Scales and models are described in the main text. Survey weights were applied to generalize to the Physician Organization of Michigan ACO. Multiple imputation was used for missing data. ACO is accountable care organization. CI is confidence interval.

Appendices

Appendix A. Changes in Spending and Quality in the Medicare Shared Savings Program After Accounting for Non-Random Exit

Table of Contents

Supplemental Methods

A. Study data, beneficiary attribution, and outcome specification

B. Calculating Medicare Shared Savings Program (MSSP) supply as an instrumental variable

C. Assessing differential trends in spending prior to the MSSP

D. Estimating the impact of the MSSP on changes in spending and quality

E. Evaluating compositional changes in the MSSP

F. Instrumental variable sensitivity analyses

G. Supplemental analyses of MSSP exit and entry by provider groups

Figure A1. Schematic of the MSSP supply instrumental variable

Figure A2. Pre-period trends in spending across the instrumental variable

Figure A3. Pre-period trends in spending across observed MSSP participation status

Figure A4. Changes in total Medicare spending according to ACOs' year of entry into the MSSP

Figure A5. Association between county-level MSSP penetration and Medicare Advantage penetration

Figure A6. Robustness tests of MSSP supply as an instrumental variable

Figure A7. Changes in hospital use for beneficiaries attributed to MSSP ACOs vs. controls

Figure A8. Sensitivity analyses of association clinician spending performance and probability of exiting an MSSP ACO

Figure A9. Association between provider group spending performance and probability of exiting or entering an MSSP ACO

Table A1. Non-linear estimation of discount factor used in MSSP supply

Table A2. Pre-period trends in spending across observed MSSP status

Table A3. Formal test of statistical differences between instrumental variable and adjusted longitudinal model

Table A4. Sensitivity analyses of adjusted longitudinal and instrumental variable models

Table A5. Estimates of the association between prior spending and probability of joining or exiting an MSSP ACO

Table A6. Simulated changes in average MSSP beneficiary spending with attrition of high-cost clinicians from MSSP ACOs (2013-2014)

Table A7. Simulated changes in average MSSP beneficiary spending when including clinicians exiting MSSP and excluding clinicians entering the MSSP (2013-2014)

Supplemental Figure legends

Supplemental Methods

Section A. Study data, beneficiary attribution, and outcome specification

i. Study data. Accountable care organizations (ACOs) entering Medicare Shared Savings Program (MSSP) contracts receive financial bonuses ("shared-savings") from the Centers for Medicare and Medicaid Services (CMS) if they lower average total Medicare spending below a spending benchmark and meet quality performance standards. To estimate the association between the MSSP and changes in spending, hospital use, and quality performance, our analysis strategy required longitudinal Medicare claims data at the beneficiary-quarter level before and after the staggered start of the MSSP (to create a pre-post study design that precisely captured entry to and exit from MSSP contracts); MSSP data on who belongs to which ACOs; spatial data from CMS and the US Census on where clinicians work and beneficiaries live (to calculate MSSP supply); and, to control for confounders, beneficiary information from the American Community Survey (ACS) and CMS claims and enrollment files.

Medicare claims and enrollment files. Our analyses were based on a 20% random sample of Medicare claims, including the following Research Identifiable Files (RIFs): Medicare Provider Analysis and Review (MedPAR), which captures payments for inpatient and skilled nursing facility services; Carrier, which captures payments for physician services; Outpatient; Master Beneficiary Summary File, which captures patient enrollment and sociodemographic information; the ACO Beneficiary-Level file,

which identifies beneficiaries attributed to the MSSP and the ACOs to which they were attributed; and the ACO Provider-Level file, which identifies the clinicians and provider groups, (as identified by Taxpayer Identification Number [TIN]), participating in MSSP ACOs.

American Community Survey (ACS). We used ACS data to define beneficiary area-level poverty and educational attainment for each year of the study. We obtained five-year estimates of ZIP code tabulation area (ZCTA) poverty (proportion living below 100% of the federal poverty level) and educational attainment (proportion with a high school degree or equivalent, proportion graduated from college).[72] We used the final year of the 5-year estimate for each study period (e.g., ACS 2014-2010 data for study year 2014). Because complete, ZCTA-level 5-year ACS estimates only began in the 2006-2010 wave, we used 2006-2010 data for study years 2008 through 2010. For each year of the analysis, we linked beneficiaries to ZCTA-level ACS data using beneficiary ZIP codes provided in CMS Master Beneficiary Summary File.

***US Census Gazetteer*.** For each year of the study, we identified the latitude and longitude of all beneficiary residencies and clinician practice locations. To do so, we drew on ZCTA-level geographic data provided by the US Gazetteer data and collected by the US Census (further linkage information is provided below).

Rural-urban continuum areas (RUCA). We used RUCA codes to categorize beneficiaries and clinicians as residing in urban, large rural, small rural, or isolated ZIP

codes.[70] We followed the RUCA categorization scheme to categorize the 33 RUCA codes in the following manner: urban (RUCA codes 1.0, 1.1, 2.0, 2.1, 3.0, 4.1, 5.1, 7.1, 8.1, 10.1); large rural cities/towns (4.0, 4.2, 5.0, 5.2, 6.0, 6.1); small rural towns (7.0, 7.2, 7.3, 7.4, 8.0, 8.2, 8.3, 8.4, 9.0, 9.1, 9.2); and Isolated (10.0, 10.2, 10.3, 10.4, 10.5, 10.6).

Hospital Referral Region (HRR). We linked beneficiary and clinician ZIP codes to HRRs using a crosswalk provided by the Dartmouth Atlas of Health Care.[103] We excluded beneficiaries who resided outside of HRRs or were attributed to clinicians who practiced outside of HRRs (see Figure A1 for CONSORT diagram).

ii. Beneficiary Attribution

We defined our main exposure of interest, beneficiary attribution to an MSSP ACO, as a time-varying indicator that equaled 1 if the beneficiary was in an MSSP ACO in the given quarter, 0 otherwise. We used a time-varying quarterly ACO indicator rather than stably attributing beneficiaries to an ACO or control TIN for the entire study period for two reasons. First, this captured the staggered entry of ACOs into MSSP contracts, which occurred in April 2012, July 2012, January 2013, and January 2014. Second, this reflected the fact that beneficiaries and clinicians can enter, exit, and transition between MSSP ACOs each year.

For each year of the study, we attributed beneficiaries to the provider group and clinician from which they received the plurality of eligible primary care services. Following MSSP specifications[104] and prior work,[21,22,58] we first linked beneficiaries to the provider TIN from which they received the plurality of eligible primary care services. To do so, we identified whether the beneficiary was attributed to an MSSP ACO using the ACO Beneficiary-Level RIF. Next, for those beneficiaries attributed to an MSSP ACO, we identified the TINs participating in the ACO for that given year. Among those TINs, we then attributed beneficiaries to the TIN from which they received the plurality of eligible primary care services, defined by the MSSP as the sum of eligible outpatient evaluation and management services (Healthcare Common Procedure Coding System codes: 99201 through 99215; 99304 through 99350; G0402; G0438; and G0439.)[104]

We next attributed beneficiaries to a specific clinician within the ACO-TIN to which they were attributed above. Although the MSSP does not attribute beneficiaries to individual clinicians within a TIN or ACO, our study needed to form beneficiary-clinician units in order to define the instrumental variable (MSSP supply) at the level of the individual clinician. To ensure that beneficiaries were attributed to clinicians within the ACO-TIN considered accountable for beneficiary outcomes by the MSSP, we mirrored MSSP attribution methodology in attributing beneficiaries to the clinician within the attributed ACO-TIN who provided the plurality of eligible primary care services, as defined above. Because CMS extended the performance for ACOs entering MSSP contracts in either April or July 2012 through the end of 2013, we attributed MSSP beneficiaries to MSSP TINs and clinicians from which they received the plurality of eligible services across the

years 2012 and 2013. We followed an analogous procedure for attributing control beneficiaries to non-MSSP TINs and clinicians, defining controls as those beneficiaries not listed in the ACO Beneficiary-Level RIF in the given year of interest.

As specified by the MSSP, we divided the attribution process described above into two steps. In step 1 of the MSSP algorithm, we attributed beneficiaries to a TIN on the basis of only those eligible services furnished by a primary care physician (PCP), defined by claim specialty codes for general practice, family practice, internal medicine, or geriatric medicine. If a beneficiary could be attributed in this manner, we then attributed the beneficiary to the PCP who furnished the plurality of eligible services within the attributed TIN.

Following prior work[21,22,58] and in order to make ACO and control beneficiaries more comparable, we attributed beneficiaries to TINs on the basis of outpatient claims alone and excluded beneficiaries who were: 1) attributed on basis of eligible services furnished by non-PCP clinicians eligible for attribution in the MSSP (i.e., step 2 of the MSSP algorithm); 2) attributed to a clinician who was ineligible for attribution in the MSSP; or 3) not attributed to any clinician (because they did not receive eligible services).

iii. Outcome specifications

Spending. We defined total spending per Medicare beneficiary as the total price-standardized spending for inpatient, outpatient, physician, and skilled nursing facility services. We price-standardized to account for variations resulting from regional wage indices and payments for indirect medical education, Disproportionate Share Hospitals, and new technologies. We defined component spending as inpatient care (from the MedPAR RIF), outpatient care (from the Outpatient RIF), professional services (from the Carrier RIF), and skilled nursing facility spending (from the MedPAR RIF).

Hospitalization. We defined all-cause hospitalizations as any inpatient stay during the quarter of interest. We defined 30-day readmissions according to MSSP specifications.[62] We included 30-day all-cause readmissions as this outcome is included in MSSP contracts. We defined preventable hospitalizations as hospitalization for any 1 of 11 ambulatory care-sensitive conditions (ACSCs), defined by the Agency for Healthcare Quality and Research (AHRQ) as conditions for which "good outpatient care can potentially prevent the need for hospitalization or for which early intervention can prevent complications or more severe disease". We followed AHRQ specifications in constructing the numerators and denominators for each ACSC measure.[61] We created a composite measure comprising 11 ACSCs that we chose on the basis of meeting one of the following three criteria: 1) specifically included in ACO contracts (heart failure, chronic obstructive pulmonary disorder or asthma); 2) related to ACO quality measures for diabetes or heart disease that could not be assessed using claims data (uncontrolled diabetes, diabetes short-term complications, diabetes long-term complications, diabetes lower-leg amputation, hypertension, angina); or 3) captured quality performance for

acute conditions (dehydration, pneumonia, and urinary tract infection). For all hospitalization outcomes, we defined hospitalization as the count per beneficiary per quarter.

Clinical quality indicators. We chose 4 clinical quality process indicators that can be captured accurately in claims and are either included as quality measures in MSSP contracts or are in the clinical pathway toward quality measures included in MSSP contracts.[62] Because the MSSP includes several measures pertaining to diabetes care, we included 3 measures of preventive care services for beneficiaries with diabetes: (1) glycated hemoglobin testing (chosen as a process measure related to ACO quality measure DM-2, hemoglobin A1c poor control); (2) low-density lipoprotein cholesterol (LDL-C) testing (related to DM-13, LDL-C control); and (3) diabetic retinal examinations (included in MSSP contracts as DM-7). To match our quarterly estimation strategy, we defined the beneficiary as meeting each respective measure if he or she received the service in that quarter (1=yes, 0=no). We followed National Quality Form specifications for defining receipt of glycated hemoglobin testing, LDL-C testing, and diabetes retinal examinations.[105] We used a claims-based method for identifying beneficiaries with diabetes according to Chronic Disease Warehouse specifications, defining beneficiaries as those with the presence of specified claims in the previous year.[106] In addition to the 3 diabetes quality indicators, we included mammography provision among female beneficiaries ages 65-69 (ACO quality measure ACO-Prev-5; NQF 0031).[107] Following NQF specifications, we defined the eligible beneficiaries as meeting the measure if they

received a mammogram within the quarter of interest or any of the prior 7 quarters (1=yes, 0=no).

Hip fracture. As a falsification test, we evaluated the impact of the MSSP on changes in the rate of hospitalization for hip fracture. Following prior work,[27] we defined hip fracture as the presence of an acute care hospital claim with a primary diagnosis of hip fracture (820.xx) in a given quarter (1=yes, 0=no).

Section B. Calculating MSSP supply as an instrumental variable

A schematic depicting calculation of MSSP supply is presented below (Figure A1). To capture market-level variation, we calculated the total count of MSSP and non-MSSP clinicians within 50 miles of the clinician's practice location. We defined clinician practice location using the modal 5-digit ZIP code listed in Part B claims for that clinician for that study year. We then obtained latitude and longitude data by linking Part B ZIP codes to ZCTA-level geographic data from the US Census Gazetteer files.[108] For ZIP codes that could not be linked to US Census latitude and longitude data, we used nearby ZIP codes (indicated by numeric sequence of the ZIP code) within that same year and then across study years. We then calculated the distance between two clinicians' practice locations using Stata's user-written –geonear- command, [109] which estimates the geodetic distance between length of the shortest curve between two points (specified by latitude and longitude) along the surface of a mathematical model of the earth.

To capture small-area variation, we then discounted clinicians relatively further away within the 50-mile radius by a discount factor β.[110] Because we are unaware of any data on how distance to other ACO clinicians affects clinicians' willingness to join an ACO, we empirically estimated this discount factor. Under the hypothesis that the effect of distance on the likelihood of two clinicians entering a joint MSSP contract varies by population density, we estimated this relationship separately by urbanicity, as classified by the four RUCA categorizations defined above.

Specifically, we calculated the discount factor β by estimating how the likelihood of two clinicians participating in the same MSSP ACO varied by the inverse distance between those two clinicians. For clinicians j and k, separated by distance d, and whose practice location has urbanicity r, we used non-linear least-squares[67] to estimate the following equation:

$$\Pr\left(Same\_ACO_{jkdr}\right) = \frac{1}{{d_{jk}}^{\beta_r}}$$

where $Same\_ACO$ is 1 if clinician j belongs to the same MSSP ACO as clinician k, 0 otherwise. We implemented this non-linear analysis using Stata's –nl– command.[111] For computational reasons, we estimated the beneficiary and clinician discount factors on a 2% sample of unique clinicians, although the effective analytic size is larger, as the random samples were drawn before matching each clinician to all other clinicians within the 50-mile radius. As expected, distance exerted a strong negative effect on the probability of two clinicians participating in the same MSSP ACO, with this effect

greatest in large urban areas, intermediate in rural cities or towns, and smallest in isolated areas (eTable1). As a robustness test, we estimated the impact of the MSSP on spending using instrumental variable models that used MSSP supply restricted to within a 10-, 25-, 50-, 75-, and 150-mile radius of the clinician's practice, with or without discounting clinicians who practiced further away within the given radius.

Section C. Assessing differential trends in spending prior to the MSSP

We tested for whether spending trends prior to the start of the MSSP varied across the instrumental variable (high vs. low MSSP supply) and observed MSSP participation status (beneficiaries who ultimately joined the MSSP vs. those that never joined the MSSP). High vs. low MSSP supply was defined by whether the attributed clinician's median MSSP value was above (1=high) or equal to or below (0=low) the median MSSP supply value for the HRR over the entire post-period, which was designated independently for each HRR as beginning with the first quarter in which the HRR contained at least one clinician participating in the MSSP. Observed MSSP participation status was defined using an indicator for whether the beneficiary was ultimately attributed to an MSSP ACO (1=ever, 0=never).

We tested for differential pre-period spending trends across the instrumental by estimating a linear spline model that included an indicator for high MSSP supply, a linear term for time, a pre-period linear spline (demarcated by Quarter 1/Quarter 2, 2012, i.e., April 2012), a post-period linear spline, interaction terms between the high

MSSP supply indicator and the pre- and post-period linear splines, seasonal indicators, year fixed effects, and market fixed effects. In these spline models, differential trends in the pre-period were identified by the coefficient for the interaction term between the high MSSP supply indicator and the pre-period linear spline. We also graphically depicted adjusted trends in average spending in each quarter across the instrumental variable. Adjusted spending was estimated using a model that included market fixed effects, year fixed effects, seasonality, beneficiary characteristics, an indicator for each quarter, an indicator for high MSSP supply, and an interaction term between high MSSP supply and the quarter of interest, but did not impose an underlying time trend on the adjusted spending data. From these models, Stata's –margins- command was used to estimate predicted spending for beneficiaries with high- vs. low-MSSP supply in each quarter.

Differential pre-period spending trends across observed MSSP participation status were tested using an analogous spline model that included an indicator for whether the beneficiary was ever attributed to an MSSP ACO or never attributed to an MSSP ACO. We also graphically depicted adjusted trends in average spending in each quarter across the 4 MSSP ACO cohorts using a model similar to instrumental variable model described above, now interacting the quarter of interest with an indicator for the beneficiary's ultimate MSSP cohort status (0=never in an MSSP ACO, 1=2012 April cohort, 2=2012 July cohort, 3=2013 cohort, 4=2014 cohort).

Spline and adjusted trend analyses demonstrated that spending trends in the pre-period were similar across beneficiaries with high vs. low MSSP supply (differential trend, –$2

per beneficiary-quarter; 95% CI: -$4 to -$1; Figure A2 and Table A2 in Appendix A). In contrast, spending trends in pre-period differed substantially across MSSP beneficiaries and controls (differential trend, $18 per beneficiary-quarter; 95% CI: $16 to $19; Figure A3 and Table A2 in Appendix A), a violation of the parallel trends assumption.

Section D. Estimating the impact of the MSSP on spending and quality

We analyzed associations between changes in spending, hospital utilization, and quality and beneficiary attribution to an MSSP ACO. To identify the degree to which selection bias affects current conclusions about the impact of the MSSP on spending, we estimated 2 linear regression models for each outcome of interest. First, we estimated adjusted longitudinal models that accounted for market-level factors, secular trends, seasonality, and beneficiary characteristics. Specifically, for outcome Y (e.g., total spending), clinician i, attributed beneficiary j, in HRR k, quarter q, season r, and year t, we estimated the following model:

$$Y_{ijkqrt} = \beta_1(MSSP_{ijqt}) + \beta_2 X_{jt} + \mu_k + \omega_t + \theta_r + \zeta_{ijkqrt}$$

where $MSSP$ is 1 if the beneficiary was attributed to an MSSP ACO, 0 otherwise, with the quarterly subscript indicating that MSSP ACOs enter MSSP contracts within a given year; X is a vector of beneficiary characteristics (described in the main article); μ represents HRR fixed effects accounting for fixed sources of confounding across markets; ω represents year fixed effects accounting for common shocks across markets

in a given year; θ is a quarter indicator accounting for seasonal trends; and ζ is the idiosyncratic error term.

Second, we estimated instrumental variable models that are analogous to the adjusted longitudinal models above but used MSSP supply as an instrumental variable (defined in the section B of the eAppendix). By replacing clinicians' *observed* MSSP status with their *predicted probability* of MSSP participation, these two-stage least squares models may address residual sources of time-varying confounding that exist between clinician-beneficiary pairs who share similar observable characteristics and reside in the same area. For clinician i, attributed beneficiary j, in HRR k, quarter q, season r, and year t, we estimated the first-stage relationship between MSSP ACO participation and MSSP supply:

$$MSSP_{ijkqrt} = \delta_1(MSSP\_Supply_{ijqt}) + \delta_2 X_{jt} + \mu_k + \omega_t + \theta_r + \varepsilon_{ijkqrt}$$

where $MSSP$ is 1 if the beneficiary was attributed to an MSSP ACO, 0 otherwise; $MSSP\_Supply$ was the MSSP supply to which the clinician was exposed; X is a vector of beneficiary characteristics; μ represents HRR fixed effects; ω represents year fixed effects; θ is the quarter indicator; and ε is the idiosyncratic error term. In the second stage, we incorporate predicted MSSP participation from the first stage, $\widehat{MSSP}$, to identify the effect of the MSSP on outcome Y:

$$Y_{ijkqrt} = \gamma_1(\widehat{MSSP}_{ijqt}) + \gamma_2 X_{jt} + \mu_k + \omega_t + \theta_r + \varphi_{ijkqrt}$$

As before, we included a vector of beneficiary (X) covariates, HRR fixed effects (μ); year fixed effects (ω); quarter indicators (θ), and an idiosyncratic error term (φ). In this instrumental variable framework, the effect of the MSSP on spending and quality is identified under the assumption that within-market changes to MSSP supply are unrelated to potential changes in spending and quality after accounting for secular trends, seasonality, fixed differences across markets, and observed beneficiary characteristics. In contrast to standard longitudinal models, instrumental variable models capture treatment effects among "marginal" clinicians, i.e., those participating in the MSSP due to greater exposure to MSSP supply.

Section E. Evaluating compositional changes in the MSSP

To identify the degree to which changes in the composition of MSSP participants drives performance estimates, we evaluated changes in hip fracture and spending using across five sets of longitudinal models. First, we estimated a market-year fixed effects model that included fixed effects for each HRR* year, thus capturing market-specific spending trends. Second, we estimated a beneficiary fixed effects model that included beneficiary fixed effects and year fixed effects. Third, we estimated a model that included market-year fixed effects, ACO fixed effects, and year fixed effects. In the ACO fixed effects models, each ACO was defined as a fixed collection of all TINs that ultimately formed an MSSP ACO together; control TINs that never participated in the MSSP all received the same (control) indicator when defining ACO fixed effects. For

TINs that participated in multiple MSSP ACOs over the study period, we attributed TINs to the MSSP ACO that encompassed the plurality of the TIN's attributed beneficiaries during the post-MSSP period. To improve the computational feasibility of estimating models that included both ACO fixed effects and market-year fixed effects, we designated market-years as the panel variable (using Stata's –xtset– command) and included covariates that were mean-centered for each ACO (to capture ACO fixed effects).

Fourth, we estimated a clinician fixed effects model that included clinician fixed effects and year fixed effects. As in our instrumental variable analysis, beneficiaries were first attributed to TINs and then to clinicians within the attributed TIN based on the plurality of eligible primary care services (see section A). Finally, we estimated an adjusted longitudinal model that additionally adjusted for the share of beneficiaries in the county participating in MA (i.e., MA penetration). For all of the models described above, we included a time-varying indicator for whether the beneficiary was attributed to an MSSP ACO in a given quarter, a quarter indicator (to capture seasonal trends), and the beneficiary characteristics described in the main article. Models including beneficiary fixed effects omitted time-invariant characteristics (i.e., race/ethnicity, disability as the original reason for Medicare entitlement, end-stage renal disease as the original reason for Medicare entitlement). All analyses specified robust standard errors to account for clustering at the market level.

Section F. Instrumental variable sensitivity analyses

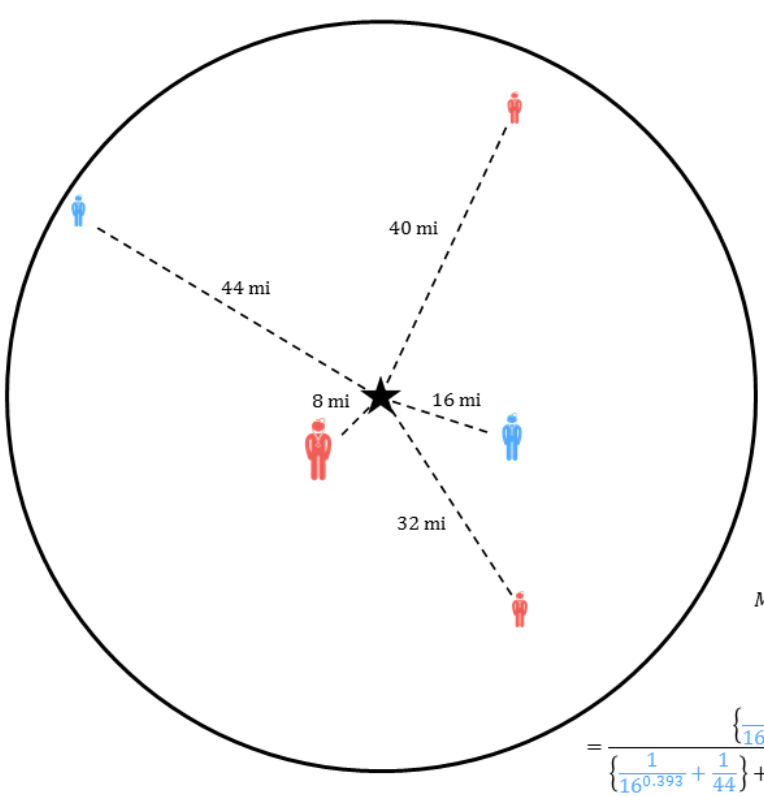
We performed several sensitivity analyses to test the robustness of our instrumental variable model. First, in order to address potential area-level relationships between participation in managed care and MSSP ACOs, we estimated an instrumental variable model that adjusted for county-level MA penetration, a proxy for managed care. Second, we estimated an instrumental variable model that included market-year fixed effects instead of market fixed effects and year fixed effects. This accounted for potential market-specific trends in spending or quality that may be associated with market-level changes in MSSP supply. Third, we estimated an instrumental variable model that included both market-year fixed effects and ACO fixed effects (using covariates that were mean-centered for each ACO). This accounted for potential small-area variations in practices patterns or beneficiaries that may confound the relationship between MSSP supply and performance that may exist at a more local level not captured by market fixed effects. Finally, we estimated models that were analogous to our base instrumental variable model but that used MSSP supply restricted to within 10, 25, 50, 75, or 150 miles from the clinician's practice, with or without discounting the supply of providers relatively further away within a given radius (defined in the section B of the eAppendix). For each instrumental variable model described above, we estimated the impact of the MSSP on spending and hip fracture (the falsification outcome). Results were similar to those from the main instrumental variable model (Table A3 and Figure A6).

