

Randomized trial shows healthcare payment reform has equal-sized spillover effects on patients not targeted by reform

Liran Einav^{a,b}, Amy Finkelstein^{b,c,1}, Yunan Ji^d, and Neale Mahoney^{a,b}

^aDepartment of Economics, Stanford University, Stanford, CA 94305; ^bNational Bureau of Economic Research, Cambridge, MA 02138; ^cDepartment of Economics, Massachusetts Institute of Technology, Cambridge, MA 02139; and ^dGraduate School of Arts and Sciences, Harvard University, Cambridge, MA 02138

This contribution is part of the special series of Inaugural Articles by members of the National Academy of Sciences elected in 2018.

Contributed by Amy Finkelstein, June 9, 2020 (sent for review March 16, 2020; reviewed by Jeffrey Clemens and Craig Garthwaite)

Changes in the way health insurers pay healthcare providers may not only directly affect the insurer's patients but may also affect patients covered by other insurers. We provide evidence of such spillovers in the context of a nationwide Medicare bundled payment reform that was implemented in some areas of the country but not in others, via random assignment. We estimate that the payment reform—which targeted traditional Medicare patients—had effects of similar magnitude on the healthcare experience of nontargeted, privately insured Medicare Advantage patients. We discuss the implications of these findings for estimates of the impact of healthcare payment reforms and more generally for the design of healthcare policy.

bundled payments | healthcare | randomized controlled trial | spillover

In the US healthcare system, most doctors and hospitals treat patients that are covered by multiple health insurers. This raises the possibility that a change in the way one health insurer pays healthcare providers may not only affect the healthcare received by its own patients but could also affect the healthcare received by patients covered by other insurers. Such spillovers could potentially have the same or opposite sign as the direct effect of the payment reform. To the extent they exist, spillovers have implications for estimates of the impact of healthcare payment reforms and for the design of healthcare payment policy.

Yet, credibly estimating spillovers is challenging. In the canonical research design, which compares outcomes for directly targeted patients to outcomes for nontargeted patients, spillovers cannot be identified. When a research design does permit the identification of spillovers, a skeptical reader may interpret effects on nontargeted patients as evidence of a flawed research design, rather than evidence of spillovers. For good reason, therefore, the bar for credibly identifying spillovers is high.

In this article, we estimate spillovers from a Medicare payment reform for hip and knee replacement that directly affected patients in the traditional Medicare (TM) program. This program provides an ideal setting to study spillovers because it was implemented via random assignment in some areas of the country and not others, and because it targets enough patients to potentially affect nontargeted patients as well.

In particular, the program, known as Comprehensive Care for Joint Replacement (CJR), was designed and implemented in 2016 by the Centers for Medicare and Medicaid Services (CMS) as a nationwide randomized trial. Randomization took place at the metropolitan statistical area (MSA) level, with 67 MSAs randomly assigned to treatment and 104 MSAs randomly assigned to control. In the treatment group, Medicare makes a single bundled payment to hospitals for all services related to the episode of care, including the initial hospital stay and subsequent care by other medical providers during the recovery period. In the control

group, Medicare continues to reimburse providers under the status quo fee-for-service system in which each provider receives a separate payment based on the amount of care delivered.

In the first 2 years of the program, about 400,000 patients and 1,500 hospitals were assigned to the treatment or control group. Prior work has examined the direct impact of this bundled payment reform on targeted TM patients and estimated a ~3% decline in TM spending, driven predominantly by reduced discharges from the hospital to postacute care facilities, particularly skilled nursing facilities (1–5).

We take advantage of the random assignment to study the spillover impacts of this payment reform on patients who are covered by private insurers through the Medicare Advantage (MA) program, which provides private insurance coverage in lieu of TM to about one-third of Medicare-eligible enrollees. Our work builds on Wilcock et al. (6) and Meyers et al. (7), who were the first to document spillovers of this program onto MA. We enrich this analysis by examining heterogeneity across hospitals and by considering implications for healthcare payment reform.

We study the direct effect on TM patients and spillover effects on MA patients using data from CMS. These data provide comprehensive claims-level information for TM patients and encounter-level information for MA patients. Although not as detailed, the

Significance

In the US healthcare system, most medical providers treat patients covered by multiple insurers. As a result, changes in the way one health insurer pays for medical services could, by changing provider behavior, spill over onto patients covered by other insurers. In this article, we present evidence of such spillovers in the context of a large Medicare payment reform that was implemented in some areas of the country but not in others, via random assignment. We show that the payment reform had effects of similar magnitude on both the directly covered patients and on patients covered by other insurers. We discuss implications of these findings for our interpretation of existing estimates as well as for the design of healthcare policy.

Author contributions: L.E., A.F., Y.J., and N.M. designed research, performed research, analyzed data, and wrote the paper.

Reviewers: J.C., University of California San Diego; and C.G., Northwestern University.

The authors declare no competing interest.

This open access article is distributed under [Creative Commons Attribution-NonCommercial-NoDerivatives License 4.0 \(CC BY-NC-ND\)](#).

See Profile on page 18909.

¹To whom correspondence may be addressed. Email: afink@mit.edu.

This article contains supporting information online at <https://www.pnas.org/lookup/suppl/doi:10.1073/pnas.2004759117/-DCSupplemental>.

First published July 27, 2020.

data on MA patients still allow us to quantify spillovers along several important margins.

Our results reveal spillover effects of the same sign and magnitude as the direct effects. For nontargeted MA patients, we estimate that the payment reform reduces discharges to postacute care facilities, such as skilled nursing facilities, by a statistically significant 3.3 percentage points, or about 12%. This average spillover effect is remarkably similar to the average direct effect on TM patients. We also estimate a strong, positive correlation between the direct effects and the spillover effects at the hospital level: Hospitals with larger direct effects on the discharge destination for TM patients also have larger spillover effects on the discharge destination for MA patients. The effects—both direct and indirect—tend to be larger for hospitals with a large number of directly affected patients and for higher-quality hospitals.