G. Supplemental analyses of MSSP exit and entry. To determine whether MSSP ACOs strategically recruit or prune clinicians according to prior spending performance, we tested whether a clinician's spending performance predicted the probability of the clinician entering or exiting an existing MSSP ACO in the subsequent year. We determined clinician participation in the MSSP by linking clinicians via their NPI to the provider-level SSP RIF. In the main analysis, prior spending performance was defined as the clinician's average Medicare spending per beneficiary per year for the three years prior to MSSP exit or entry determination. In these analyses, we directly attributed beneficiaries to clinicians (vs. TINs) from whom they received the plurality of eligible evaluation and management services in the outpatient setting, as defined by the MSSP.[57]

In our models, we estimated the probability of entering or exiting the MSSP as a function of market fixed effects, year fixed effects, average spending, and average beneficiary characteristics of the clinician's panel. These characteristics included average beneficiary age, sex, race/ethnicity (white, black, Hispanic, or other/unknown), disability, end-stage renal disease, dual-eligibility for Medicaid (months of enrollment), the CMS Hierarchical Condition Category (HCC) risk score, and area-level poverty (proportion below federal poverty level) and education (proportion graduated from high school, college). We included quadratic spending terms to allow for any potential non-linearities in the relationship between spending and MSSP exit or entry.

Analyses of MSSP exit were restricted to clinicians participating in the MSSP during the year prior to analysis. Analyses of MSSP entry were restricted to clinicians not participating in the MSSP during the year prior to analysis and did not include ACO formation, i.e., participation in an ACO's first contract year. Both sets of analyses were restricted to ACOs entering MSSP contracts in 2012 or 2013, as the use of 2012-2014 MSSP data precluded the ability to observe exit or entry into ACOs formed in 2014. Results were robust when using clinician's average spending in the prior year (Figure A8 in Appendix A, Panel A) or prior two years (Panel B), or when not adjusting for the average characteristics of the clinician's attributed beneficiaries (Panel C). We also analyzed MSSP exit at the level of the provider group (i.e., TIN), finding no consistent relationship between group's prior spending performance and subsequent MSSP exit (Figure A9 in Appendix A).

Figure A1. Schematic of the MSSP supply instrumental variable



$$MSSP\ Supply = \frac{\sum_{k=1}^{N} \frac{1}{d_{i,k}^{\beta}}}{\sum_{k=1}^{N} \frac{1}{d_{i,k}^{\beta}} + \sum_{j=1}^{N} \frac{1}{d_{i,j}^{\beta}}} \times 100\%$$

$$= \frac{\left\{\frac{1}{16^{0.393}} + \frac{1}{44^{0.393}}\right\}}{\left\{\frac{1}{16^{0.393}} + \frac{1}{44}\right\} + \left\{\frac{1}{8^{0.393}} + \frac{1}{32^{0.393}} + \frac{1}{40^{0.393}}\right\}} \times 100\% = 35.9\%$$

Figure A2. Pre-period trends in spending across the instrumental variable

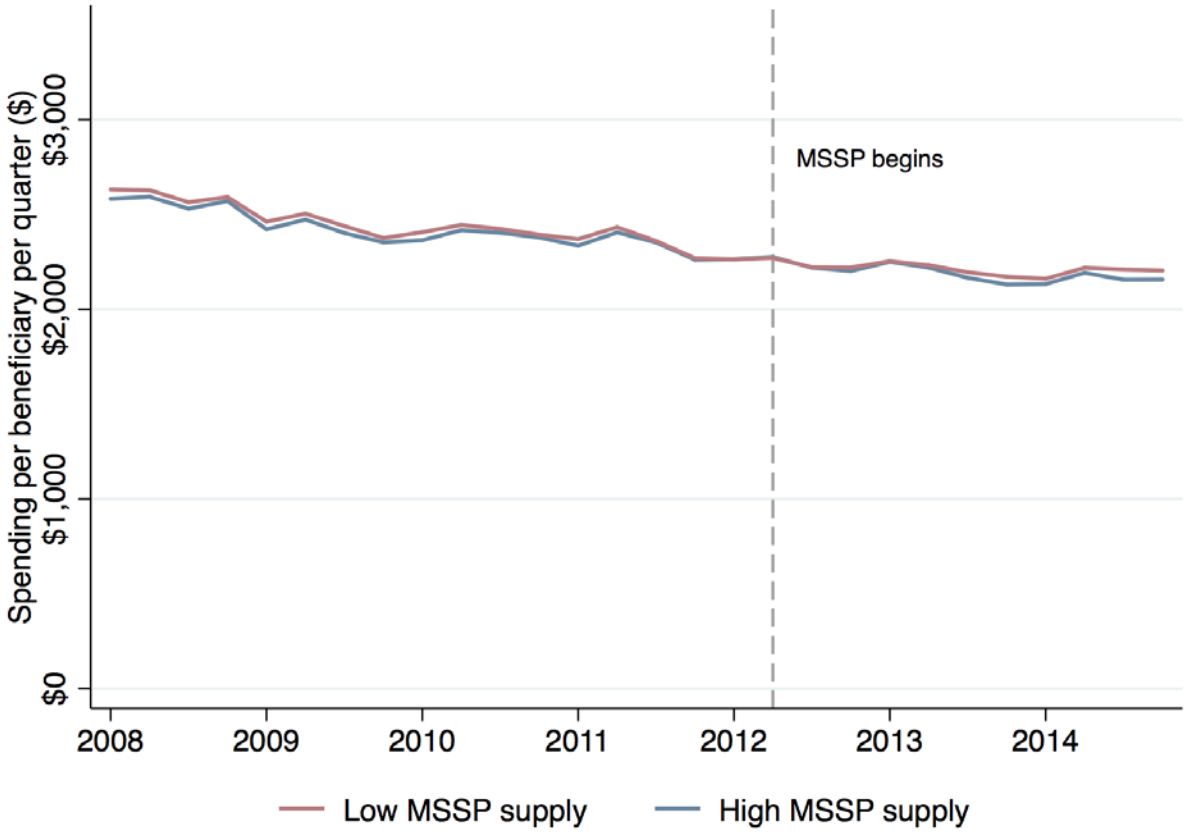


Figure A3. Pre-period trends in spending across observed MSSP participation status

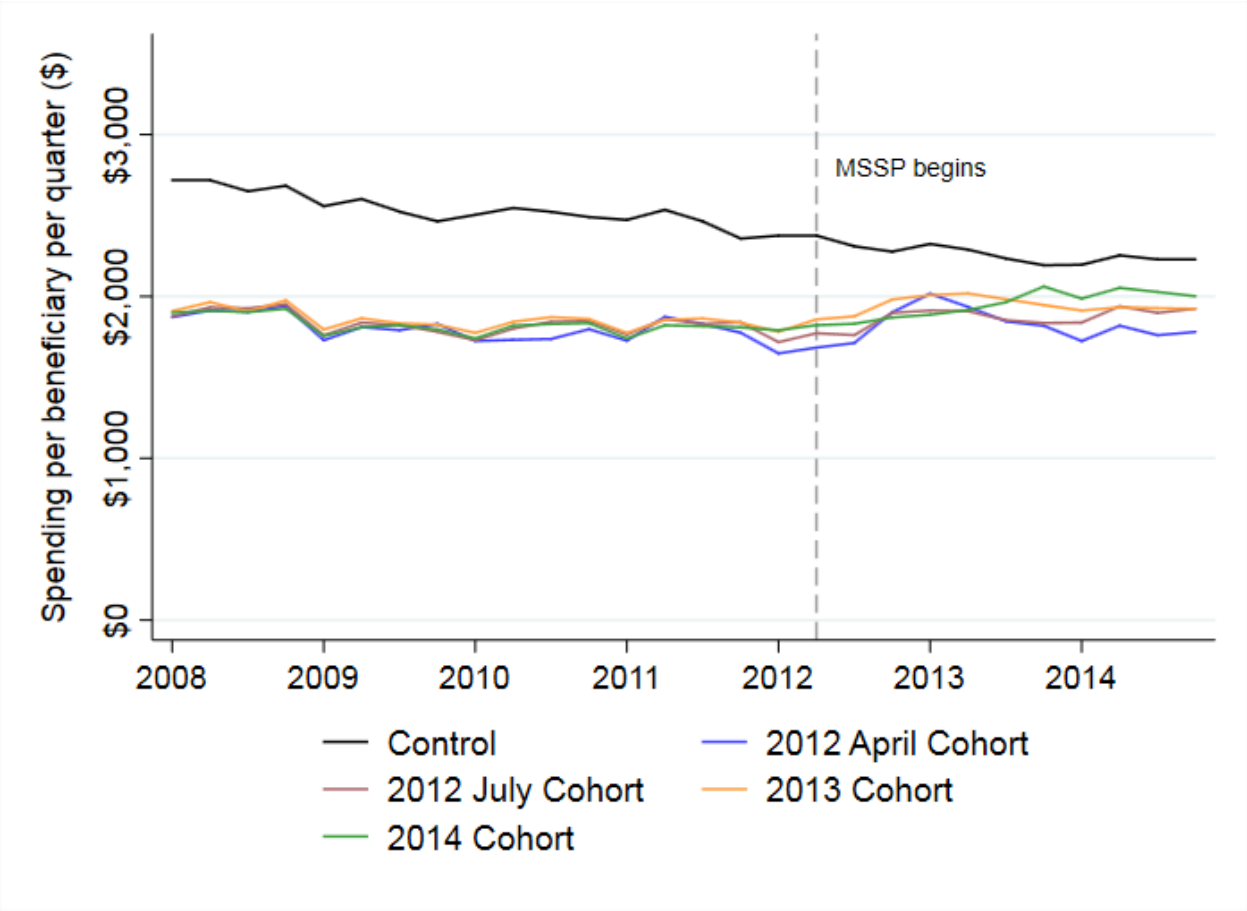


Figure A4. Changes in total Medicare spending according to ACOs' year of entry into the MSSP

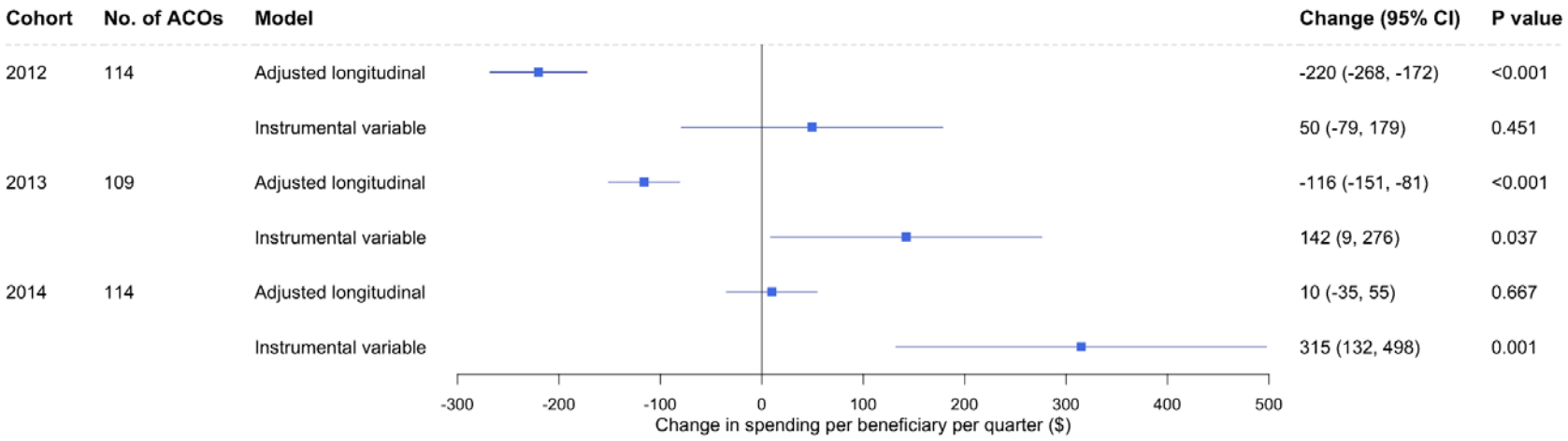


Figure A5. Association between county-level MSSP penetration and Medicare Advantage penetration

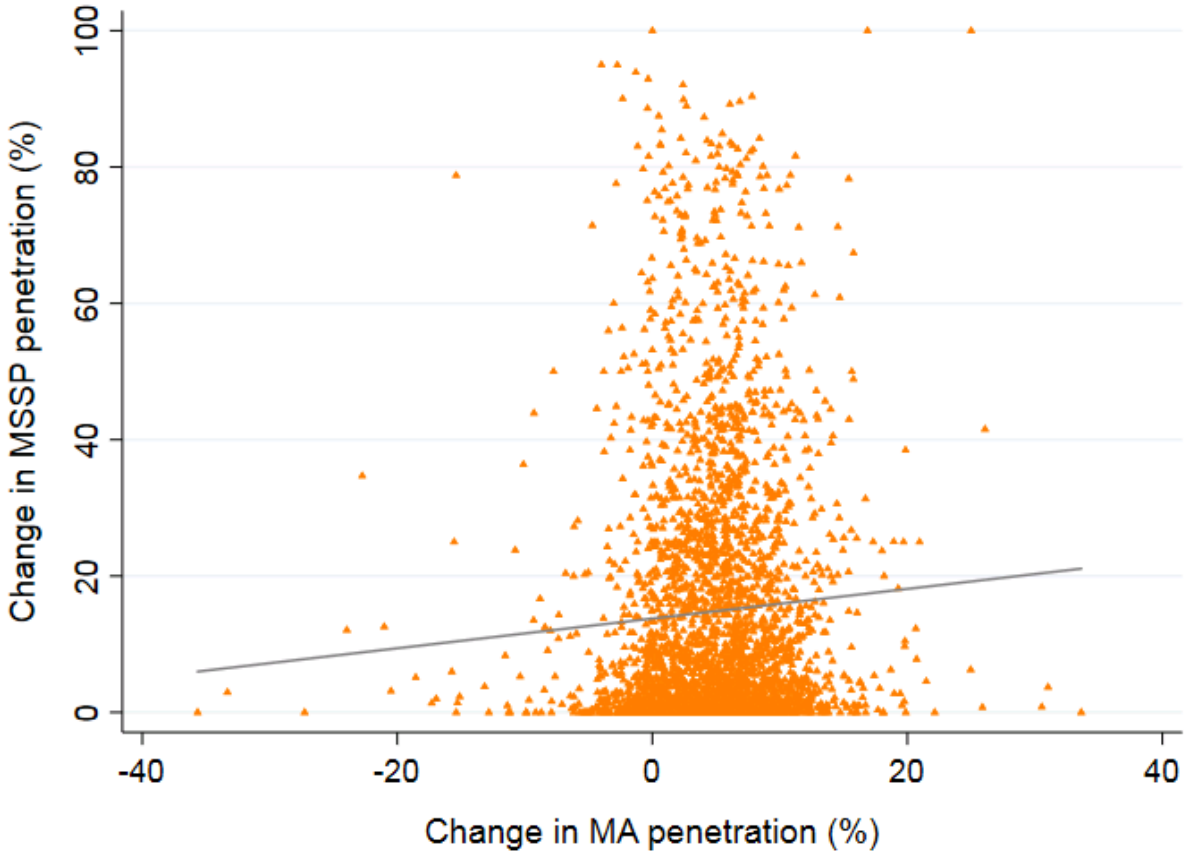


Figure A6. Robustness tests of MSSP supply as an instrumental variable

Panel A. Changes in hip fracture

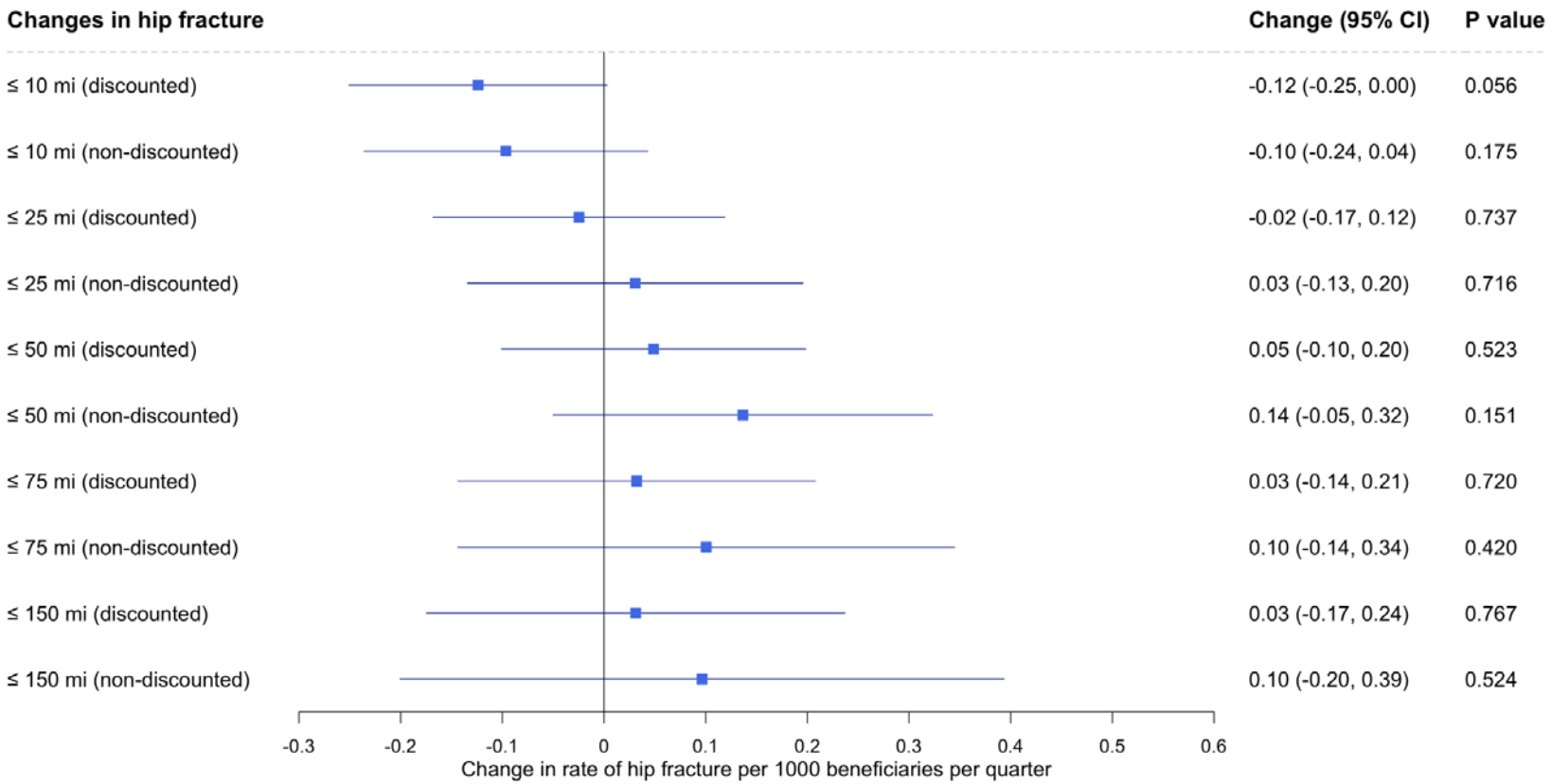


Figure A6. Robustness tests of MSSP supply as an instrumental variable

Panel B. Changes in total spending

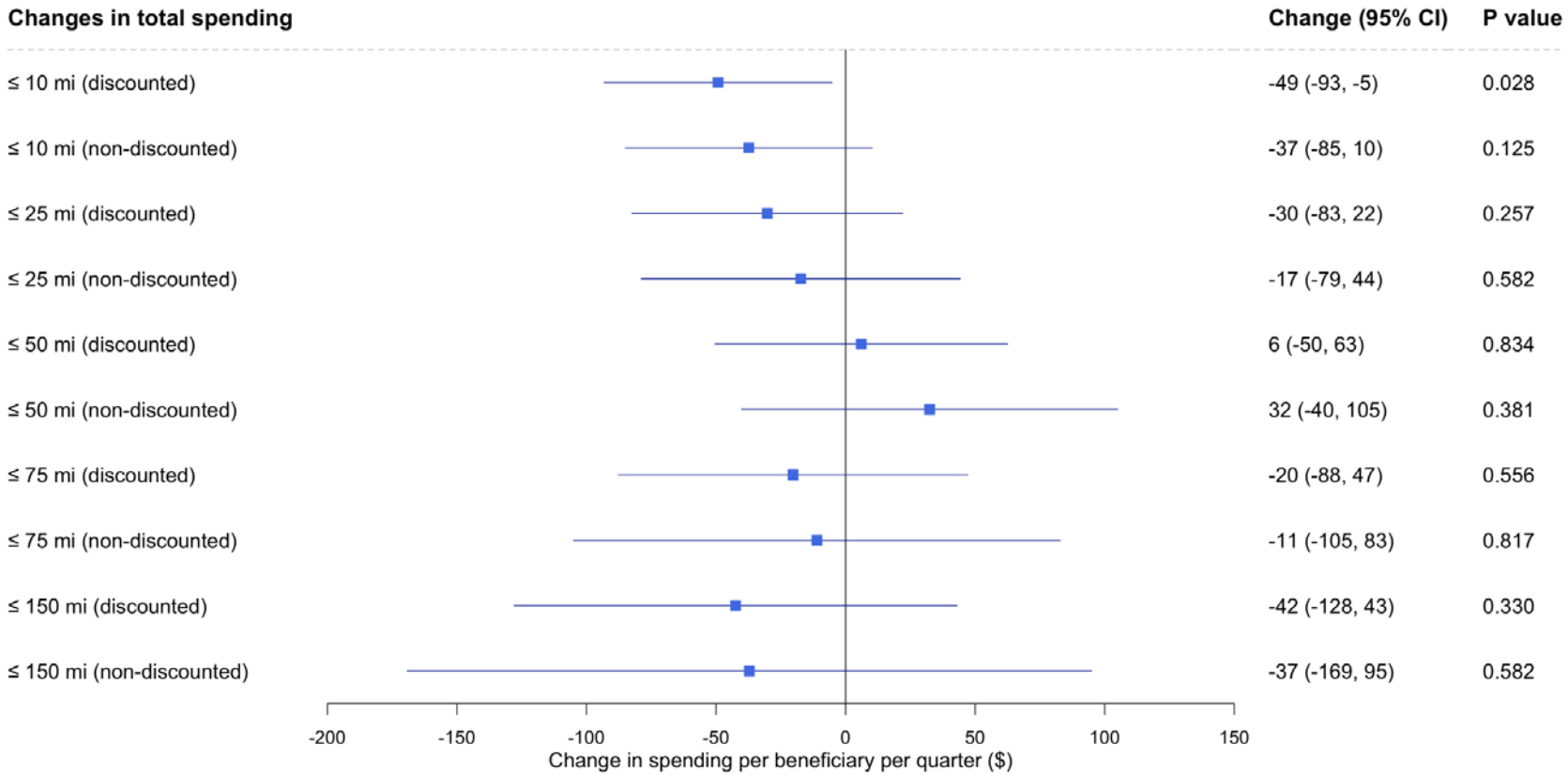


Figure A7. Changes in hospital use for beneficiaries attributed to MSSP ACOs vs. controls

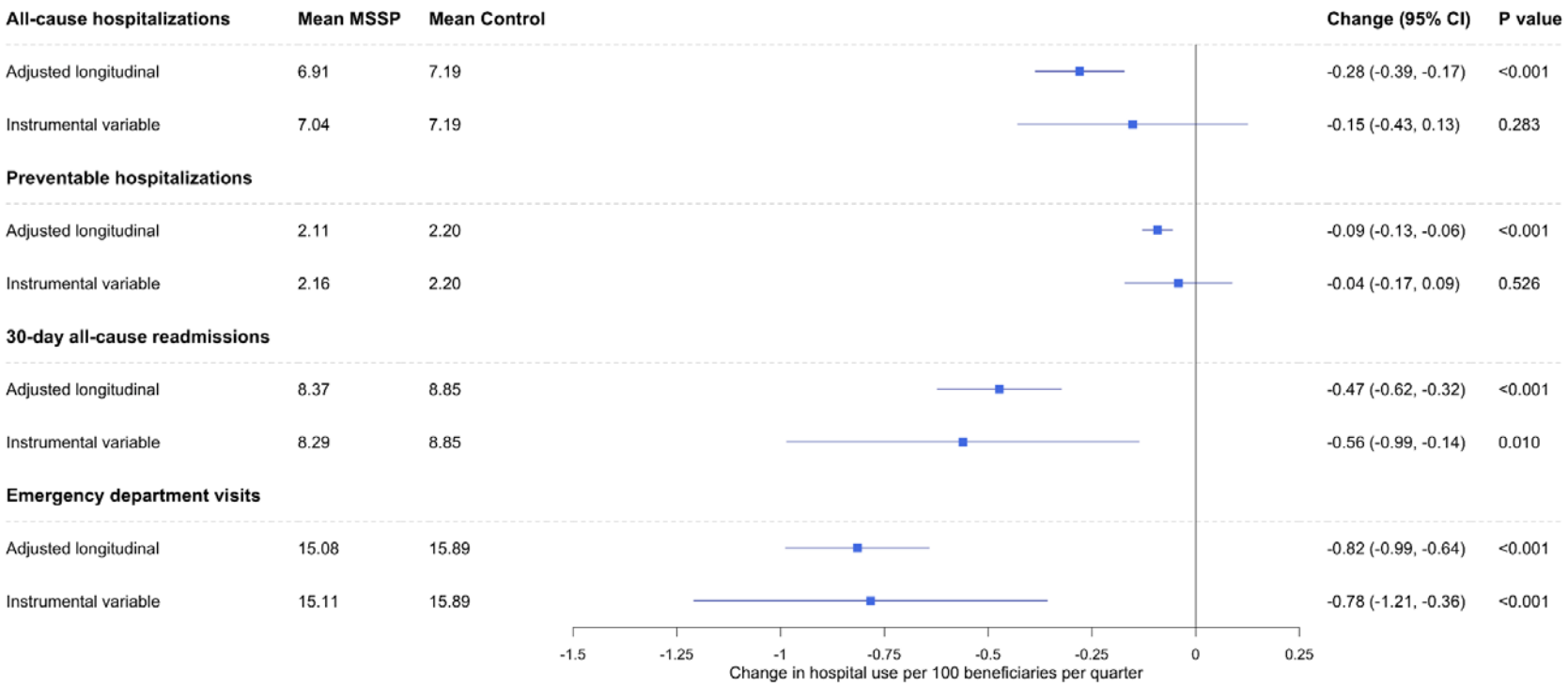


Figure A8. Sensitivity analyses of association between clinician spending performance and probability of exiting an MSSP ACO

Panel A. Clinician entry and exit as a function of average spending in prior one year

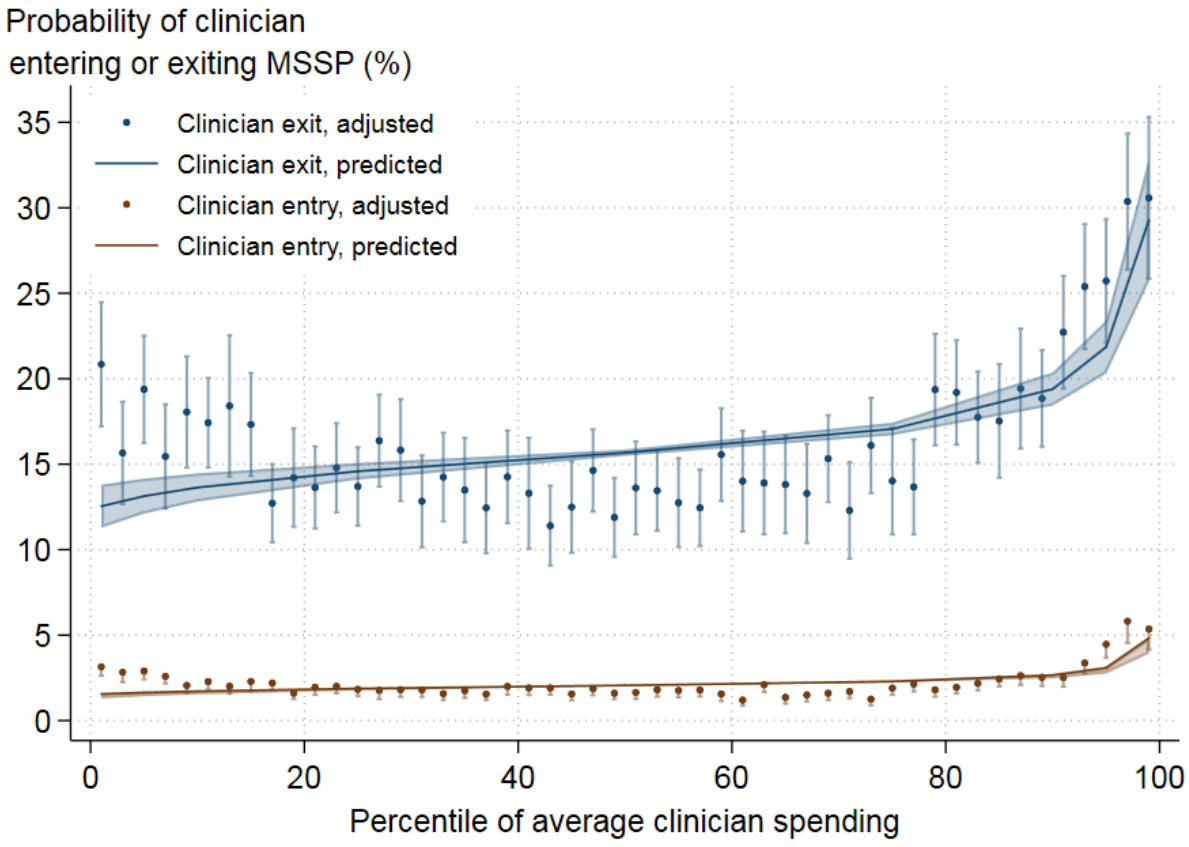


Figure A8.

Panel B. Clinician entry and exit as a function of average spending in prior two years

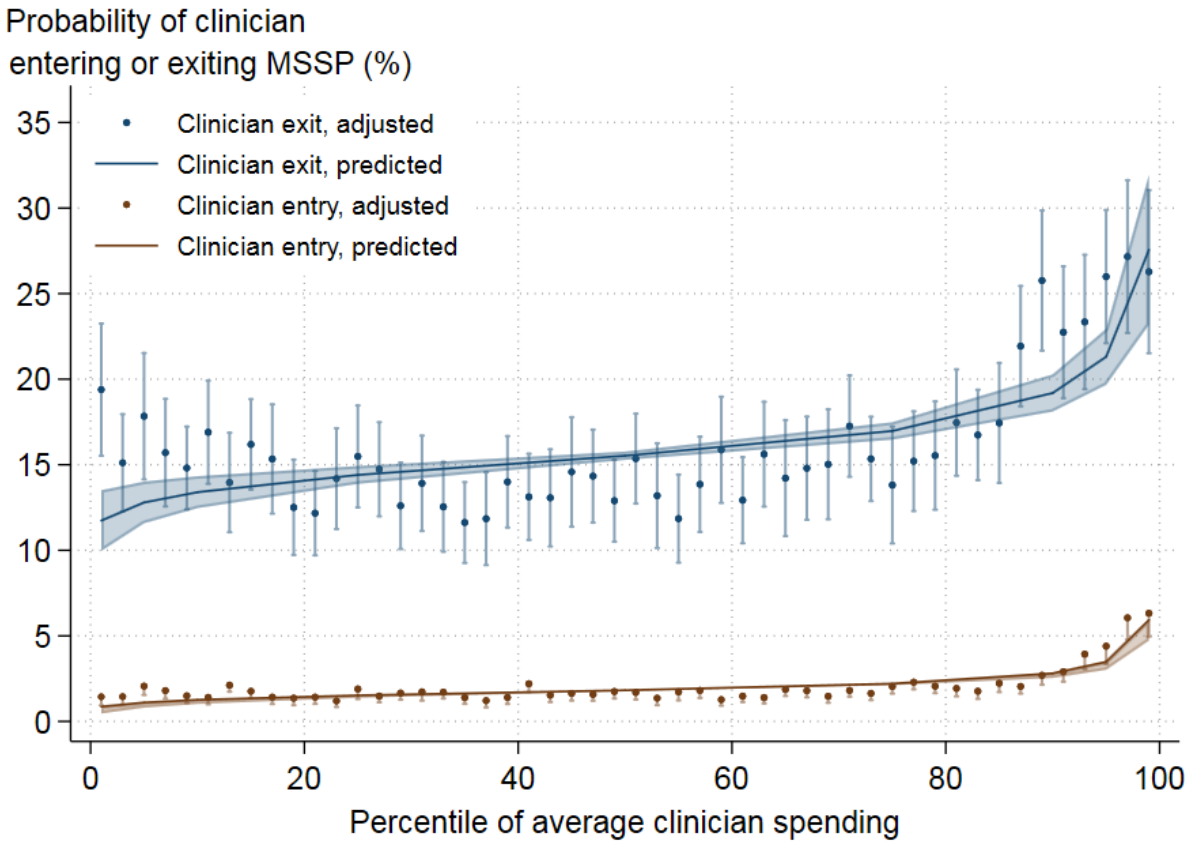


Figure A8.

Panel C. Clinician entry and exit without adjustment for average beneficiary characteristics

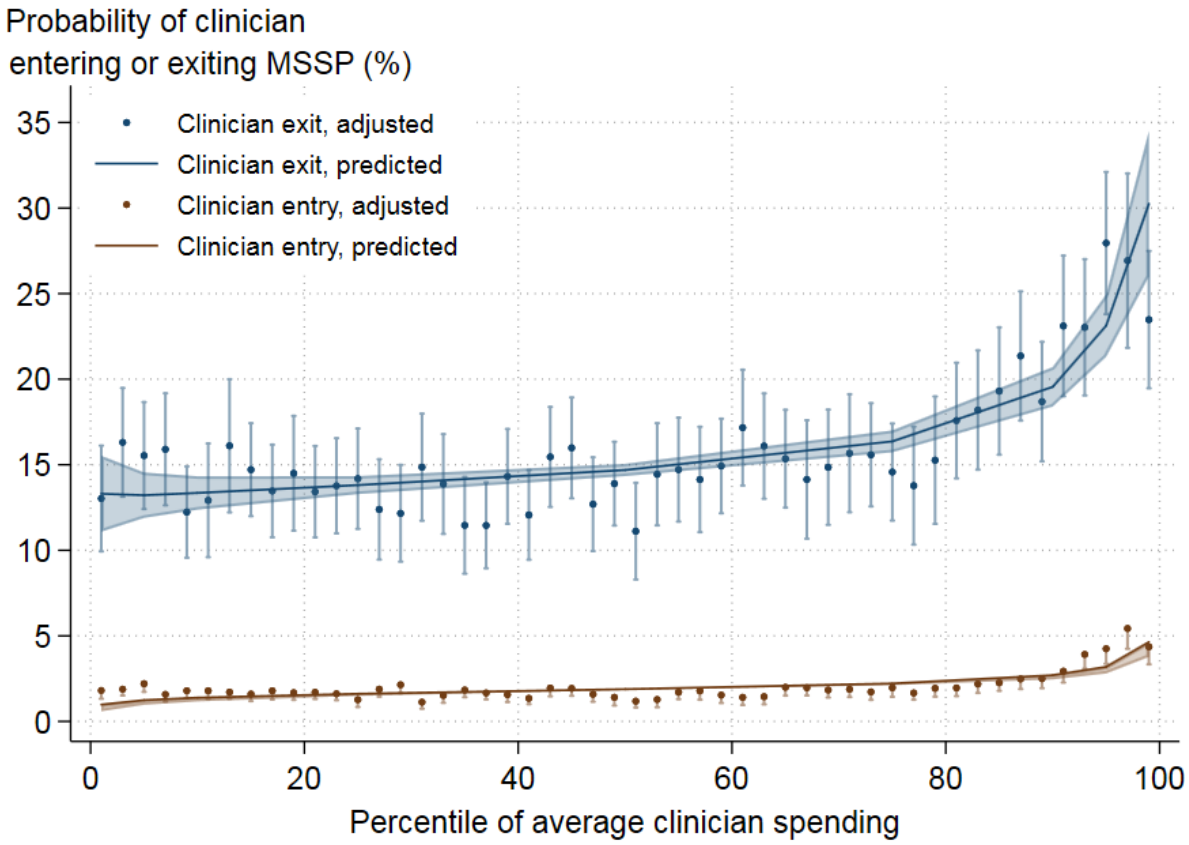


Figure A9. Association between provider group spending performance and probability of exiting or entering an MSSP ACO

Panel A. Group exiting an MSSP ACO

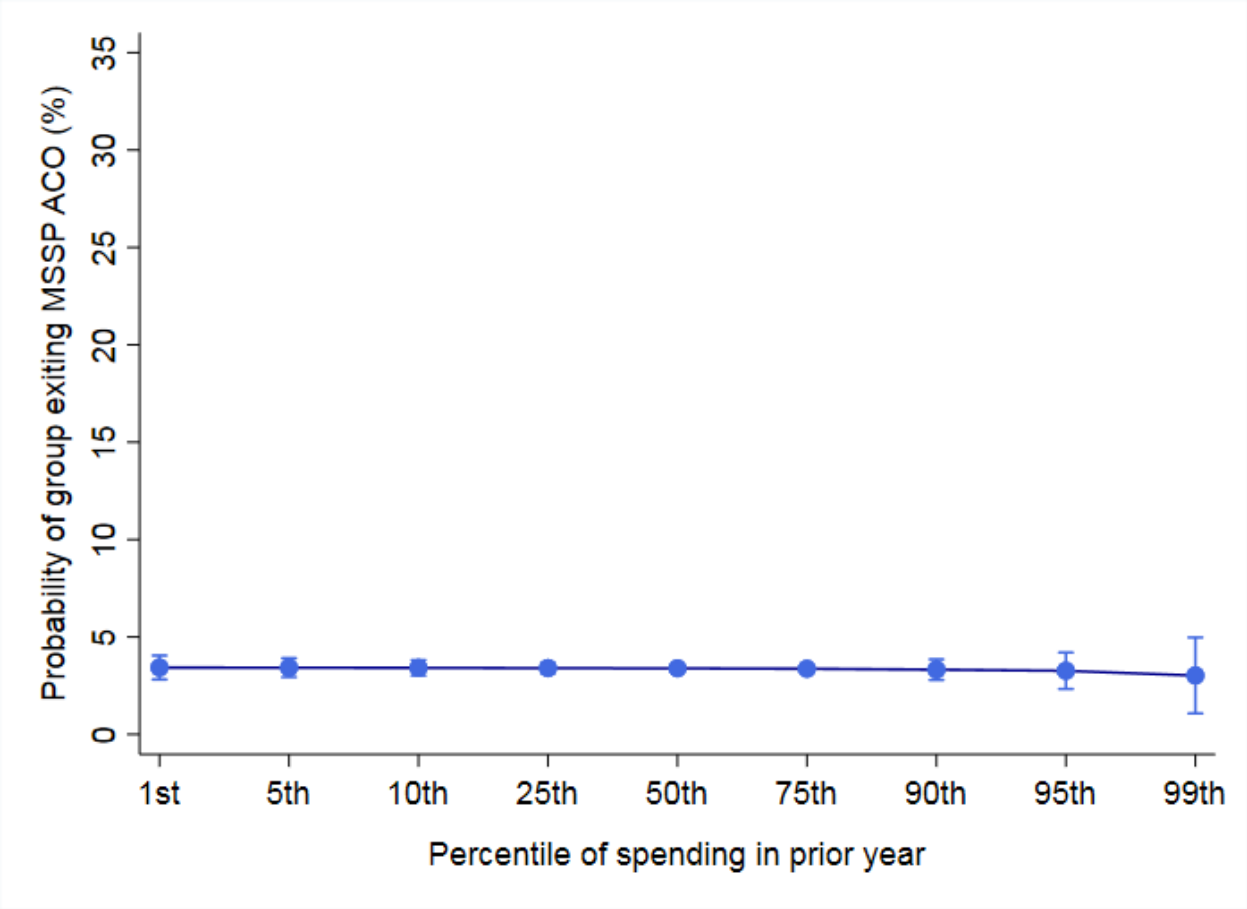


Figure A9.

Panel B. Group entering an MSSP ACO

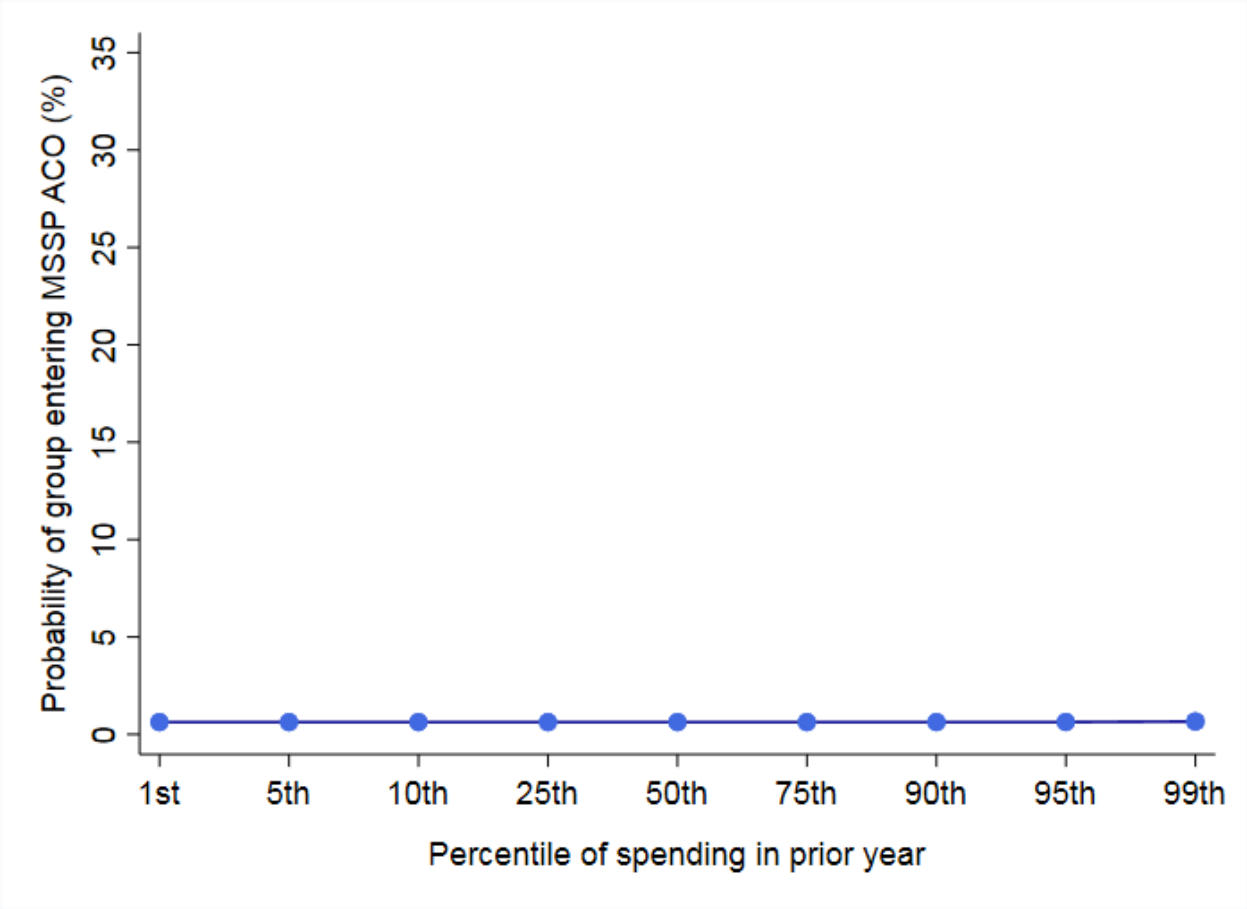


Table A1. Non-linear estimation of discount factor used in MSSP supply

	Rural-Urban Commuting Area categorization			
Probability of two clinicians belonging to same MSSP ACO	**Urban**	**Large rural cities/towns**	**Small rural towns**	**Isolated**
Discount factor β (95% CI)	0.263 (0.262, 0.263)	0.158 (0.156, 0.160)	0.171 (0.167, 0.175)	0.137 (0.132, 0.143)
Sample size	2,668,375	69,273	20,634	9,963

Table A2. Pre-period trends in spending across observed MSSP status

Observed MSSP status vs. instrumental variable	Difference in spending levels in pre-period, $ (95% CI)	Difference in spending trends in pre-period, $ per quarter (95% CI)	P value
MSSP participants vs. non-participants (observed MSSP status	-642 (-670, -614)	18 (16, 19)	<0.001
High vs. low MSSP supply (instrumental variable)	48 (19, 76)	-2 (-4, -1)	0.001

Table A3. Formal test of statistical differences between instrumental variable and adjusted longitudinal model

Study outcome	Difference-in-Sargan, P value[a]
Total Medicare spending	<0.001
Hip fracture	0.001
Glycated hemoglobin testing	0.021
LDL cholesterol testing	0.021
Diabetic retinal examination	0.015
All three diabetes measures	0.019
Mammography	0.007
All-cause hospitalizations	0.335
Preventable hospitalizations	0.404
Emergency department visits	0.873

[a] Statistical difference between estimates from the instrumental variable model and adjusted longitudinal model was tested via the difference-in-Sargan statistic.[75]

Table A4. Sensitivity analyses of adjusted longitudinal and instrumental variable models

Hip fracture	**Change (95% CI)**	***P* value**
Adjusted longitudinal + MA penetration	-0.28 (-0.36, -0.19)	<0.001
Instrumental variable + MA penetration	0.11 (-0.06, 0.27)	0.210
Instrumental variable + market-year fixed effects	-0.17 (-0.43, 0.09)	0.205
Instrumental variable + market-year + ACO fixed effects	0.04 (-0.32, 0.39)	0.840
Total spending		
Adjusted longitudinal + MA penetration	-116 (-148, -83)	<0.001
Instrumental variable + MA penetration	12 (-44, 68)	0.680
Instrumental variable + market-year fixed effects	32 (-44, 109)	0.406
Instrumental variable + market-year + ACO fixed effects	-19 (-149, 111)	0.776

Table A5. Estimates of the association between prior spending and probability of joining or exiting an MSSP ACO

Average spending in prior three years	Probability of clinician entering MSSP ACO, % (95% CI)	Risk difference*	*P* value*
1st percentile	0.6 (0.2, 1.1)	-1.2	<0.001
5th percentile	1.0 (0.6, 1.3)	-0.9	<0.001
10th percentile	1.2 (0.9, 1.4)	-0.7	<0.001
25th percentile	1.5 (1.3, 1.6)	-0.4	<0.001
50th percentile	1.8 (1.8, 1.9)	-- †	-- †
75th percentile	2.3 (2.2, 2.4)	0.5	<0.001
90th percentile	3.0 (2.7, 3.3)	1.1	<0.001
95th percentile	3.6 (3.1, 4.1)	1.8	<0.001
99th percentile	5.6 (4.5, 6.7)	3.7	<0.001
Average spending in prior three years	**Probability of clinician exiting MSSP ACO, % (95% CI)**	**Risk difference***	***P* value***
1st percentile	13.2 (11.0, 15.3)	-1.5	0.176
5th percentile	13.2 (11.8, 14.5)	-1.6	0.029
10th percentile	13.3 (12.2, 14.4)	-1.4	0.008
25th percentile	13.8 (13.2, 14.5)	-0.9	0.001
50th percentile	14.7 (14.3, 15.1)	-- †	-- †
75th percentile	16.4 (15.8, 17.0)	1.7	<0.001
90th percentile	19.6 (18.2, 20.9)	4.9	<0.001
95th percentile	23.1 (20.8, 25.4)	8.4	<0.001
99th percentile	30.0 (24.9, 35.1)	15.3	<0.001

* Risk differences were calculated as the difference between the probability of exit or entry for a given percentile of spending (e.g., 95th) to the probability of exit or entry for the 50th percentile of spending.
‡ Clinicians in the 50th percentile of average spending was the referent.
CI is Confidence Interval.

Table A6. Simulated changes in average MSSP beneficiary spending with attrition of high-cost clinicians from MSSP ACOs (2013-2014)

Sample of MSSP ACO participants (2013-2014)	Average MSSP ACO spending per beneficiary-year ($)	Difference in calculated average MSSP ACO spending (%)	Sample size (clinician-years)	Sample size (beneficiary-years)
All clinicians	8,278	--	37,243	3,803,968
Excluding clinicians with ≥ 99th of average spending	8,229	-0.6	36,871	3,800,764
Excluding clinicians with ≥ 95th of average spending	8,024	-3.1	35,381	3,761,548

Table A7. Simulated changes in average MSSP beneficiary spending when including clinicians exiting MSSP and excluding clinicians entering the MSSP (2013-2014)

Sample of MSSP ACO participants (2013-2014)	Average MSSP ACO spending per beneficiary-year ($)	Difference in calculated average MSSP ACO spending (%)	Sample size (clinician-years)	Sample size (beneficiary-years)
Observed sample of MSSP clinicians	8,278	--	37,243	3,803,968
Simulated sample including clinicians exiting MSSP and clinicians entering MSSP	8,370	1.1	44,027	4,341,368

Supplemental Figure and Table Notes

Figure A2.
Abbreviations: MSSP, Medicare Shared Savings Program; CI, confidence interval. High vs. low MSSP supply was defined by whether the attributed clinician's median MSSP value was above (1=high) or equal to or below (0=low) the median MSSP supply value for the HRR over the entire post-period, which was designated independently for each HRR as beginning with the first quarter in which the HRR contained at least one clinician participating in the MSSP. Adjusted spending was estimated using a model that included market fixed effects, year fixed effects, seasonality, beneficiary characteristics, an indicator for each quarter, an indicator for high MSSP supply, and an interaction term between high MSSP supply and the quarter of interest, but did not impose an underlying time trend on the adjusted spending data. From these models, Stata's –margins- command was used to estimate predicted spending for beneficiaries with high- vs. low-MSSP supply in each quarter.

Figure A3.
Abbreviations: MSSP, Medicare Shared Savings Program; CI, confidence interval. Adjusted spending trends were estimated using models that included market fixed effects, year fixed effects, seasonality, beneficiary characteristics, an indicator for each time period, an indicator for the beneficiary's ultimate MSSP cohort status (0=never in an MSSP ACO, 1=2012 April cohort, 2=2012 July cohort, 3=2013 cohort, 4=2014 cohort), and a fully-interacted term between the MSSP cohort indicator and the quarter of interest. From these models, Stata's –margins- command was used to estimate predicted spending for beneficiaries in each MSSP cohort in each quarter.

Figure A4.
Abbreviations: MSSP, Medicare Shared Savings Program; CI, confidence interval.

Figure A5.
Abbreviations: MSSP, Medicare Shared Savings Program; MA, Medicare Advantage. In a supplemental analysis, changes in county-level MA penetration were estimated as a linear function of year fixed effects and county fixed effects. For purposes of graphical representation, changes in county-level MSSP and MA penetration were calculated using a first-differences model that used changes between the pre-MSSP period (2011) and the post-MSSP period (2014).

Figure A6.
Abbreviations: MSSP, Medicare Shared Savings Program; CI, confidence interval.

Figure A7
Abbreviations: MSSP, Medicare Shared Savings Program; CI, confidence interval.

Mean outcomes for control beneficiaries were calculated as the intercept coefficient from adjusted longitudinal models with mean-centered covariates. Mean outcomes for MSSP beneficiaries were calculated as the sum of the control mean and the estimated effect of the MSSP on the outcome. All-cause hospitalization was defined as any inpatient stay. Preventable hospitalization was defined as hospitalization for any 1 of 11 ambulatory care-sensitive conditions (ACSCs), as specified by the Agency for Healthcare Quality and Research (see Methods of Appendix A). All-cause 30-day readmission was defined according to MSSP specifications and limited to hospitalized beneficiaries (n= 5,902,749 beneficiary-quarters).

.

Figure A8.

Abbreviations: MSSP, Medicare Shared Savings Program; ACO, accountable care organization.

Spending in the prior 2 years (Panel A) and 3 years (Panel B) was defined as a rolling average of each clinician's average Medicare care spending per beneficiary in the 2 or 3 years prior to analysis, respectively. The probability of a clinician exiting the MSSP was estimated as a linear function of market fixed effects, year fixed effects, and prior spending. Adjusted analyses (Panel C) included the average beneficiary characteristics of the clinician's panel, including average beneficiary age, sex, race/ethnicity (white, black, Hispanic, or other/unknown), disability, end-stage renal disease, dual-eligibility for Medicaid (months of enrollment), the CMS Hierarchical Condition Category (HCC) risk score, and area-level poverty (proportion below federal poverty level) and education (proportion graduated from high school, college). A quadratic spending term was included to allow for any potential non-linearities in the effect of spending performance on MSSP participation. Analyses of MSSP exit were restricted to clinicians participating in the MSSP during the year prior to analysis and to ACOs entering MSSP contracts in 2012 or 2013.

Figure A9.

Groups were defined by Taxpayer Identification Number (TIN). Prior spending performance was defined as each group's average Medicare spending per beneficiary per year in the one year prior to MSSP exit or entry determination. The probability of a group exiting or entering the MSSP was estimated as a linear function of market fixed effects, year fixed effects, and prior spending. A quadratic spending term was included to allow for any potential non-linearities in the effect of spending performance on MSSP participation. Analyses of MSSP exit (Panel A; N = 6,199 group-years) were restricted to groups participating in the MSSP during the year prior to analysis. Analyses of MSSP entry (Panel B; N = 89,026 group-years) were restricted to groups not participating in the MSSP during the year prior to analysis and did not include ACO formation, i.e., participation in an ACO's first contract year. Both sets of analyses were restricted to ACOs entering MSSP contracts in 2012 or 2013.

Appendix B. Risk adjustment in Medicare ACO program deters coding increases but may lead ACOs to drop high-risk beneficiaries

Table of Contents

Supplemental Methods

Figure B1. CONSORT diagram

Figure B2. Sensitivity analyses of relationship between beneficiary attribution to the Medicare Shared Savings Program (MSSP) and change in risk score

Figure B3. Relationship between beneficiary risk score and beneficiary exit or entry in the Medicare Shared Savings Program (MSSP) among beneficiaries attributed via outpatient claims submitted by primary care providers

Figure B4. Relationship between clinician's average patient panel risk score and clinician entry and exit to and from the Medicare Shared Savings Program (MSSP)

Figure B5. Trends in risk score across beneficiaries in the Medicare Shared Savings Program (MSSP) and controls (2008-2014)

Figure B6. Event study of change in risk score before and after attribution to the Medicare Shared Savings Program (MSSP)

Table B1. Characteristics of Medicare beneficiaries in the analytic sample vs. excluded sample (2008-2011).

Table B2. Relationship between beneficiary risk score and beneficiary exit or entry in the Medicare Shared Savings Program (MSSP)

Figure B7. Heterogeneity in relationship between beneficiary risk score and exit from Medicare Shared Savings Program (MSSP)

Figure B8. Relationship between beneficiary's average risk score in prior two or three years and entry and exit in the Medicare Shared Savings Program (MSSP)

Table B3. Growth in risk score and beneficiary entry and exit to and from the Medicare Shared Savings Program (MSSP) between 2012 and 2014

Supplemental Methods

Analysis of clinician average patient panel risk and entry and exit in the MSSP. We evaluated the relationship between the clinician's patient panel and entry and exit in the MSSP in the following manner. First, we defined clinician participation in the MSSP using CMS' Shared Savings Program Provider-Level Research Identifiable File, which lists the ACOs, provider groups, and clinicians participating in the MSSP. Second, defined each clinician's patient panel, adapting MSSP attribution specifications to directly beneficiaries to the clinician from whom they received the plurality of eligible evaluation and management services (Healthcare Common Procedure Coding System codes: 99201 through 99215; 99304 through 99350; G0402; G0438; and G0439.)[104] Finally, we estimated the probability of clinician entry or exit as a linear function of the clinician's patient panel's average risk score, year fixed effects, market fixed effects, and average patient characteristics (age, sex, race/ethnicity, disability, end-stage renal disease, dual-eligibility for Medicaid, and area-level poverty, education, and Medicare Advantage penetration).

Analyses of MSSP entry were restricted to clinicians not participating in the MSSP during the year prior to analysis and did not include ACO formation, i.e., participation in an ACO's first contract year. Analyses of MSSP exit were restricted to clinicians participating in the MSSP during the year prior to analysis. Both sets of analyses were restricted to ACOs entering MSSP contracts in 2012 or 2013, as 2012-2014 MSSP data could not be used to determine clinician exit or entry for ACOs formed in 2014. These analyses also required that clinicians have beneficiaries attributed to their patient panel

in the analytic year(s), thus excluding clinicians who exit the MSSP due to retirement or death, for example. To improve statistical reliability of clinician-level estimates, our main specification used average patient characteristics in the three years prior to determination of MSSP entry or exit. We obtained similar results when using average patient characteristics in the prior year, average patient characteristics in the prior two years, or when weighting by the size of clinician's patient panel.

Figure B1. CONSORT diagram

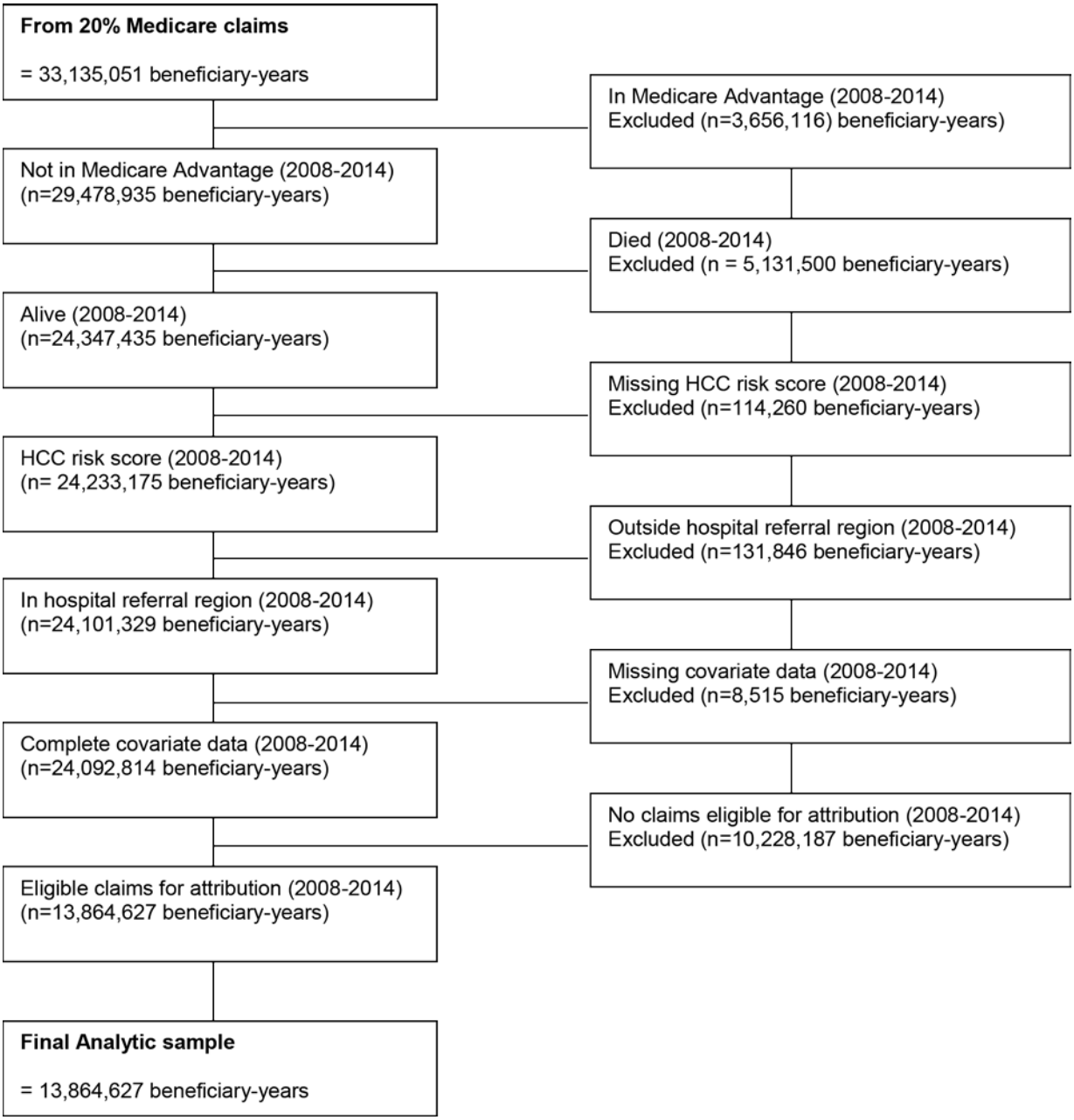
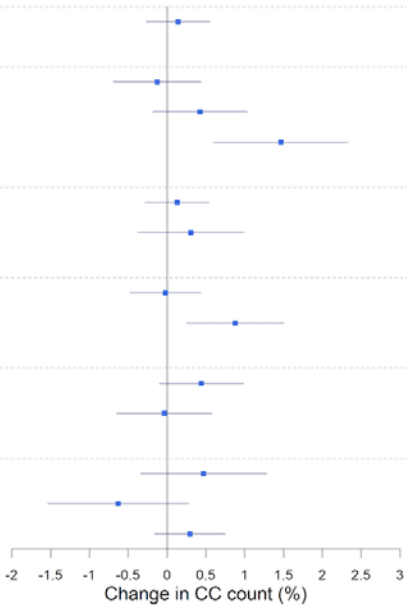


Figure B2. Sensitivity analyses of relationship between beneficiary attribution to the Medicare Shared Savings Program (MSSP) and change in risk score

Panel A. Change in count of Condition Categories (CC)

Characteristic	Mean Controls	Mean MSSP	Change in risk score (95% CI)	Change relative to base (95% CI)
Average change	2.057	2.060	0.1 (-0.3, 0.6)	N/A
MSSP Cohort				
2012	2.081	2.078	-0.1 (-0.7, 0.4)	--
2013	2.010	2.019	0.4 (-0.2, 1.0)	0.4 (-0.4, 1.3)
2014	2.044	2.074	1.5 (0.6, 2.3)	2.0 (0.9, 3.0)
Earned shared savings				
No	2.024	2.026	0.1 (-0.3, 0.5)	--
Yes	2.123	2.130	0.3 (-0.4, 1.0)	0.1 (-0.6, 0.8)
High Medicare Advantage area				
No	2.061	2.060	0.0 (-0.5, 0.4)	--
Yes	2.042	2.060	0.9 (0.3, 1.5)	1.2 (0.5, 1.8)
Commercial ACO contract				
No	2.072	2.081	0.4 (-0.1, 1.0)	--
Yes	2.038	2.038	0.0 (-0.6, 0.6)	-0.6 (-1.4, 0.3)
Organizational structure				
Physician-led	2.097	2.107	0.5 (-0.3, 1.3)	--
Hospital-led	2.040	2.027	-0.6 (-1.5, 0.3)	-1.2 (-2.4, 0.0)
Physician-hospital partnership	2.032	2.038	0.3 (-0.2, 0.7)	-0.2 (-1.2, 0.7)



Panel B. Change in risk score adjusted for total price-standardized spending

Characteristic	Mean Controls	Mean MSSP	Change in risk score (95% CI)	Change relative to base (95% CI)
Average change	1.158	1.160	0.1 (-0.1, 0.4)	N/A
MSSP Cohort				
2012	1.165	1.166	0.0 (-0.3, 0.4)	--
2013	1.147	1.150	0.2 (-0.1, 0.6)	0.2 (-0.3, 0.7)
2014	1.157	1.162	0.4 (-0.1, 0.8)	0.4 (-0.2, 1.0)
Earned shared savings				
No	1.152	1.152	0.0 (-0.2, 0.3)	--
Yes	1.172	1.176	0.3 (-0.1, 0.7)	0.3 (-0.1, 0.7)
High Medicare Advantage area				
No	1.159	1.159	0.0 (-0.3, 0.3)	--
Yes	1.156	1.163	0.6 (0.2, 1.0)	1.0 (0.6, 1.4)
Commercial ACO contract				
No	1.162	1.165	0.3 (-0.1, 0.6)	--
Yes	1.155	1.154	0.0 (-0.3, 0.3)	-0.3 (-0.8, 0.2)
Organizational structure				
Physician-led	1.168	1.170	0.2 (-0.3, 0.7)	--
Hospital-led	1.152	1.150	-0.2 (-0.6, 0.3)	-0.5 (-1.1, 0.2)
Physician-hospital partnership	1.155	1.157	0.2 (-0.1, 0.4)	-0.1 (-0.7, 0.5)

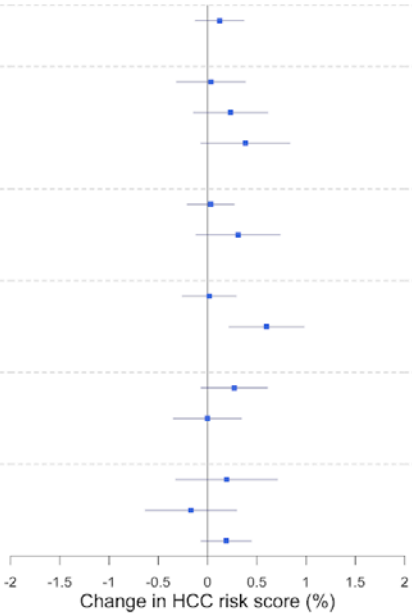


Figure B3. Relationship between beneficiary risk score and beneficiary exit or entry in the Medicare Shared Savings Program (MSSP) among beneficiaries attributed via outpatient claims submitted by primary care providers

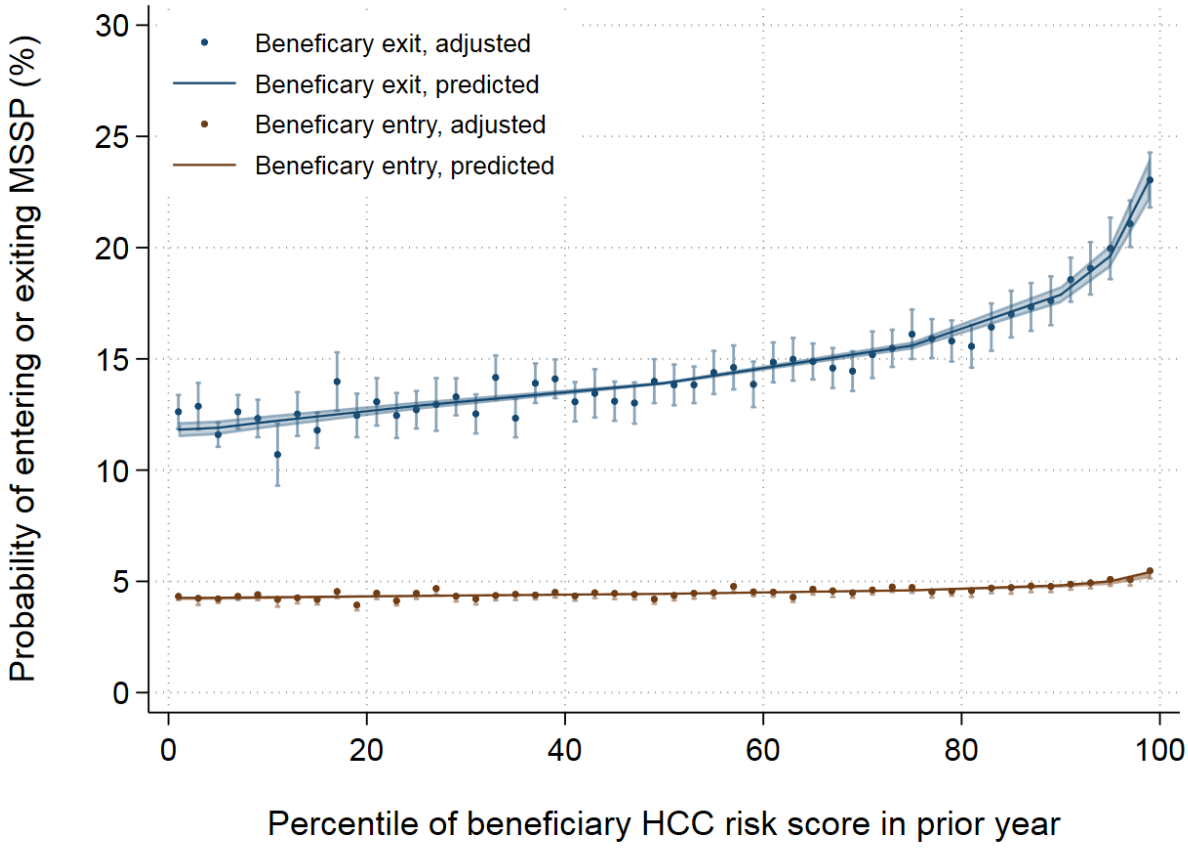


Figure B4. Relationship between clinician's average patient panel risk score and clinician entry and exit to and from the Medicare Shared Savings Program (MSSP)

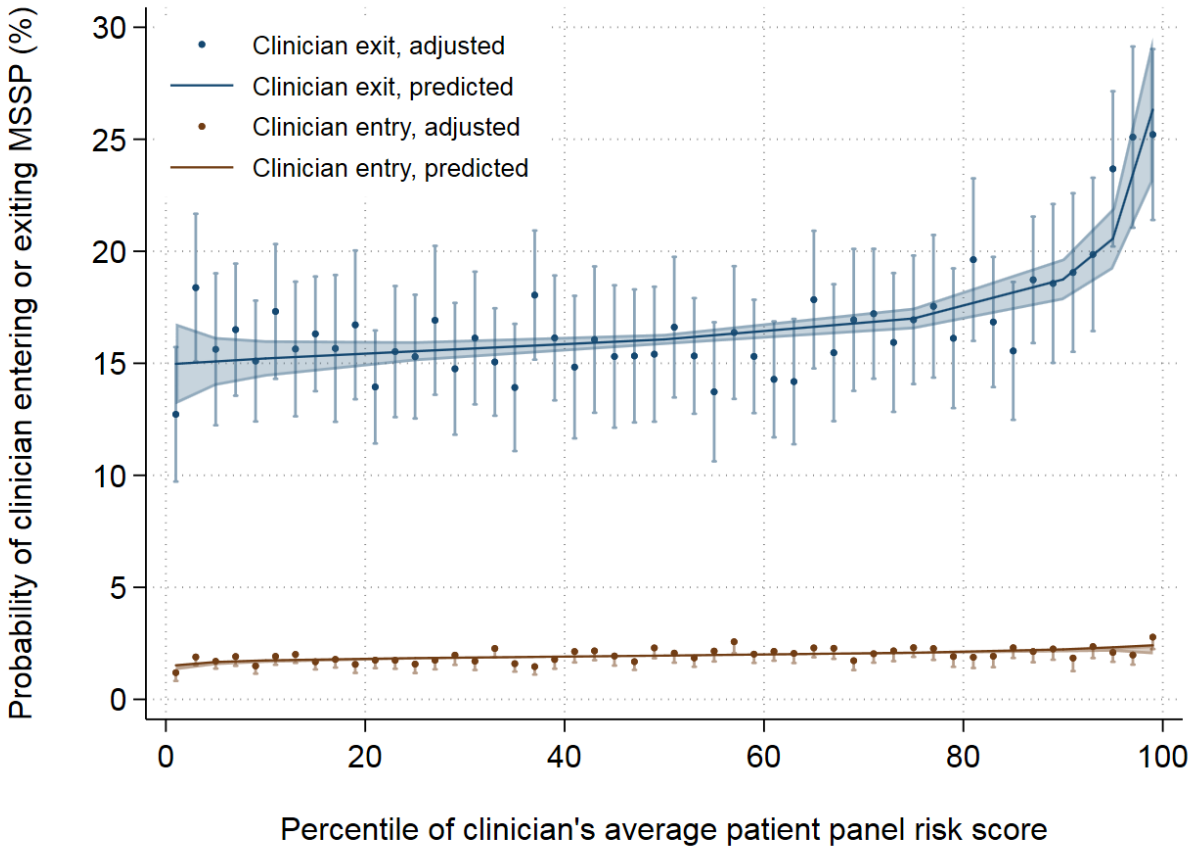


Figure B5. Trends in risk score across beneficiaries in the Medicare Shared Savings Program (MSSP) and controls (2008-2014)

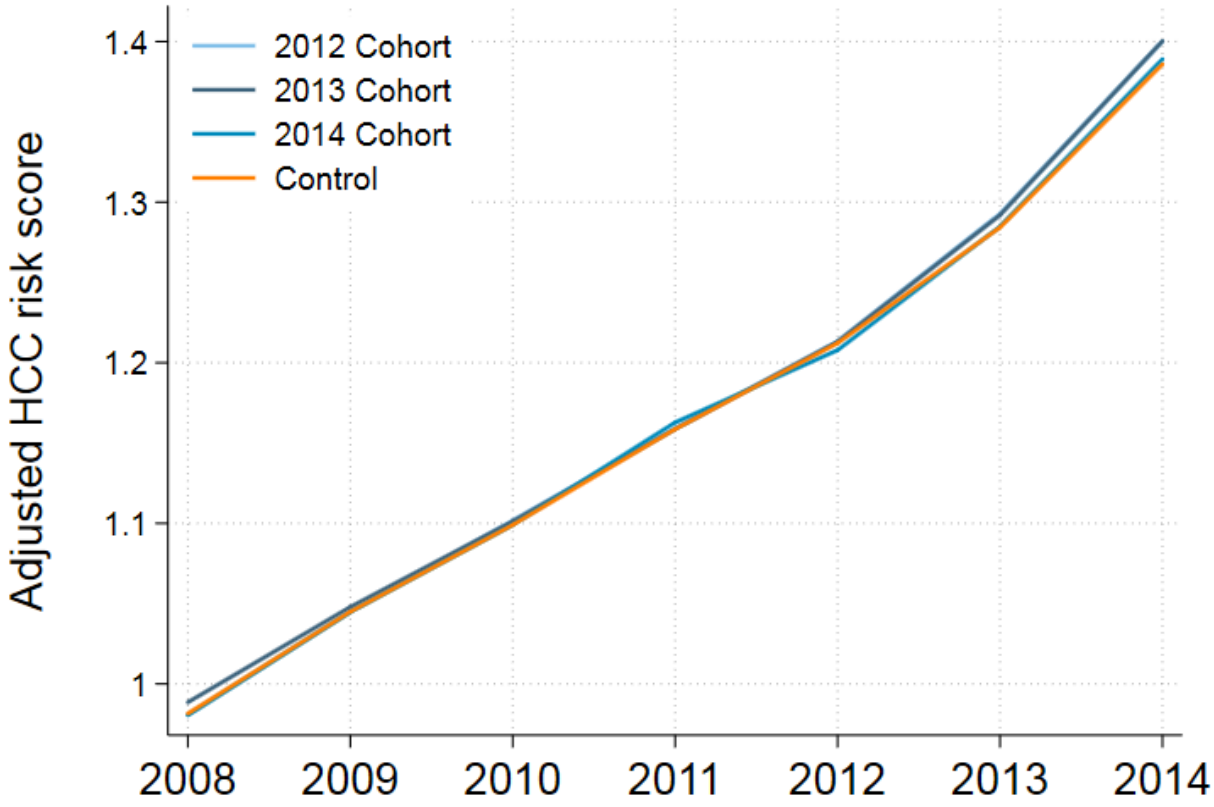


Figure B6. Event study of change in risk score before and after attribution to the Medicare Shared Savings Program (MSSP)

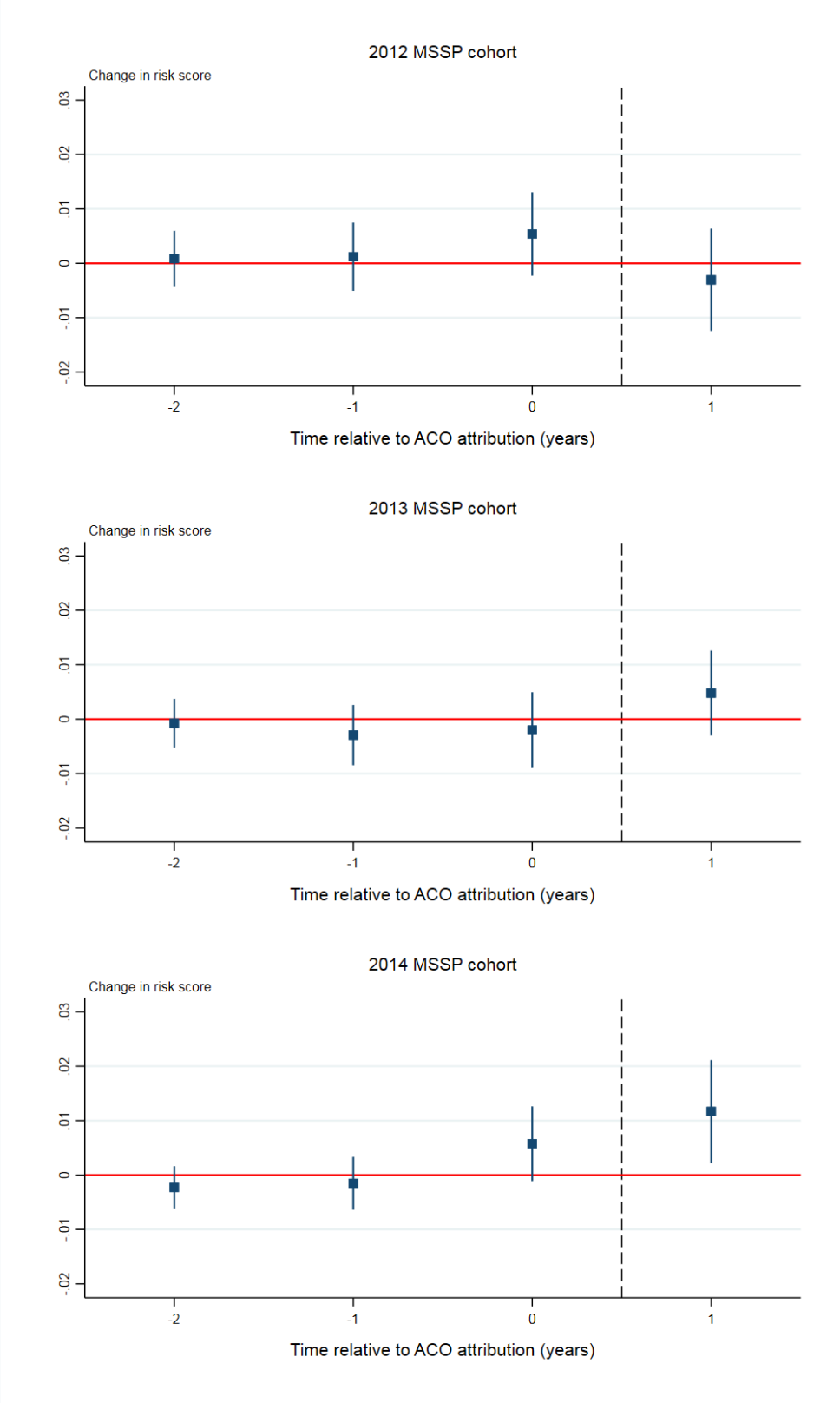


Table B1. Characteristics of Medicare beneficiaries in the analytic sample vs. excluded sample (2008-2011).

Characteristics	Analytic sample (N=13,864,627)	Excluded sample (N=19,270,424)	P value
Patient characteristics	Unadjusted mean	Unadjusted mean	
Age, y (SD)	74.3 (10.6)	71.5 (12.8)	<0.001
Female	60.7%	56.5%	<0.001
Race/Ethnicity	86.2%		
Non-Hispanic white	86.2%	80.5%	<0.001
Non-Hispanic black	7.2%	9.8%	
Hispanic	3.7%	5.9%	
Other	2.9%	3.8%	
Dual-eligibility for Medicaid (months per year)	1.8 (4.2)	2.5 (4.7)	<0.001
Disability	19.7%	24.5%	<0.001
End-stage renal disease	0.4%	1.4%	<0.001
Area-level characteristics	23.1 (12.8)		
Medicare Advantage	23.1 (12.8)	23.8 (13.2)	<0.001
Below federal poverty level	13.9 (8.5)	14.8 (9.1)	<0.001
With high school degree	28.6 (16.0)	27.4 (15.9)	<0.001
With college degree	86.9 (8.3)	86.0 (9.0)	<0.001
Beneficiary outcomes	1.167 (0.901)		
HCC risk score	1.167 (0.901)	1.235 (1.142)	<0.001
Total annual spending	7,507 (14,329)	10,819 (21,725)	<0.001

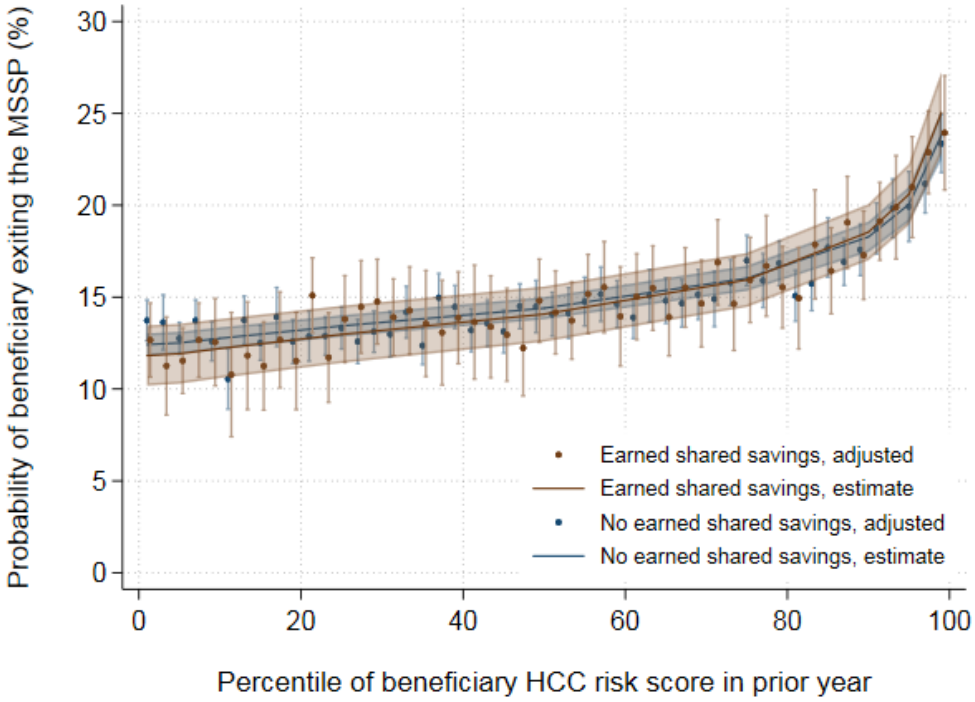
Table B2. Relationship between beneficiary risk score and beneficiary exit or entry in the Medicare Shared Savings Program (MSSP)

Average risk score in prior year	Probability of beneficiary entering MSSP, % (95% CI)	Risk difference: Pr(entry|Nth percentile)-Pr(entry|50th percentile)	P value	Risk ratio: Pr(entry|Nth percentile)/Pr(entry|50th percentile)	P value
1st percentile	4.1 (4.0, 4.3)	-0.3	0.002	0.94	<0.001
5th percentile	4.3 (4.2, 4.3)	-0.2	<0.001	0.96	<0.001
10th percentile	4.3 (4.3, 4.4)	-0.1	<0.001	0.97	<0.001
25th percentile	4.4 (4.4, 4.4)	0.0	<0.001	0.99	<0.001
50th percentile	4.4 (4.4, 4.4)	--	--	--	--
75th percentile	4.5 (4.5, 4.5)	0.1	<0.001	1.02	<0.001
90th percentile	4.7 (4.6, 4.7)	0.2	<0.001	1.05	<0.001
95th percentile	4.8 (4.7, 4.9)	0.4	<0.001	1.09	<0.001
99th percentile	5.4 (5.1, 5.6)	1.0	<0.001	1.21	<0.001

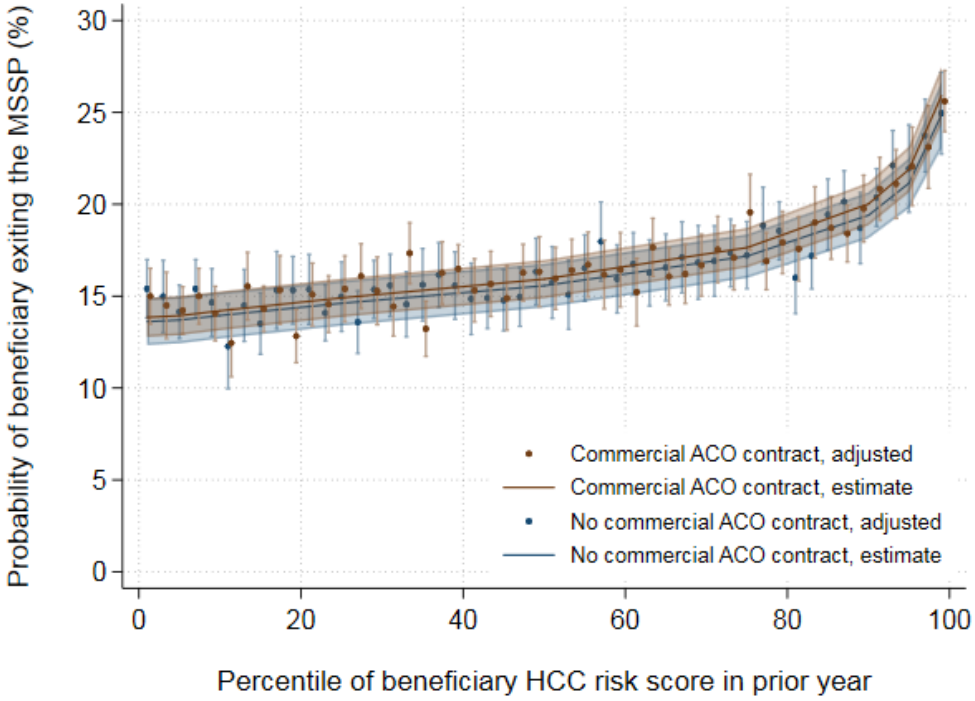
Average risk score in prior year	Probability of beneficiary exiting MSSP, % (95% CI)	Risk difference: Pr(exit|Nth percentile)-Pr(exit|50th percentile)	P value	Risk ratio: Pr(exit|Nth percentile)/Pr(exit|50th percentile)	
1st percentile	13.9 (13.5, 14.3)	-2.1	<0.001	0.87	<0.001
5th percentile	14.0 (13.7, 14.3)	-2.0	<0.001	0.88	<0.001
10th percentile	14.3 (14.0, 14.6)	-1.7	<0.001	0.89	<0.001
25th percentile	15.0 (14.8, 15.2)	-1.0	<0.001	0.94	<0.001
50th percentile	16.0 (15.9, 16.0)	--	--	--	--
75th percentile	17.6 (17.4, 17.8)	1.7	<0.001	1.10	<0.001
90th percentile	19.9 (19.5, 20.3)	3.9	<0.001	1.25	<0.001
95th percentile	21.6 (21.0, 22.2)	5.6	<0.001	1.35	<0.001
99th percentile	25.1 (24.1, 26.1)	9.1	<0.001	1.57	<0.001

Figure B7. Heterogeneity in relationship between beneficiary risk score and exit from Medicare Shared Savings Program (MSSP)

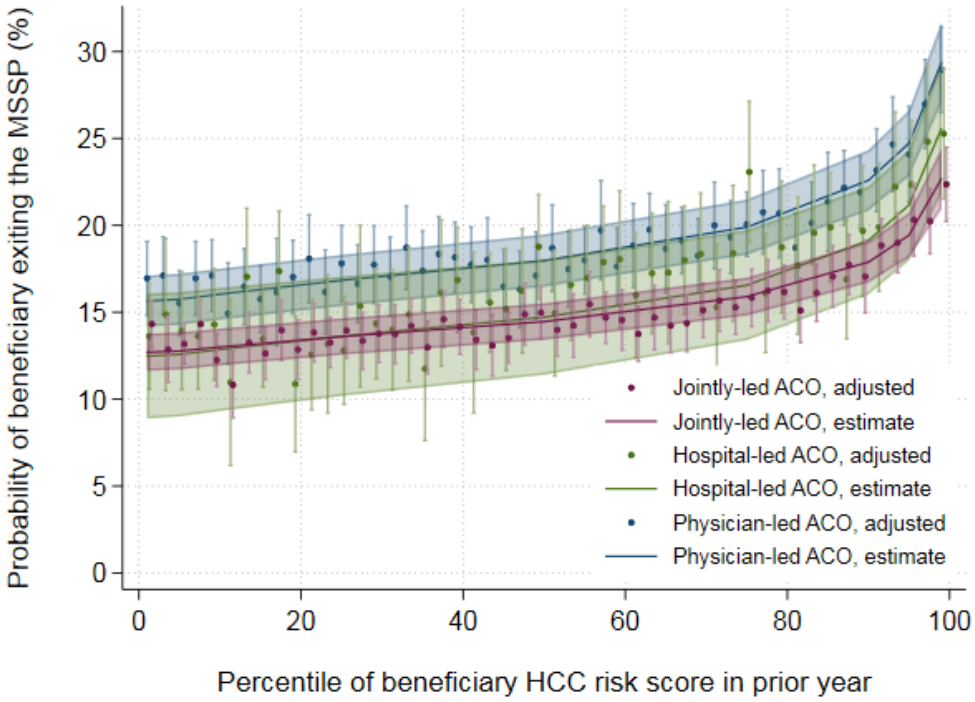
Panel A. ACO earned shared savings



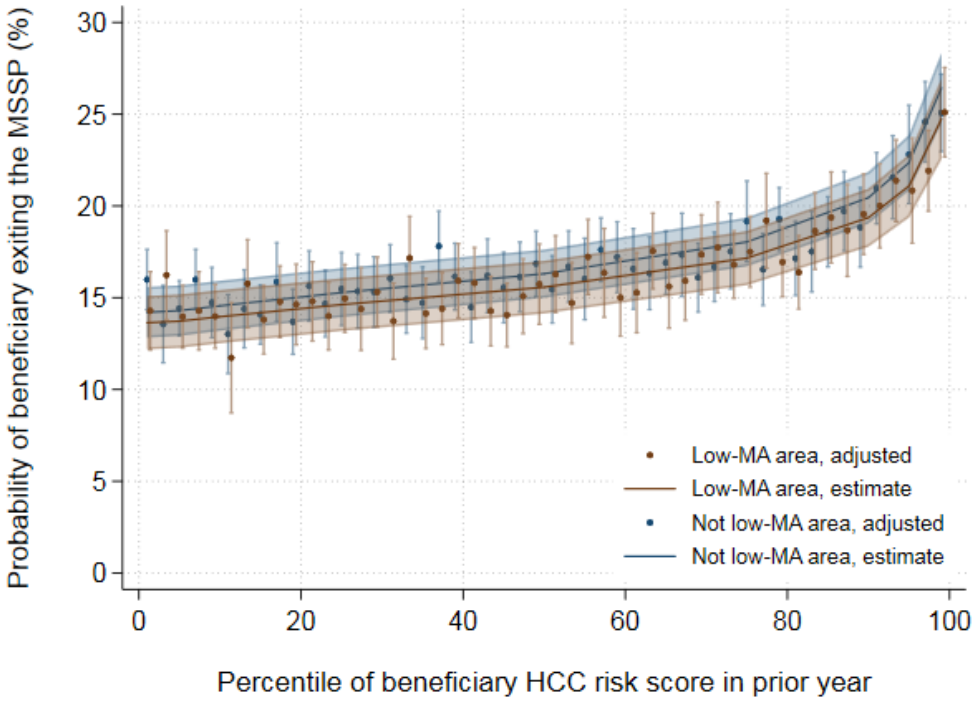
Panel B. Concurrent commercial ACO contracts



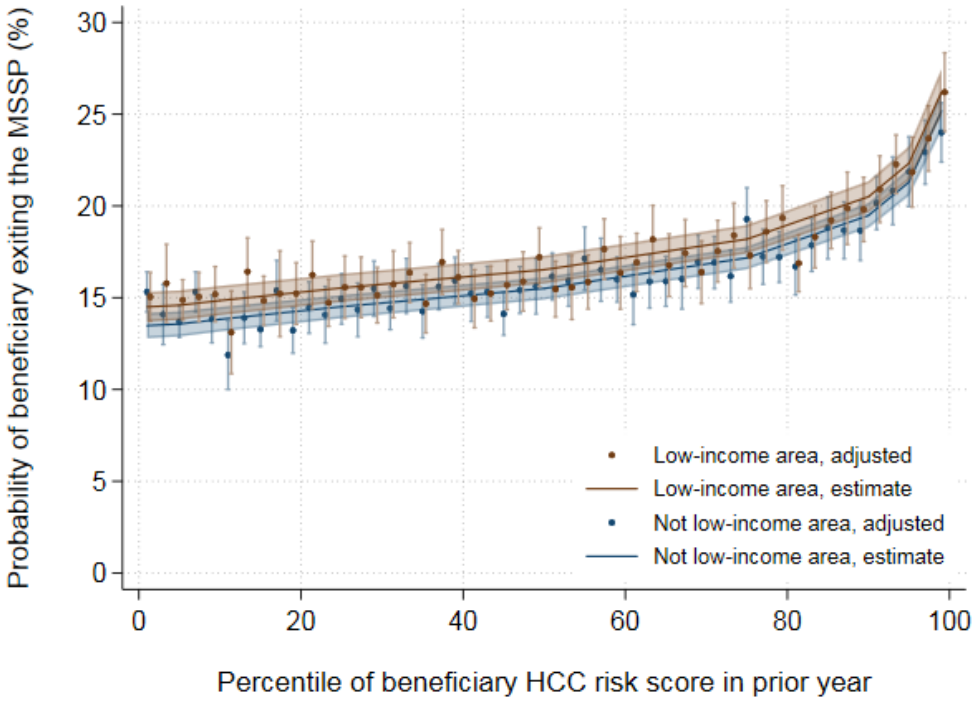
Panel C. ACO organizational structure



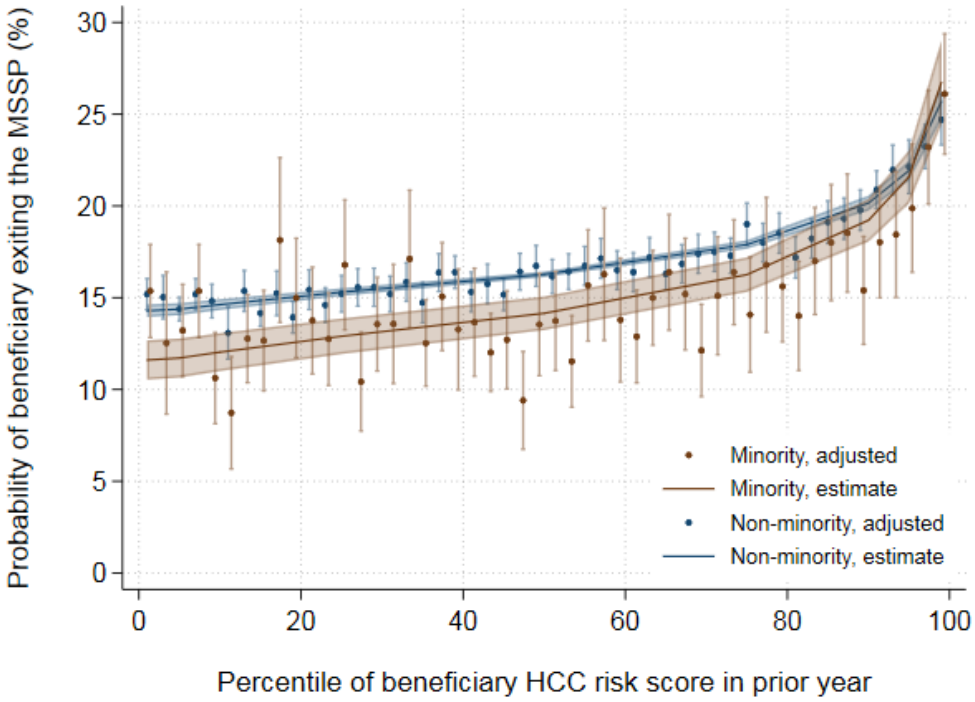
Panel D. Beneficiary county-level Medicare Advantage (MA) penetration



Panel E. Beneficiary area-level income



Panel F. Beneficiary minority status



Panel G. Beneficiary dual-eligibility for Medicaid

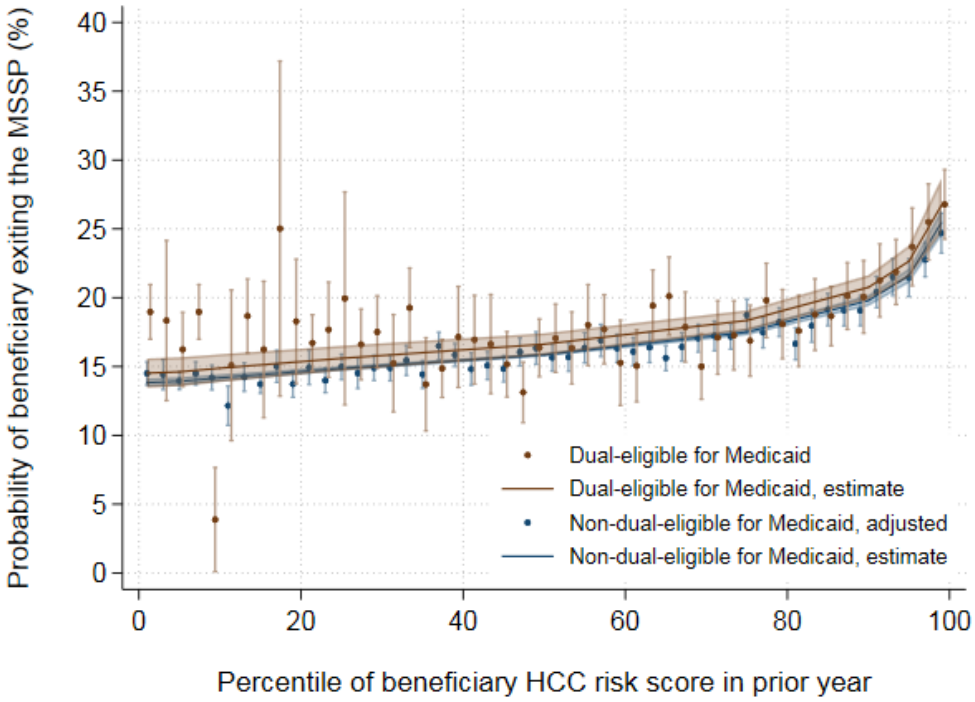
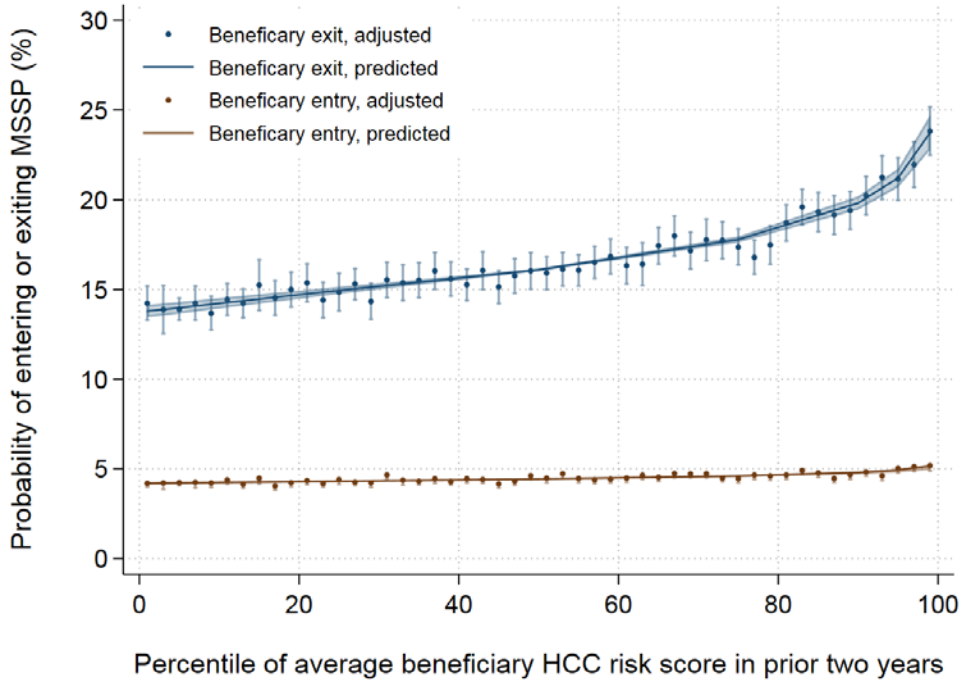


Figure B8. Relationship between beneficiary's average risk score in prior two or three years and entry and exit in the Medicare Shared Savings Program (MSSP)

Panel A. Risk score in prior two years



Panel B. Risk score in prior three years

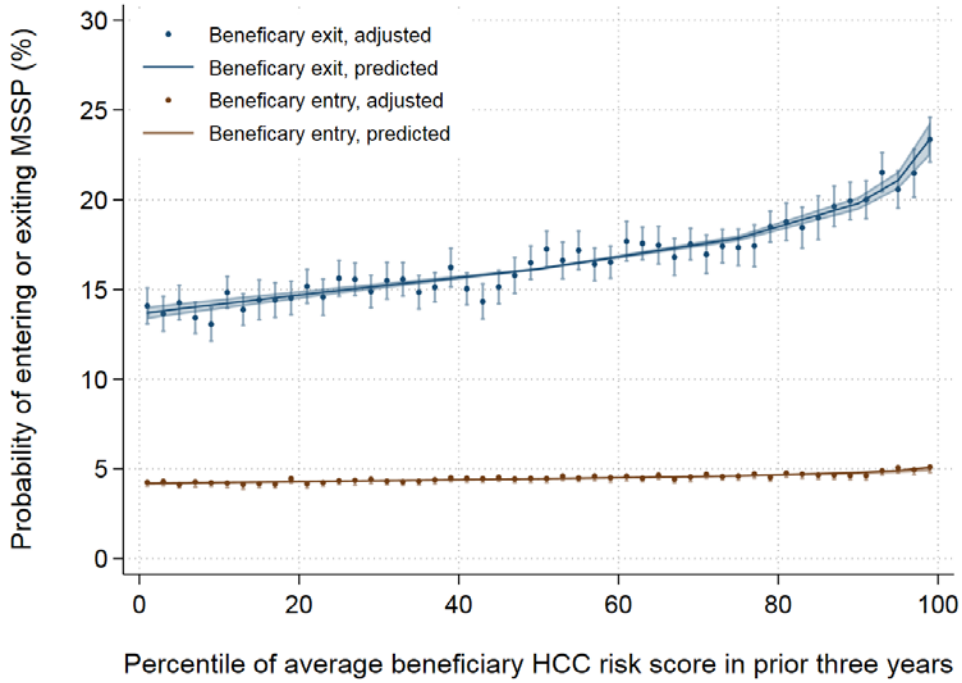


Table B3. Growth in risk score and beneficiary entry and exit to and from the Medicare Shared Savings Program (MSSP) between 2012 and 2014

	Growth in beneficiary risk score between 2012 and 2013 (%)	P value	Growth in beneficiary risk score between 2013 and 2014 (%)	P value
MSSP status	**Risk growth across MSSP status (%)**		**Risk growth across MSSP status (%)**	
Never in MSSP	3.7 (3.5, 3.8)	<0.001	5.8 (5.6, 6.0)	<0.001
Always in MSSP	3.3 (3.0, 3.7)	<0.001	5.6 (5.2, 6.0)	<0.001
Entered MSSP	5.5 (4.9, 6.1)	<0.001	7.2 (6.4, 7.9)	<0.001
Exited MSSP	8.2 (7.2, 9.2)	<0.001	8.7 (7.6, 9.9)	<0.001
	Difference in growth relative to beneficiaries never in MSSP (pp)		**Difference in growth relative to beneficiaries never in MSSP (pp)**	
Always in MSSP	-0.3 (-0.7, 0.0)	0.085	-0.1 (-0.5, 0.2)	0.436
Entered MSSP	1.9 (1.3, 2.4)	<0.001	1.4 (0.6, 2.2)	<0.001
Exited MSSP	4.5 (3.5, 5.5)	<0.001	3.0 (1.8, 4.1)	<0.001
	Difference in growth relative to beneficiaries always in MSSP (pp)		**Difference in growth relative to beneficiaries always in MSSP (pp)**	
Entered MSSP	2.2 (1.6, 2.8)	<0.001	1.6 (0.7, 2.4)	<0.001
Exited MSSP	4.8 (3.8, 5.9)	<0.001	3.1 (1.9, 4.3)	<0.001
	Difference in growth relative to beneficiaries entering MSSP (pp)		**Difference in growth relative to beneficiaries entering MSSP (pp)**	
Exited MSSP	2.7 (1.5, 3.9)	<0.001	1.6 (0.0, 3.1)	0.046

Supplemental Figure and Table Notes

Figure B1.
SOURCE: Authors' analysis of 2008-2014 data from: 20% sample of Medicare claims; the American Community Survey; CMS' Beneficiary-level Shared Savings Program File; Leavitt Partners ACO Database; CMS' Shared Savings Program Public-Use File.
NOTES: HCC is Hierarchical Condition Category.

Figure B2.
SOURCE: Authors' analysis of 2008-2014 data from: 20% sample of Medicare claims; the American Community Survey; CMS' Beneficiary-level Shared Savings Program File; Leavitt Partners ACO Database; CMS' Shared Savings Program Public-Use File.
NOTES: Model specification is provided in the main text. High Medicare Advantage penetration was defined as residing in a county > 80th percentile for the share of fee-for-service beneficiaries enrolled in Medicare Advantage. Percent change in risk score was measured as the estimated change in risk score relative to the average adjust risk score for MSSP beneficiaries in 2011 (prior to the program's start). To better isolate the influence of coding on risk score, we estimated changes in the risk score component plausibly affected by coding practice (i.e., the count of Condition Categories that originate from provider-reported diagnoses) and excluded the components originating from administrative data (e.g., age, sex, disability status) (Panel A). Because ACOs also have an incentive to lower spending in the MSSP, we also evaluated changes in risk scores normalized by total price-standardized spending (Panel B). Total annual spending was the sum of spending for inpatient, outpatient, professional, and skilled nursing facility services and was price-standardized to account for variation resulting from regional wage indices and payments for indirect medical education, Disproportionate Share Hospitals, and new technologies. The error bars indicate 95% confidence intervals. MSSP is Medicare Shared Saving Program. ACO is accountable care organization. CC is Condition Category.

Figure B3.
SOURCE: Authors' analysis of 2012-2014 data from: 20% sample of Medicare claims; the American Community Survey; CMS' Beneficiary-level Shared Savings Program File.
NOTES: Model specification is described in the Supplemental Methods. The error bars indicate 95% confidence intervals. MSSP is Medicare Shared Saving Program.

Figure B4.
SOURCE: Authors' analysis of 2012-2014 data from: 20% sample of Medicare claims; the American Community Survey; CMS' Provider-level Shared Savings Program File.
NOTES: Model specification is described in the Supplemental Methods. The error bars indicate 95% confidence intervals. MSSP is Medicare Shared Saving Program.

Figure B5.
SOURCE: Authors' analysis of 2008-2014 data from: 20% sample of Medicare claims; the American Community Survey; CMS' Beneficiary-level Shared Savings Program File.

NOTES: These analyses graphed trends in adjusted risk score across four groups of beneficiaries: beneficiaries who joined in the MSSP in 2012; beneficiaries who joined in the MSSP in 2013; beneficiaries who joined in the MSSP in 2014; and controls who were never in the MSSP. Adjusted risk score was estimated using a model that included beneficiary fixed effects, time-varying area-level characteristics, an indicator for each year, an indicator for MSSP cohort (2012 cohort, 2013 cohort, 2014 cohort, control), and an interaction term between MSSP cohort and the year of interest, but did not impose an underlying time trend. From these models, Stata's –margins– command was used to estimate predicted risk score for each cohort and in each year. The error bars indicate 95% confidence intervals. MSSP is Medicare Shared Saving Program. HCC is Hierarchical Condition Category.

Figure B6.
SOURCE: Authors' analysis of 2008-2014 data from: 20% sample of Medicare claims; the American Community Survey; CMS' Beneficiary-level Shared Savings Program File.
NOTES: We performed an event study to test for differential changes in risk score between MSSP beneficiaries vs. controls in the years prior to beneficiary attribution to MSSP ACOs. Estimates represent the association between eventual attribution to the MSSP and change in risk score in the years prior to and following attribution. Estimates are from a regression that includes indicators for the interaction between beneficiary MSSP attribution (ever/never) and time relative to MSSP attribution (where 1 = first year of attribution to MSSP), with additional controls for beneficiary fixed effects, year fixed effects, and time-varying beneficiary characteristics (described in main text). The error bars indicate 95% confidence intervals. MSSP is Medicare Shared Saving Program. ACO is accountable care organization.

Table B1.
SOURCE: Authors' analysis of 2008-2014 data from: 20% sample of Medicare claims; the American Community Survey; CMS' Beneficiary-level Shared Savings Program File.
NOTES: Study inclusion and exclusion criteria for analytic sample are described in the main text.

Table B2.
SOURCE: Authors' analysis of 2012-2014 data from: 20% sample of Medicare claims; the American Community Survey; CMS' Beneficiary-level Shared Savings Program File.
NOTES: Model specification is described in the Main Text. From these models, we estimated the probability of MSSP entry or exit across beneficiaries at the 1st, 5th, 10th, 25th, 50th, 75th, 90th, 99th, and 99th percentile of prior-year risk score. Risk ratios and risk differences were calculated by testing the difference between the probability of beneficiary exit or entry at a given risk score percentile (e.g., 95^{th} percentile) vs. the probability of beneficiary exit or entry at the median risk score (50^{th} percentile). MSSP is Medicare Shared Saving Program. ACO is accountable care organization.

Figure B7.
SOURCE: Authors' analysis of 2012-2014 data from: 20% sample of Medicare claims; the American Community Survey; CMS' Beneficiary-level Shared Savings Program File.

NOTES: To measure heterogeneity in the relationship between risk score and exit, we estimated a single fully-interacted model that added to the base regression model (explained in the text) a series of interaction terms between risk score and each ACO and beneficiary characteristic listed in the exhibits. From this fully-interacted model, we estimated the probability of MSSP exit for beneficiaries in each group (e.g., in ACOs that earned shared savings) at the 1st, 5th, 10th, 25th, 50th, 75th, 90th, 99th, and 99th percentile of prior-year risk score. The error bars indicate 95% confidence intervals. MSSP is Medicare Shared Saving Program. ACO is accountable care organization. HCC is Hierarchical Condition Category.

Figure B8.
SOURCE: Authors' analysis of 2012-2014 data from: 20% sample of Medicare claims; the American Community Survey; CMS' Beneficiary-level Shared Savings Program File.
NOTES: Models are estimated as described in the main text. However, in these sensitivity analyses, we estimated the probability of beneficiary exit or entry using average beneficiary risk score in either the prior two years (Panel A) or prior three years (Panel B). The error bars indicate 95% confidence intervals. MSSP is Medicare Shared Saving Program. HCC is Hierarchical Condition Category.

Table B3.
SOURCE: Authors' analysis of 2012-2014 data from: 20% sample of Medicare claims; the American Community Survey; CMS' Beneficiary-level Shared Savings Program File.
NOTES: Comparisons of beneficiaries who were always in the MSSP (through 2014), entered the MSSP (in 2014) or exited the MSSP (in 2014) were restricted to ACOs that entered MSSP contracts in 2012 or 2013. We excluded beneficiaries who formed ACOs that entered MSSP contracts in 2014, as we could not observe subsequent exit or entry in ACOs formed in 2014 using 2008-2014 data. Differences in risk score growth were examined using a linear spline model that included market fixed effects, year fixed effects, the previously described beneficiary characteristics, beneficiary MSSP status (always vs. never vs. enter vs. exit), splines for the years 2012-2013 (when no entry or exit occurred) and 2013-2014 (when entry or exit could occur), and an interaction between MSSP status and the two splines. We then tested for differences in risk score growth estimated from this fully-interacted spline model.

Appendix C. Low-value care and clinician engagement in the Medicare Shared Savings Program: a survey of frontline clinicians

Table of Contents

Supplemental Methods

Figure C1. CONSORT diagram

Table C1. Survey domains and instrument sources

Table C2. Awareness of ACO incentives

Table C3. Perceived influence of ACO on practice

Table C4. Perceived effect of ACO on quality improvement

Table C5. List of recommendations against low-value care

Table C6. Characteristics of survey respondents vs. non-respondents in the ACO

Supplemental Methods

The survey's target population included all clinicians listed as participants in the POM ACO administrative roster. This includes physicians, physician assistants, advanced practice nurses (nurse practitioners, certified nurse anesthetists, certified nurse midwives), and other clinicians (clinical social workers, psychologists, audiologists, podiatrists, optometrists, chiropractors, and physical therapists). We followed American Association for Public Research (AAPR) guidelines and standard definitions for determining response rates.[98] We adopted the following standards for defining surveys as complete surveys, partial surveys, or break-off: 1) less than 50% of all critical questions were considered break-off; 2) 50%-75% were considered partial surveys; and 3) more than 75% were considered complete surveys.[98]

Of the 5,139 clinicians listed in the POM ACO administrative roster, we ruled 320 clinicians ineligible for participation in the survey: 10 clinicians who were either members of the study team and/or participated in survey pre-testing; and 310 clinicians who we failed to contact with an invitation to take the survey (e.g., undeliverable/return to sender mailings). This resulted in 4,819 clinicians being ruled as eligible for the study. Of these 4,819 eligible clinicians, 1,620 clinicians submitted completed surveys. This yielded an AAPR RR1 response rate[98] of 34% (complete responses/eligible respondents=1,620/4,819). Of these 4,819 eligible clinicians, we excluded 3,075 clinicians who did not respond to the survey and 77 clinicians who began the survey but were considered survey break-offs. Our final sample of eligible respondents included 1,620 complete responses and 50 partial responses, for a total of 1,667 eligible respondents.

After survey administration, for the present analysis, we made the following additional exclusions to our initial sample of 1,667 eligible respondents. First, we included only those clinician types most frequently represented in the *Choosing Wisely*® campaign (i.e., physicians, physician assistants, advanced practice nurses), excluding the 112 respondents who were other types of clinicians, namely clinical social workers, psychologists, audiologists, optometrists, podiatrists, chiropractors, physical therapists, and other/unknown (PhD, MBA, MHSA). Second, we excluded 111 pediatricians, as they are not typically included in Medicare initiatives targeting adult beneficiaries such as ACOs. Third, we excluded 90 respondents who selected specialties for which we did not present a block of *Choosing Wisely*® recommendations in the survey (i.e., medical genetics, nuclear medicine, other). Finally, we excluded 68 clinicians who responded "does not apply to my practice" for each *Choosing Wisely*® recommendation that they were shown during the survey, leaving 1,289 respondents. This left a final analytic sample of 1,289 respondents. From these respondents, we excluded responses of "does not apply to my practice" (18%; n=1,958 responses) from regression analyses. This left a final analytic sample of 8,448 recommendation-responses.

Figure C1. CONSORT diagram

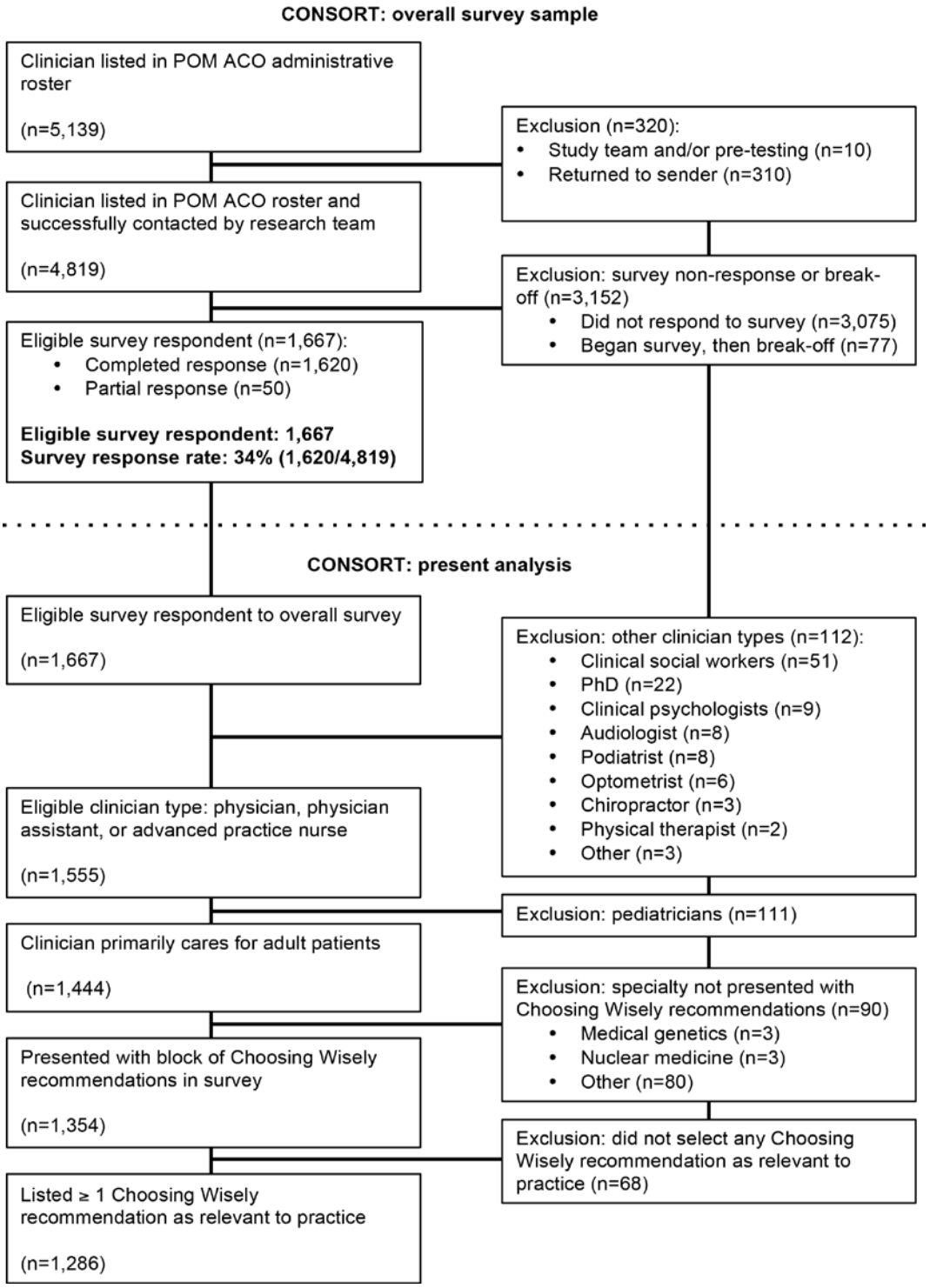


Table C1. Survey domains and instrument sources

Survey domain	Source of adapted instruments
Domain 1: Reported ability to implement recommendations against low-value care	*The Choosing Wisely*® campaign (Zikmund-Fisher et al., 2017)[86,92]
Domain 2: Involvement in ACO decision-making	"Survey of Physicians Participating in Medicare ACOs" (e.g., Schur, Sutton, 2017)[48]
Domain 3: Awareness of ACO incentives	"National Survey of Accountable Care Organization" (e.g., Colla et al., 2013)[20] "Survey of Physicians Participating in Medicare ACOs" (e.g., Schur, Sutton, 2017)[48]
Domain 4: Perceived influence of ACO on practice	"National Survey of Accountable Care Organization" (e.g., Colla et al., 2013)[20] "Survey of Physicians Participating in Medicare ACOs" (e.g., Schur, Sutton, 2017)[48]
Domain 5: Perceived effect of ACO on quality	"National Survey of Accountable Care Organization" (e.g., Colla et al., 2013)[20] "Commonwealth Fund International Health Policy Survey of Primary Care Physicians" (Osborn et al., 2015)[88]
Domain 6: Perceived effect of ACO on satisfaction and practice outcomes	(SteelFisher et al., 2009)[87] "Commonwealth Fund International Health Policy Survey of Primary Care Physicians" (Osborn et al., 2015)[88]
Domain 7: Perceived state of health care system	"Commonwealth Fund International Health Policy Survey of Primary Care Physicians" (Osborn et al., 2015)[88]
Domain 8: Perceived responsibility to address health care costs, health care quality, and population health	"Physicians, Health Care Costs, and Society" Survey (e.g., Tilburt et al., 2013)[89] "Survey of Physicians Participating in Medicare ACOs" (e.g., Schur, Sutton, 2017)[48] "Commonwealth Fund International Health Policy Survey of Primary Care Physicians" (Osborn et al., 2015)[88]
Domain 9: Participation in ACO and other value-based reforms	N/A
Domain 10: Attitude toward health policy reform	"Physicians, Health Care Costs, and Society" Survey (e.g., Tilburt et al., 2013)[89] "Commonwealth Fund International Health Policy Survey of Primary Care Physicians" (Osborn et al., 2015)[88]
Domain 11: Clinician and organization characteristics	"Commonwealth Fund International Health Policy Survey of Primary Care Physicians" (Osborn et al., 2015)[88] "Community Tracking Study Physician Survey, 2004-2005)"[91] "American Medical Association Physician Practice 2016 Benchmark Survey"[90]

Domain 1: Reported ability to implement recommendations against low-value care

"Don't recommend cancer screening in adults with life expectancy of less than 10 years."

Do **you** find this recommendation easy or difficult to **follow** most of the time?

- Easy to follow
- Difficult to follow
- Does not apply to my practice

Do most **patients** find this recommendation easy or difficult to **accept**?

- Easy to accept
- Difficult to accept
- Does not apply to my practice

Domain 2: Involvement in ACO decision-making

Which of the following best describes your decision to participate in an ACO?

- I was **involved** in the decision-making process.
- I was **not involved** but was **aware** of the decision-making process.
- I was **not involved or aware** of the decision-making process.

Domain 3: Awareness of ACO incentives

Please answer the following questions regarding the POM ACO.

	Yes	No	Don't know
The POM ACO is held accountable for both spending and quality.	○	○	○
The POM ACO lowered spending in the most recent performance year (2016).	○	○	○
The POM ACO received a financial bonus from Medicare in the most recent performance year (2016).	○	○	○
The POM ACO emphasized conducting Welcome to Medicare and Annual Wellness visits in the most recent performance year (2016).	○	○	○
The POM ACO faces both upside and downside financial risk (i.e., could either gain or lose money).	○	○	○
The POM ACO is accountable for both Medicare and non-Medicare patients.	○	○	○

Domain 4: Perceived influence of ACO on practice

Please indicate your degree of agreement or disagreement with the following statements.

	Strongly agree	Moderately agree	Moderately disagree	Strongly disagree
Joining an ACO has changed how I practice medicine.	○	○	○	○
Joining an ACO has improved the quality of care my patients receive.	○	○	○	○
Joining an ACO has made me more aware of controlling treatment costs.	○	○	○	○
Joining an ACO has changed how I am compensated.	○	○	○	○
I receive useful feedback on ACO cost and quality performance.	○	○	○	○
ACO financial bonuses are large enough to influence my behavior.	○	○	○	○

Domain 5: Perceived effect of ACO on quality improvement

Has joining an ACO had a negative impact, positive impact, or no impact on each of the following?

	Negative impact	Positive impact	No impact	Don't know
Coordinate care across care settings	○	○	○	○
Decrease unnecessary hospitalizations	○	○	○	○
Help patients manage their care between visits	○	○	○	○
Engage in shared decision-making with patients	○	○	○	○
Reduce inappropriate or harmful care	○	○	○	○
Improve the health of low-income patients	○	○	○	○
Improve the health of medically complex patients	○	○	○	○

Domain 6: Perceived effect of ACO on satisfaction and practice outcomes

Has joining an ACO had a negative impact, positive impact, or no impact on each of the following?

	Negative impact	Positive impact	No impact	Don't know
Your professional satisfaction	○	○	○	○
Your autonomy as a clinician	○	○	○	○
The administrative burden on your practice or hospital	○	○	○	○
Staff morale	○	○	○	○
The financial situation of your practice or hospital as a whole	○	○	○	○

Domain 7: Perceived state of health care system

Which of the following statements comes closest to expressing your overall view of the United States health care system?

On the whole the health care system works pretty well and only minor changes are necessary to make it work better.

There are some good things in our health system, but fundamental changes are needed to make it work better.

Our health care system has so much wrong with it that we need to completely rebuild it.

Domain 8: Perceived responsibility to address health care costs, health care quality, and population health

Please indicate your degree of agreement or disagreement with the following statements:

	Strongly agree	Moderately agree	Moderately disagree	Strongly disagree
The local health care market is moving rapidly toward paying physicians based on quality of care rather than fee-for-service alone.	○	○	○	○
I should be held financially accountable for patients' health outcomes, even if they result largely from unhealthy behaviors.	○	○	○	○
There is currently too much emphasis on costs of tests and procedures.	○	○	○	○
Clinicians should play a more prominent role in promoting the health of all individuals in their communities.	○	○	○	○
The way I practice medicine is compatible with an environment that pays for quality of care rather than entirely on a fee-for-service basis.	○	○	○	○
It is unfair to ask clinicians to be cost-conscious and still keep the welfare of their patients foremost in their minds.	○	○	○	○

When you think about your treatment decisions, how often would you say...

	Often	Sometimes	Rarely	Never	Don't know
You are aware of how much the tests or treatments that you recommend to your patients actually cost.	○	○	○	○	○
You consider the cost to the health care system when making treatment decisions.	○	○	○	○	○

Domain 9: Participation in ACO and other value-based reforms

Which of the following value-based payment programs do you participate in?

	Yes	No	Don't know
Bundled payments or **episode-based payments**, i.e., reimbursing providers for expected costs for clinically-defined episodes of care.	○	○	○
Meaningful Use of electronic health records.	○	○	○
Patient-centered medical home ("PCMH") designation for providing team-based, physician-led care that emphasizes coordination, integration, and access.	○	○	○
Accountable care organizations ("ACOs"), i.e., holding groups of clinicians financially accountable for the total health care spending of a defined patient population.	○	○	○
Pay-for-performance, i.e., linking financial incentives to specified quality or cost metrics (for example, Medicare's Hospital Readmissions Reduction Program).	○	○	○
Public reporting of clinicians cost and quality data.	○	○	○

Domain 10: Attitude toward health policy reform

Given what you know about the Affordable Care Act ("ACA" or "Obamacare"), do you have a generally favorable or unfavorable opinion about the following programs/regulations?

	Favorable	Unfavorable	Don't know
Medicaid expansion coverage for low-income adults	○	○	○
Protections against insurers charging higher prices to individuals with **pre-existing conditions**	○	○	○
Individual mandate, i.e., tax penalty on individuals who do not purchase insurance	○	○	○
Premium tax credits (subsidies) for purchasing insurance on ACA marketplaces	○	○	○
The Affordable Care Act **as a whole**	○	○	○

How much potential does each of the following reforms have for improving health care value, i.e., lowering costs while maintaining quality?

	No potential	Some potential	Great potential	Don't know
Penalizing providers for avoidable readmissions	○	○	○	○
Higher patient co-pays	○	○	○	○
Reducing compensation for the highest-paid specialties	○	○	○	○
ACOs	○	○	○	○
Limiting access to expensive treatments with little net benefit	○	○	○	○
High-deductible health plans	○	○	○	○
Enacting single-payer health care in the US	○	○	○	○
Bundled payments	○	○	○	○

Domain 11: Clinician and organization characteristics

What is your gender?

- Female
- Male

How old are you?

- Under 35
- 35-44
- 45-54
- 55-64
- 65 or older

Which best describes your primary professional activity?

- Direct patient care
- Teaching (excluding teaching that occurs during patient care)
- Research
- Administration/management
- Other (please specify)

Which best describes your practice arrangement?

- Employed by a medical school or teaching institution
- Employed by a hospital
- Employed by a managed care organization
- Solo practice
- Single specialty group practice
- Multi-specialty group practice
- Other (please specify)

What percent of your income is from your base salary (0-100%)?

Do you receive bonuses from any of the following? *Please select as many as apply.*

- Personal productivity (e.g., RVUs)
- Practice financial performance
- Other factors (e.g., patient satisfaction, clinical performance measures)
- Something else (please describe)

Compared to other patients in your region, is your panel of patients more complex, less complex, or about the same, on average? *Please provide your best guess.*

- More complex
- Less complex
- As complex

Table C2. Awareness of ACO incentives

Six-item ACO Awareness scale item (Cronbach alpha=0.72)	Incorrect (=0)	Correct (=1)	Don't know (=0)	Interitem correlation
Held accountable for both spending and quality.	296 (23.0%)	24 (1.9%)	969 (75.2%)	0.31
Lowered spending in the most recent performance year	118 (9.2%)	20 (1.6%)	1151 (89.3%)	0.29
Did not receive a financial bonus from Medicare in most recent performance year	218 (16.9%)	12 (0.9%)	1059 (82.2%)	0.31
Emphasized conducting Welcome to Medicare and Annual Wellness visits	42 (3.3%)	124 (9.6%)	1123 (87.1%)	0.31
Does not face downside financial risk, i.e., cannot lose money	38 (2.9%)	294 (22.8%)	957 (74.2%)	0.30
Held accountable for only Medicare patients	78 (6.1%)	137 (10.6%)	1074 (83.3%)	0.29
Average response: mean, 0.6, median, 0.0)				

Table C3. Perceived influence of ACO on practice

Six-item ACO Practice Change Scale (Cronbach alpha=0.88)	Strongly disagree (=1)	Moderately disagree (=2)	Moderately agree (=3)	Strongly agree (=4)	Interitem correlation
Joining an ACO has changed how I practice medicine	652 (50.6%)	402 (31.2%)	216 (16.8%)	19 (1.5%)	0.54
Joining an ACO has improved the quality of care my patients receive	727 (56.4%)	372 (28.9%)	178 (13.8%)	12 (0.9%)	0.56
Joining an ACO has made me more aware of controlling treatment costs	840 (65.2%)	349 (27.1%)	96 (7.4%)	4 (0.3%)	0.58
Joining an ACO has changed how I am compensated	471 (36.5%)	453 (35.1%)	344 (26.7%)	21 (1.6%)	0.56
I receive useful feedback on ACO cost and quality performance	609 (47.2%)	429 (33.3%)	227 (17.6%)	24 (1.9%)	0.56
ACO financial bonuses are large enough to influence my behavior	536 (41.6%)	420 (32.6%)	307 (23.8%)	26 (2.0%)	0.54
Average response: mean, 10.3, median, 10)					

Table C4. Perceived effect of ACO on quality improvement

Seven-item ACO Quality Improvement Scale	Negative impact (=-1)	No impact (=0)	Positive impact (=1)	Don't know (=0)	Interitem correlation
Coordinate care across care settings	18 (1.4%)	277 (21.5%)	210 (16.3%)	784 (60.8%)	0.49
Decrease unnecessary hospitalizations	14 (1.1%)	252 (19.6%)	150 (11.6%)	873 (67.7%)	0.51
Help patients manage care between visits	15 (1.2%)	240 (18.6%)	193 (15.0%)	841 (65.2%)	0.49
Engage in shared decision-making	13 (1.0%)	296 (23.0%)	185 (14.4%)	795 (61.7%)	0.49
Reduce inappropriate or harmful care	12 (0.9%)	261 (20.2%)	166 (12.9%)	850 (65.9%)	0.50
Improve health of low-income patients	16 (1.2%)	224 (17.4%)	114 (8.8%)	935 (72.5%)	0.52
Improve health of medically complex patients	21 (1.6%)	202 (15.7%)	202 (15.7%)	864 (67.0%)	0.49
Average response: mean, 0.9, median, 0.0)					

Table C5. List of recommendations against low-value care presented to respondents

Specialty	Recommendation against low-value care
Allergy/Immunology	Don't diagnose or manage asthma without spirometry.
Allergy/Immunology	Don't routinely prescribe antibiotics for acute mild-to-moderate sinusitis unless symptoms last for seven or more days, or symptoms worsen after initial clinical improvement.
Allergy/Immunology	Don't recommend cancer screening in adults with life expectancy of less than 10 years.
Anesthesiology	Avoid admission or preoperative chest X rays for ambulatory patients with unremarkable history and physical exam.
Anesthesiology	Don't obtain baseline laboratory studies in patients without significant systemic disease (ASA I or II) undergoing low-risk surgery – specifically complete blood count, basic or comprehensive metabolic panel, coagulation studies when blood loss (or uid shifts) is/are expected to be minimal.
Anesthesiology	Don't administer packed red blood cells (PRBCs) in a young healthy patient without ongoing blood loss and hemoglobin of ≥ 6 g/dL unless symptomatic or hemodynamically unstable.
Anesthesiology	Don't recommend cancer screening in adults with life expectancy of less than 10 years.
Cardiology	Don't perform annual stress cardiac imaging or advanced non-invasive imaging as part of routine follow-up in asymptomatic patients.
Cardiology	Don't routinely prescribe lipid-lowering medications in individuals with a limited life expectancy.
Cardiology	Don't image for suspected pulmonary embolism (PE) without moderate or high pre-test probability of PE.
Cardiology	Don't recommend cancer screening in adults with life expectancy of less than 10 years.
Critical Care	Don't administer packed red blood cells (PRBCs) in a young healthy patient without ongoing blood loss and hemoglobin of ≥ 6 g/dL unless symptomatic or hemodynamically unstable.
Critical Care	Don't image for suspected pulmonary embolism (PE) without moderate or high pre-test probability of PE.
Critical Care	Don't order diagnostic tests at regular intervals (such as every day), but rather in response to specific clinical questions.
Critical Care	Don't recommend cancer screening in adults with life expectancy of less than 10 years.
Dermatology	Don't treat uncomplicated, nonmelanoma skin cancer less than 1 centimeter in size on the trunk and extremities with Mohs micrographic surgery.
Dermatology	Don't routinely use topical antibiotics on a surgical wound.
Dermatology	Don't use oral antibiotics for treatment of atopic dermatitis unless there is clinical evidence of infection.
Dermatology	Don't recommend cancer screening in adults with life expectancy of less than 10 years.
Emergency Medicine	Don't routinely prescribe antibiotics for acute mild-to-moderate sinusitis unless symptoms last for seven or more days, or symptoms worsen after initial clinical improvement.
Emergency Medicine	Don't do imaging for low back pain within the first six weeks, unless red flags are present.
Emergency Medicine	Don't recommend cancer screening in adults with life expectancy of less than 10 years.
Emergency Medicine	Don't image for suspected pulmonary embolism (PE) without moderate or high pre-test probability of PE.
Endocrinology	Don't prescribe testosterone therapy unless there is biochemical evidence of testosterone deficiency.
Endocrinology	Don't routinely measure 1,25-dihydroxyvitamin D unless the patient has hypercalcemia or decreased kidney function.
Endocrinology	Don't routinely order a thyroid ultrasound in patients with abnormal thyroid function tests if there is no palpable abnormality of the thyroid gland.

Endocrinology	Don’t recommend cancer screening in adults with life expectancy of less than 10 years.
Gastroenterology	For pharmacological treatment of patients with gastroesophageal reflux disease (GERD), long-term acid suppression therapy (proton pump inhibitors or histamine2 receptor antagonists) should be titrated to the lowest effective dose needed to achieve therapeutic goals.
Gastroenterology	For a patient with functional abdominal pain syndrome (as per ROME IV criteria) computed tomography (CT) scans should not be repeated unless there is a major change in clinical findings or symptoms.
Gastroenterology	Do not repeat colorectal cancer screening (by any method) in average risk individuals for 10 years after a high-quality colonoscopy that does not detect neoplasia.
Gastroenterology	Don’t recommend cancer screening in adults with life expectancy of less than 10 years.
Hematology	Don’t treat patients with immune thrombocytopenic purpura (ITP) in the absence of bleeding or a very low platelet count (<30,000/microL).
Hematology	Don’t administer packed red blood cells (PRBCs) in a young healthy patient without ongoing blood loss and hemoglobin of ≥ 6 g/dL unless symptomatic or hemodynamically unstable.
Hematology	Don’t treat with an anticoagulant for more than three months in a patient with a first venous thromboembolism (VTE) occurring in the setting of a major transient risk factor.
Hematology	Don’t recommend cancer screening in adults with life expectancy of less than 10 years.
Hepatology	Don’t perform surveillance esophagogastroduodenoscopy (EGD) in patients with compensated cirrhosis and small varices without red signs treated with non-selective beta blockers for preventing a first variceal bleed.
Hepatology	Don’t perform computed tomography or magnetic resonance imaging routinely to monitor benign focal lesions in the liver unless there is a major change in clinical findings or symptoms.
Hepatology	Don’t recommend cancer screening in adults with life expectancy of less than 10 years.
Hepatology	Don’t continue treatment for hepatic encephalopathy indefinitely after an initial episode with an identifiable precipitant.
Infectious Disease	Avoid testing for a Clostridium difficile infection in the absence of diarrhea.
Infectious Disease	Avoid prophylactic antibiotics for the treatment of mitral valve prolapse.
Infectious Disease	Don’t routinely prescribe antibiotics for acute mild-to-moderate sinusitis unless symptoms last for seven or more days, or symptoms worsen after initial clinical improvement.
Infectious Disease	Don’t recommend cancer screening in adults with life expectancy of less than 10 years.
Nephrology	Don’t administer erythropoiesis-stimulating agents (ESAs) to chronic kidney disease (CKD) patients with hemoglobin levels greater than or equal to 10 g/dL without symptoms of anemia.
Nephrology	Don’t recommend cancer screening in adults with life expectancy of less than 10 years.
Nephrology	Don’t administer packed red blood cells (PRBCs) in a young healthy patient without ongoing blood loss and hemoglobin of ≥ 6 g/dL unless symptomatic or hemodynamically unstable.
Nephrology	Don’t do imaging for low back pain within the first six weeks, unless red flags are present.
Neurology	Don’t use opioid or butalbital treatment for migraine except as a last resort.
Neurology	Don’t perform imaging of the carotid arteries for simple syncope without other neurologic symptoms.
Neurology	Don’t do imaging for uncomplicated headache.
Neurology	Don’t recommend cancer screening in adults with life expectancy of less than 10 years.
Obstetrics and Gynecology	Don’t schedule elective, non-medically indicated inductions of labor or Cesarean deliveries before 39 weeks 0 days gestational age.

Obstetrics and Gynecology	Don’t administer packed red blood cells (PRBCs) in a young healthy patient without ongoing blood loss and hemoglobin of ≥ 6 g/dL unless symptomatic or hemodynamically unstable.
Obstetrics and Gynecology	Don’t treat patients who have mild dysplasia of less than two years in duration.
Obstetrics and Gynecology	Don’t recommend cancer screening in adults with life expectancy of less than 10 years.
Oncology	Don’t perform PET, CT, and radionuclide bone scans in the staging of early prostate cancer at low risk for metastasis.
Oncology	Don’t perform surveillance testing (biomarkers) or imaging (PET, CT, and radionuclide bone scans) for asymptomatic individuals who have been treated for breast cancer with curative intent.
Oncology	Don’t use combination chemotherapy (multiple drugs) instead of chemotherapy with one drug when treating an individual for metastatic breast cancer unless the patient needs a rapid response to relieve tumor-related symptoms.
Oncology	Don’t recommend cancer screening in adults with life expectancy of less than 10 years.
Ophthalmology	Don’t routinely provide antibiotics before or after intravitreal injections.
Ophthalmology	Don’t place punctal plugs for mild dry eye before trying other medical treatments.
Ophthalmology	Don’t routinely order imaging tests for patients without symptoms or signs of significant eye disease.
Ophthalmology	Don’t recommend cancer screening in adults with life expectancy of less than 10 years.
Otolaryngology	Don’t obtain computed tomography (CT) or magnetic resonance imaging (MRI) in patients with a primary complaint of hoarseness prior to examining the larynx.
Otolaryngology	Don’t order computed tomography (CT) scan of the head/brain for sudden hearing loss.
Otolaryngology	Don’t prescribe oral antibiotics for uncomplicated acute external otitis.
Otolaryngology	Don’t recommend cancer screening in adults with life expectancy of less than 10 years.
Pathology	Don’t test vitamin K levels unless the patient has an abnormal international normalized ratio (INR) and does not respond to vitamin K therapy.
Pathology	Don’t perform low risk HPV testing.
Pathology	Don’t order multiple tests in the initial evaluation of a patient with suspected non-neoplastic thyroid disease. Order thyroid-stimulating hormone (TSH), and if abnormal, follow up with additional evaluation or treatment depending on the findings.
Pathology	Don’t recommend cancer screening in adults with life expectancy of less than 10 years.
Physical Medicine and Rehabilitation	Don’t order repeat epidural steroid injections without evaluating the individual’s response to previous injections.
Physical Medicine and Rehabilitation	Don’t do imaging for low back pain within the first six weeks, unless red flags are present.
Physical Medicine and Rehabilitation	Don’t prescribe opiates in acute disabling low back pain before evaluation and a trial of other alternatives is considered.
Primary Care	Don’t routinely prescribe antibiotics for acute mild-to-moderate sinusitis unless symptoms last for seven or more days, or symptoms worsen after initial clinical improvement.
Primary Care	Don’t do imaging for low back pain within the past six weeks, unless red flags are present.
Primary Care	Don’t recommend cancer screening in adults with life expectancy of less than 10 years.
Primary Care	Don’t image for suspected pulmonary embolism (PE) without moderate or high pre-test probability of PE.

Psychiatry	Don’t use benzodiazepines or other sedative-hypnotics in older adults as a first choice for insomnia, agitation or delirium.
Psychiatry	Don’t prescribe antipsychotic medications to patients for any indication without appropriate initial evaluation and appropriate ongoing monitoring.
Psychiatry	Don’t routinely prescribe two or more antipsychotic medications concurrently.
Psychiatry	Don’t recommend cancer screening in adults with life expectancy of less than 10 years.
Pulmonology	Don’t diagnose or manage asthma without spirometry.
Pulmonology	For patients recently discharged on supplemental home oxygen following hospitalization for an acute illness, don’t renew the prescription without assessing the patient for ongoing hypoxemia.
Pulmonology	Don’t recommend cancer screening in adults with life expectancy of less than 10 years.
Pulmonology	Don’t image for suspected pulmonary embolism (PE) without moderate or high pre-test probability of PE.
Radiation Oncology	Don’t routinely recommend proton beam therapy for prostate cancer outside of a prospective clinical trial or registry.
Radiation Oncology	Don’t routinely use extended fractionation schemes (>10 fractions) for palliation of bone metastases.
Radiation Oncology	Don’t recommend radiation following hysterectomy for endometrial cancer patients with low-risk disease.
Radiation Oncology	Don’t recommend cancer screening in adults with life expectancy of less than 10 years.
Radiology	Don’t recommend follow-up imaging for clinically inconsequential adnexal cysts.
Radiology	Don’t do imaging for low back pain within the first six weeks, unless red flags are present.
Radiology	Don’t image for suspected pulmonary embolism (PE) without moderate or high pre-test probability of PE.
Radiology	Don’t recommend cancer screening in adults with life expectancy of less than 10 years.
Rheumatology	Don’t perform MRI of the peripheral joints to routinely monitor inflammatory arthritis.
Rheumatology	Don’t test for Lyme disease as a cause of musculoskeletal symptoms without an exposure history and appropriate exam findings.
Rheumatology	Don’t prescribe biologics for rheumatoid arthritis before a trial of methotrexate (or other conventional non-biologic DMARDs).
Rheumatology	Don’t recommend cancer screening in adults with life expectancy of less than 10 years.
Surgery	Don’t administer packed red blood cells (PRBCs) in a young healthy patient without ongoing blood loss and hemoglobin of ≥ 6 g/dL unless symptomatic or hemodynamically unstable.
Surgery	Avoid admission or preoperative chest X rays for ambulatory patients with unremarkable history and physical exam.
Surgery	Don’t image for suspected pulmonary embolism (PE) without moderate or high pre-test probability of PE.
Surgery	Don’t recommend cancer screening in adults with life expectancy of less than 10 years.
Urology	Don’t obtain computed tomography scan of the pelvis for asymptomatic men with low-risk clinically localized prostate cancer.
Urology	Don’t prescribe testosterone to men with erectile dysfunction who have normal testosterone levels.
Urology	Don’t recommend cancer screening in adults with life expectancy of less than 10 years.
Urology	Don’t routinely perform PSA-based screening for prostate cancer.

Table C6. Characteristics of survey respondents vs. non-respondents in the ACO

	Non-respondents (66%)	Respondents (34%)	p-value
	N=3,199	N=1,620	
Gender			0.95
Male	1,492 (66%)	754 (34%)	
Female	1,707 (66%)	866 (34%)	
Clinician specialty			0.25
Non-Physician	1,058 (68%)	509 (32%)	
Physician	2,141 (66%)	1,111 (34%)	
Clinician organization			<0.001
Michigan Medicine	1,743 (57%)	1,303 (43%)	
Integrated Health Associates	430 (79%)	117 (21%)	
Huron Valley Physicians Association	406 (88%)	58 (13%)	
MidMichigan Health	181 (82%)	41 (18%)	
St Mary's of Michigan	115 (88%)	15 (12%)	
Answer Health	106 (85%)	19 (15%)	
Wexford-Crawford PHO	67 (66%)	34 (34%)	
Oakland Southfield Physicians	78 (78%)	22 (22%)	
United Physicians	58 (87%)	9 (13%)	
Olympia Medical	15 (88%)	2 (12%)	

References

1. Chopra V, Edelson DP, Saint S. Mentorship malpractice. *JAMA - J Am Med Assoc*. 2016;315(14):1453-1454. doi:10.1001/jama.2015.18884.
2. Waljee JF, Chopra V, Saint S. Mentoring millennials. *JAMA - J Am Med Assoc*. 2018. doi:10.1001/jama.2018.3804.
3. Sinha SS, Engler TA, Nallamothu BK, et al. Sprint to work: A novel model for team science collaboration in academic medicine. *Perspect Med Educ*. 2018. doi:10.1007/s40037-018-0442-9.
4. Chopra V, Arora VM, Saint S. Will you be my mentor? - Four archetypes to help mentees succeed in academic medicine. *JAMA Intern Med*. 2018;178(2):175-176. doi:10.1001/jamainternmed.2017.6537.
5. Fisher E, Wennberg D, Stukel T, Gotlieb D, Lucas F, Pinder E. The implications of regional variations in Medicare spending. Part 1: the content, quality, and accessibility of care. *Ann Intern Med*. 2003;138(4):273-287.
6. Finkelstein A, Gentzkow M, Williams H. Sources of Geographic Variation in Health Care: Evidence From Patient Migration. *Q J Econ*. 2016;131(4):1681-1726.
7. Berwick D. Launching accountable care organizations--the proposed rule for the Medicare Shared Savings Program. *N Engl J Med*. 2011;364(16):e32.
8. Clough JD, McClellan M, DJ N, Z C, JD C, M M. Implementing MACRA. *JAMA*. 2016;313(21):2152-2161. doi:10.1001/jama.2016.7041.
9. Fisher E, McClellan M, Safran D. Building the path to accountable care. *N Engl J Med*. 2011;365(26):2445-2447.
10. Fisher E, Shortell S, Kreindler S, VanCitters A, Larson B. A framework for evaluating the formation, implementation, and performance of accountable care organizations. *Health Aff*. 2012;31(11):2368-2378.
11. McWilliams JM, Hatfield LA, Chernew ME, Landon BE, Schwartz AL. Early Performance of Accountable Care Organizations in Medicare. *N Engl J Med*. 2016;374(24):2357-2366. doi:10.1056/NEJMsa1600142.
12. McWilliams JM, Chernew ME, Landon BE, Schwartz AL. Performance differences in year 1 of pioneer accountable care organizations. *N Engl J Med*. 2015;372(20):1927-1936. doi:10.1056/NEJMsa1414929.
13. Ryan AM, Krinsky S, Maurer KA, Dimick JB. Changes in Hospital Quality Associated with Hospital Value-Based Purchasing. *N Engl J Med*. 2017;376(24):2358-2366. doi:10.1056/NEJMsa1613412.
14. Colla CH, Wennberg DE, Meara E, et al. Spending differences associated with the Medicare Physician Group Practice Demonstration. *JAMA - J Am Med Assoc*. 2012;308(10):1015-1023. doi:10.1001/2012.jama.10812.

15. Ibrahim AM, Dimick JB, Sinha SS, Hollingsworth JM, Nuliyalu U, Ryan AM. Association of coded severity with readmission reduction after the hospital readmissions reduction program. *JAMA Intern Med*. 2018;178(2):83-85. doi:10.1001/jamainternmed.2017.6148.
16. Berwick D. Making good on ACOs' promise--the final rule for the Medicare shared savings program. *N Engl J Med*. 2011;365(19):1753-1756.
17. Muhlestein DB, Saunders RS, McClellan MB. Growth Of ACOs and alternative payment models in 2017. Health Affairs Blog. http://healthaffairs.org/blog/2017/06/28/growth-of-acos-and-alternative-payment-models-in-2017/. Published 2017. Accessed June 28, 2017.
18. Markovitz AM, Ramsay PP, Shortell SM, Ryan AM. Financial Incentives and Physician Practice Participation in Medicare's Value-Based Reforms. *Health Services Research*. 2017.
19. Shortell S, Colla C, Lewis V, Fisher E, Kessell E, Ramsay P. Accountable Care Organizations: The National Landscape. *J Health Polit Policy Law*. 2015;40(4):647-668.
20. Colla CH, Lewis VA, Shortell SM, Fisher ES. First national survey of ACOs finds that physicians are playing strong leadership and ownership roles. *Heal Aff*. 2014;33(6):964-971. doi:10.1377/hlthaff.2013.1463.
21. McWilliams JM, Hatfield LA, Chernew ME, Landon BE, Schwartz AL. Early performance of accountable care organizations in Medicare. *N Engl J Med*. 2016;374(24):2357-2366.
22. McWilliams JM. Changes in Medicare shared savings program savings From 2013 to 2014. *JAMA*. 2017;316(16):1711-1712.
23. Colla CH, Lewis VA, Lao L, O'Malley AJ, Chang CH, Fisher ES. Association between Medicare accountable care organization implementation and spending among clinically vulnerable beneficiaries. *JAMA Intern Med*. 2016.
24. McWilliams JM, Chernew ME, Landon BE. Medicare ACO Program Savings Not Tied To Preventable Hospitalizations Or Concentrated Among High-Risk Patients. *Heal Aff*. 2017;(12):2085-2093.
25. Hsu J, Vogeli C, Price M, et al. Substantial physician turnover and beneficiary "churn" in a large medicare pioneer ACO. *Health Aff*. 2017;36(4):640-648. doi:10.1377/hlthaff.2016.1107.
26. Prasad V, Jena AB. Prespecified falsification end points: Can they validate true observational associations? *JAMA - J Am Med Assoc*. 2013;309(3):241-242. doi:10.1001/jama.2012.96867.
27. Colla CH, Wennberg DE, Meara E, et al. Spending differences associated with the Medicare Physician Group Practice Demonstration. *JAMA*. 2012;308:1015-1023. doi:10.1001/2012.jama.10812.
28. Fisher ES, Wennberg JE, Stukel TA, Sharp SM. Hospital readmission rates for cohorts of Medicare beneficiaries in Boston and New Haven. *N Engl J Med*. 1994;331:989-995. doi:10.1056/NEJM199410133311506.
29. Wennberg JE, Freeman JL, Shelton RM, Bubolz TA. Hospital use and mortality among Medicare beneficiaries in Boston and New Haven. *N Engl J Med*. 1989;321:1168-1173. doi:10.1056/NEJM198910263211706.

30. Landon BE, Mechanic RE. The Paradox of Coding — Policy Concerns Raised by Risk-Based Provider Contracts. *N Engl J Med*. 2017;377(13):1211-1213. doi:10.1056/NEJMp1708084.
31. Altman SH. The lessons of medicare's prospective payment system show that the bundled payment program faces challenges. *Health Aff*. 2012;31(9):1923-1930. doi:10.1377/hlthaff.2012.0323.
32. Kautter J, Pope GC, Leung M et al. *Evaluation of the Medicare Physician Group Practice Demonstration: Final Report*. Research Triangle Park, NC; 2012.
33. Centers for Medicare and Medicaid Services. *Medicare Program; Medicare Shared Savings Program: Accountable Care Organizations. Final Rule.*; 2011. https://www.federalregister.gov/documents/2011/11/02/2011-27461/medicare-program-medicare-shared-savings-program-accountable-care-organizations.
34. Centers for Medicare and Medicaid Services. *Medicare Program; Medicare Shared Savings Program: Accountable Care Organizations. Final Rule*.; 2015. https://www.federalregister.gov/documents/2015/06/09/2015-14005/medicare-program-medicare-shared-savings-program-accountable-care-organizations.
35. Centers for Medicare and Medicaid Sevices. *Medicare Program; Medicare Shared Savings Program; Accountable Care Organizations--Revised Benchmark Rebasing Methodology, Facilitating Transition to Performance-Based Risk, and Administrative Finality of Financial Calculations. Final Rule*. Vol 81.; 2016. https://www.federalregister.gov/documents/2016/06/10/2016-13651/medicare-program-medicare-shared-savings-program-accountable-care-organizations-revised-benchmark.
36. Ryan AMA, Shortell SSM, Ramsay PPP, Casalino LLP. Salary and Quality Compensation for Physician Practices Participating in Accountable Care Organizations. *Ann Fam Med*. 2015;13(4):321-324. doi:10.1370/afm.1805.
37. Burns LR, Pauly M V. Transformation of the Health Care Industry: Curb Your Enthusiasm? *Milbank Q*. 2018;96(1):57-109. doi:10.1111/1468-0009.12312.
38. DeCamp M, Farber N, Torke A, et al. Ethical challenges for accountable care organizations: a structured review. *J Gen Intern Med*. 2014;29(10):1392-1399.
39. Weissman J, Bailit M, D'Andrea G, Rosenthal M. The design and application of shared savings programs: lessons from early adopters. *Health Aff*. 2012;31(9):1959-1968.
40. Ryan A, Shortell S, Ramsay P, Casalino L. Salary and Quality Compsenation for Physician Practicies Participating in Accountable Care Organizations. *Ann Fam Med*. 2015;13(4):321-324.
41. Robinson J. The end of managed care. *JAMA*. 2011;285(20):2622-2628.
42. Shortell S, Waters T, Clarke K, Budetti P. Physicians as double agents: maintaining trust in an era of multiple accountabilities. *J Am Med Assoc*. 1998;280(12):1102-1108.
43. Crosson F. Analysis & commentary: The accountable care organization: whatever its growing pains, the concept is too vitally important to fail. *Heal Aff*. 2011;30(7):1250-1255.
44. Reschovsky JD, Rich EC, Lake TK. Factors contributing to variations in physicians' use of evidence at the point of care: A conceptual model. *J Gen Intern Med*. 2015;30:555-561. doi:10.1007/s11606-015-3366-7.

45. Shortell S, McClellan S, Ramsay P, Casalino L, Ryan A, Copeland K. Physician Practice Participation in Accountable Care Organizations: The Emergence of the Unicorn. *Health Serv Res*. 2014;49(5):1519-1536.
46. Lin MP, Muhlestein D, Carr BG, Richardson LD, Wiler JL, Schuur JD. Engagement of Accountable Care Organizations in Acute Care Redesign: Results of a National Survey. *J Gen Intern Med*. June 2018. doi:10.1007/s11606-018-4525-4.
47. Lewis VA, Tierney KI, Fraze T, Murray GF. Care Transformation Strategies and Approaches of Accountable Care Organizations. *Medical Care Research and Review*. 2017.
48. Schur CL, Sutton JP. Physicians in medicare ACOs offer mixed views of model for health care cost and quality. *Health Aff*. 2017;36(4):649-654. doi:10.1377/hlthaff.2016.1427.
49. Centers for Medicare and Medicaid Services. Shared Savings Program Accountable Care Organizations Public Use File. 2016a. https://www.cms.gov/research-statistics-data-and-systems/downloadable-public-use-files/sspaco/index.html. Accessed August 8, 2018.
50. Robinson J. Theory and practice in the design of physician payment incentives. *Milbank Q*. 2001;79(2):149-177.
51. McWilliams JM, Hsu J, Newhouse JP. New risk-adjustment system was associated with reduced favorable selection in medicare advantage. *Health Aff*. 2012;31(12):2630-2640. doi:10.1377/hlthaff.2011.1344.
52. Markovitz A, Ellimoottil C, Sukul D, et al. Risk-adjustment may lessen the financial burden imposed on hospitals that treat complex patients in Medicare's cardiac bundled payment program. *Heal Aff Invit Submiss*.
53. Ryan AM. Will Value-Based Purchasing Increase Disparities in Care ? *N Engl J Med*. 2013;369(26):2472-2474. doi:10.1056/NEJMp1312654.
54. Jha AK, Zaslavsky AM. Quality reporting that addresses disparities in health care. *JAMA*. 2014;312(3):225-226. doi:10.1001/jama.2014.7204.
55. Markovitz AA, Ramsay P, Shortell S, Ryan AM. Financial incentives and physician practice participation in Medicare's value-based reforms. *Health Serv Res*. 2017;52(6):1996-2017.
56. Lewis VA, Colla CH, Carluzzo KL, Kler SE, Fisher ES. Accountable Care Organizations in the United States: market and demographic factors associated with formation. *Heal Serv Res*. 2013;48(6 Pt 1):1840-1858. doi:10.1111/1475-6773.12102.
57. Centers for Medicare and Medicaid Services. Shared savings and losses and assignment methodology: specifications (version 3). https://www.cms.gov/Medicare/Medicare-Fee-for-Service-Payment/sharedsavingsprogram/Downloads/Shared-Savings-Losses-Assignment-Spec.pdf. Published 2014. Accessed November 8, 2017.
58. Nyweide DJ, Lee W, Cuerdon TT, Pham HH, Cox M, Rajkumar R. Association of pioneer accountable care organizations vs traditional Medicare fee for service with spending, utilization, and patient experience. *JAMA*. 2015;313(21):2152-2161.
59. Gottlieb DJ, Zhou W, Song Y, Andrews KG, Skinner JS, Sutherland JM. Prices

don’t drive regional Medicare spending variations. *Health Aff*. 2010;29(3):537-543. doi:10.1377/hlthaff.2009.0609.
60. McWilliams JM, Gilstrap LG, Stevenson DG, Chernew ME, Huskamp HA, Grabowski DC. Changes in postacute care in the Medicare shared savings program. *JAMA Intern Med*. 2017;177(4):518. doi:10.1001/jamainternmed.2016.9115.
61. Agency for Healthcare Research and Quality. Prevention quality indicators technical specifications - version 6.0. Agency for Healthcare Researcha nd Quality. https://www.qualityindicators.ahrq.gov/modules/PQI_TechSpec.aspx. Published 2016.
62. RTI International. Accountable care organization 2014 program analysis quality performance standards narrative measure specifications. http://www.healthreform.ct.gov/ohri/lib/ohri/work_groups/quality/2014-09-03/aco_narrative_measures_specs_1.pdf. Published 2014.
63. McClellan M, McNeil BJ, Newhouse JP. Does more intensive treatment of acute myocardial infarction in the elderly reduce mortality? Analysis using instrumental variables. *J Am Med Assoc*. 1994;272(11):859-866.
64. Valley TS, Sjoding MW, Ryan AM, Iwashyna TJ, Cooke CR. Association of intensive care unit admission with mortality among older patients with pneumonia. *JAMA*. 2015;314(12):1272-1278.
65. Bekelis K, Marth N, Wong K, Zhou W, Birkmeyer JK, Skinner JS. Primary Stroke Center Hospitalization for Elderly Patients With Stroke. *JAMA Intern Med*. 2016;176(9):1361-1368.
66. Crooks VA, Schuurman N, Berube M. Interpreting the results of a modified gravity model: examining access to primary health care physicians in five Canadian provinces and territories. *BMC Health Serv Res*. 2012;12(230).
67. Davidson R, MacKinnon JG. *Estimation and Inference in Econometrics*. New York: Oxford University Press; 1993.
68. Luo W, Wang F. Measures of spatial accessibility to health care in a GIS environment: synthesis and a case study in the Chicago region. *Environ Plan B*. 2003;30(6):865-884.
69. McGrail MR, Humphreys JS. Measuring spatial accessibility to primary care in rural areas: Improving the effectiveness of the two-step floating catchment area method. *Appl Geogr*. 2009;29(4):533-541.
70. WWAMI Rural Health Research Center. Rural Urban Commuting Area Data: ZIP Code RUCA Approximation (Version 2.0). http://depts.washington.edu/uwruca/ruca-approx.php.
71. Winblad U, Mor V, McHugh JP, Rahman M. ACO-affiliated hospitals reduced rehospitalizations from skilled nursing facilities faster than other hospitals. *Health Aff*. 2017;36(1):67-73. doi:10.1377/hlthaff.2016.0759.
72. United States Census Bureau. American FactFinder. https://factfinder.census.gov/faces/nav/jsf/pages/index.xhtml. Published 2017. Accessed December 14, 2017.
73. Baiocchi M, Cheng J, Small DS. Instrumental variable methods for causal inference. *Stat Med*. 2014;33(13):2297-2340. doi:10.1002/sim.6128.
74. Staiger D, Stock J. Instrumental variables regression with weak instruments.

Econometrica. 1997;65(3):557-586.
75. Baum CF, Schaffer ME, Stillman S. Instrumental variables and GMM: Estimation and testing. *Stata J*. 2003;3(1):1-31. doi:The Stata Journal.
76. Pope GC, Kautter J, Ellis RP, et al. Risk adjustment of Medicare capitation payments using the CMS-HCC model. *Health Care Financ Rev*. 2004;25:119-141. doi:hcfr-25-4-119 [pii].
77. Kronick R. Projected coding intensity in medicare advantage could increase medicare spending by $200 billion over ten years. *Health Aff*. 2017;36(2):320-327. doi:10.1377/hlthaff.2016.0768.
78. Burns A, Hayford T. Effects of Medicare Advantage Enrollment on Beneficiary Risk Scores: Working Paper 2017-08. *Congr Budg Off Work Pap 2017-08*. 2017.
79. Kronick R, Welch WP. Measuring Coding Intensity in the Medicare Advantage Program. *Medicare Medicaid Res Rev*. 2014;4(2):E1-E19. doi:10.5600/mmrr.004.02.a06.
80. Geruso M, Layton TJ. Upcoding: Evidence from Medicare on Squishy Risk Adjustment. *SSRN*. 2018.
81. The Medicare Payment Advisory Commison. *Medicare Accountable Care Organization Models: Recent Performance and Long-Term Issues*.; 2018. http://www.medpac.gov/docs/default-source/reports/jun18_ch8_medpacreport_sec.pdf?sfvrsn=0.
82. Wennberg DE, Sharp SM, Bevan G, Skinner JS, Gottlieb DJ, Wennberg JE. A population health approach to reducing observational intensity bias in health risk adjustment: cross sectional analysis of insurance claims. *BMJ*. 2014;348:g2392. doi:10.1136/bmj.g2392.
83. Department of Health and Human Services Office of the Assistant Secretary for Planning and Evaluation United States. *Report to Congress: Social Risk Factors and Performance Under Medicare's Value-Based Purchasing Programs*. Washington D.C.; 2016. https://aspe.hhs.gov/pdf-report/report-congress-social-risk-factors-and-performance-under-medicares-value-based-purchasing-programs.
84. Schwartz A, Chernew M, Landon B, McWilliams J. Changes in low-value services in year 1 of the medicare pioneer accountable care organization program. *JAMA Intern Med*. 2015;175(11):1815-1825. doi:10.1001/jamainternmed.2015.4525.
85. Reschovsky JD, Hadley J, Saiontz-Martinez CB, Boukus ER. Following the money: Factors associated with the cost of treating high-cost medicare beneficiaries. *Health Serv Res*. 2011;46(4):997-1021. doi:10.1111/j.1475-6773.2011.01242.x.
86. Zikmund-Fisher BJ, Kullgren JT, Fagerlin A, Klamerus ML, Bernstein SJ, Kerr EA. Perceived Barriers to Implementing Individual Choosing Wisely(®) Recommendations in Two National Surveys of Primary Care Providers. *J Gen Intern Med*. 2017;32(2):210-217. doi:10.1007/s11606-016-3853-5.
87. SteelFisher GK, Blendon RJ, Sussman T, Connolly JM, Benson JM, Herrmann MJ. Physicians' Views of the Massachusetts Health Care Reform Law — A Poll. *N Engl J Med*. 2009;361(19):e39. doi:10.1056/NEJMp0909851.
88. Osborn R, Moulds D, Schneider EC, Doty MM, Squires D, Sarnak DO. Primary care physicians in ten countries report challenges caring for patients with complex

health needs. *Health Aff*. 2015;34(12):2104-2112. doi:10.1377/hlthaff.2015.1018.
89. Tilburt JC, Wynia MK, Sheeler RD, et al. Views of US physicians about controlling health care costs. *J Am Med Assoc*. 2013;310(4):380-388. doi:10.1001/jama.2013.8278.
90. American Medical Association. American Medical Association Physician Practice 2016 Benchmark Survey. https://www.ama-assn.org/about/physician-practice-benchmark-survey.
91. Kemper P, Blumenthal D, Corrigan JM, et al. The Design of the Community Tracking Study: A longitudinal study of health system change and its effects on people. *Inquiry*. 1996;33(2):195-206.
92. ABIM Foundation. Choosing Wisely. http://www.choosingwisely.org/clinician-lists/. Accessed January 6, 2016.
93. Dillman D. *Internet, Phone, Mail, and Mixed-Mode Surveys: The Tailored Design Method*. New York, NY: John Wiley and Sons; 2014.
94. Admon AJ, Gupta A, Williams M, et al. Appraising the Evidence Supporting Choosing Wisely(R) Recommendations. *J Hosp Med*. 2018. doi:10.12788/jhm.2964.
95. Kerr EA, Hofer TP. Deintensification of Routine Medical Services The Next Frontier for Improving Care Quality. *JAMA Intern Med*. 2016;176:978-980. doi:10.1001/jamainternmed.2016.2292.
96. Markovitz AA, Hofer TP, Froehlich W, et al. An examination of deintensification recommendations in clinical practice guidelines: Stepping up or scaling back? *JAMA Intern Med*. 2018;178(3):414-416. doi:10.1001/jamainternmed.2017.7198.
97. Morden NE, Colla CH, Sequist TD, Rosenthal MB. Choosing wisely--the politics and economics of labeling low-value services. *N Engl J Med*. 2014;370(7):589-592. doi:10.1056/NEJMp1314965.
98. American Association for Public Opinion Research. *Standard Definitions: Final Dispositions of Case Codes and Outcome Rates for Surveys*. 9th ed. Lenexa; 2016.
99. Kolenikov S. Calibrating survey data using iterative proportional fitting (raking). *Stata J*. 2014;14(1):22-59.
100. von Hippel PT. How many imputations do you need? A two-stage calculation using a quadratic rule. *Sociol Methods Res*. 2018;in press.
101. Friedberg MW, Chen PG, White C, et al. Effects of Health Care Payment Models on Physician Practice in the United States. *Rand Heal Q*. 2015;5(1):8. http://www.ncbi.nlm.nih.gov/pubmed/28083361%0Ahttp://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=PMC5158241.
102. Cho YI, Johnson TP, VanGeest JB. Enhancing Surveys of Health Care Professionals: A Meta-Analysis of Techniques to Improve Response. *Eval Heal Prof*. 2013;36(3):382-407. doi:10.1177/0163278713496425.
103. The Dartmouth Atlas of Health Care. ZIP code crosswalks. http://www.dartmouthatlas.org/tools/downloads.aspx?tab=39#zip_crosswalks. Published 2016. Accessed December 8, 2016.
104. *Medicare Shared Savings Program: Shared Savings and Losses and Assignment Methodology Specifications, Version 4*.; 2015.
105. National Quality Forum. Diabetes Composite Measure.

https://www.qualityforum.org/Projects/n-r/Patient_Outcomes/OT1-029-09.aspx. Published 2009.
106. Chronic Conditions Data Warehouse. Condition Categories. https://www.ccwdata.org/web/guest/condition-categories. Published 2017. Accessed March 22, 2017.
107. National Quality Forum. Measure #112 (NQF 0031): Preventive Care and Screening: Breast Cancer Screening. http://www.cimronebraska.org/Main_Content_Documents/Clinical Spec PQRS 112 Breast Cancer Screening.pdf. Published 2012.
108. United States Census Bureau. U.S. Gazetteer Files. https://www.census.gov/geo/maps-data/data/gazetteer.html. Published 2017. Accessed December 17, 2016.
109. Picard R. GEONEAR: Stata module to find nearest neighbors using geodetic distances. 2010. https://ideas.repec.org/c/boc/bocode/s457146.html.
110. Guargliardo M. Spatial accessibility of primary care: concepts, methods and challenges. *Int J Health Geogr*. 2004;3(1):3-3.
111. StataCorp. *Stata 15 Base Reference Manual*. College Station, TX: Stata Press; 2017.