Our findings have implications for the optimal design of healthcare payment policies. When health insurers separately design payment policies, they may not take into account the spillovers to other patients. As a result, the separately designed payment policies may jointly result in suboptimal incentives for healthcare providers, even when the policies are individually optimal. In other words, this is a classic case of what economists call externalities.

Our evidence of substantial spillovers onto nontargeted patients also has potential implications for the existing empirical literature on the impact of healthcare policy. Because existing research tends to focus on directly targeted patients, the costs and benefits on nontargeted patients may not be counted. In addition, if nontargeted patients are used as a “control group” for targeted patients, estimates of the direct effects may be biased. In our context, if we used nontargeted MA patients as a control group, the estimated impacts on targeted TM patients would be substantially biased toward zero.

Prior research on potential spillovers from health policies is not limited to the randomized, bundled payment reform we study. There is indirect evidence of spillovers from quasi-experimental research. For instance, Finkelstein (8) finds that the introduction of Medicare for the elderly increased technology adoption and hospital construction, which could have affected care provided to nonelderly patients. Likewise, Baicker and Staiger (9) find that changes in Medicaid payment policies reduced not only the mortality rate of Medicaid-covered infants but also the post-heart-attack mortality rate of Medicare-covered heart-attack patients, perhaps through an overall improvement of hospital quality. There have also been efforts to directly examine the spillovers of health insurance policies, for example by examining the impact of managed care on the treatment of commercial fee-for-service patients (10, 11) and the impact of MA on the treatment of TM patients (12, 13). Clemens and Gottlieb (14) provide another example of spillovers, through the setting of payment rates.

A randomized evaluation is arguably the most effective way to credibly test for spillovers. Indeed, such designs have been used to examine spillovers in other contexts (15, 16). Unfortunately, from this perspective, the three previously conducted randomized evaluations of health insurance policy in the United States (17–19) have been relatively small in scale, and at the individual level, making them less conducive for estimating spillovers. Medicare’s random assignment of payment reform to some MSAs and not others thus presented a unique opportunity.

In the rest of the article, we first describe our setting, data, and framework, then present our results, and finally discuss the implication of our findings and directions for further work.

Setting, Data, and Framework

Setting. We study a payment reform in the TM program. Medicare, the US public health insurance program for the elderly and disabled, had about 60 million enrollees and \$731 billion in annual spending in 2018 (20, 21). About two-thirds of these enrollees are covered by the TM program, which offers a publicly administered insurance product. The remaining one-third of Medicare-eligible enrollees opt out of TM and enroll in the MA program, where they are covered by private insurance plans. Private insurers in MA are paid a fixed (risk-adjusted) monthly amount by Medicare for each beneficiary they enroll (21).

Under TM, providers have been paid primarily on a fee-for-service basis, in which providers are reimbursed based on claims submitted for medical services. For instance, for a patient undergoing hip replacement, Medicare would make separate payments to the hospital for the initial hospital stay, to the surgeon for performing the procedure and for each postoperative visit, to the skilled nursing facility for postacute care, and to the equipment owner for the rental of a wheelchair during the recovery period. Moreover, within most of these categories, the payment would depend on the specific services provided.*

A key, well-appreciated concern with such fee-for-service payments is that they may encourage providers to deliver care that has low or little value to the patient. Patients have to pay only a small fraction of the costs of the treatment, and providers get paid more the more care they deliver. Over the last decade, Medicare has responded to these concerns by increasing the use of alternative, arguably more efficient payment models—such as accountable care organizations, bundled payments, and primary care coordination models—which reduce incentives to provide excessive care. This shift has been quite dramatic. In 2015, the Obama administration announced the goal of shifting 50% of TM reimbursement from fee-for-service to alternative payment models by the end of 2018, and it reached 38% in 2017.[†]

In this article, we focus on an alternative payment model under which Medicare pays for hip and knee replacement with a single bundled payment. Bundled payments represent a middle ground between fee-for-service and fully capitated models (which pay providers a fixed per-capita amount per annum). Under bundled payments, Medicare makes a single payment for all services related to a clearly defined episode of care. The payments are sometimes adjusted to reflect predictable variation in patient health or in costs in the local medical market, or to limit the hospital’s exposure to potential losses.

Proponents of bundled payments argue that it will improve coordination of care and reduce unnecessary healthcare utilization. Yet, some are concerned that because providers do not receive payments based on the quantity of care they provide, they may cut back on necessary care or cherry-pick patients with a lower cost of provision (23, 24). Most bundled payment programs have been implemented on a voluntary basis, with a small number of hospitals choosing to participate. These programs have covered about four dozen conditions (25), including coronary bypass (26), hip and knee replacements (27–29), and others (30).

In this article we study the first mandatory participation bundled payment program, CJR. The program started in 2016 and covers hip and knee replacement (also known as lower extremity joint replacement, or LEJR). LEJR is a large category, with 530,000 procedures and \$7.1 billion in Medicare inpatient

*One exception to this system is hospital reimbursements. Since 1982, when the Prospective Payment System was adopted, hospitals are paid a single payment for the entire hospital stay. This payment is based on the patient’s diagnosis but not on the length of stay or the services provided during the stay (22).

[†]<https://www.hhs.gov/about/budget/fy2020/performance/performance-plan-goal-1-objective-1/index.html>.

spending in 2016, representing about 5% of TM admissions and 5% of its inpatient spending.[‡] Under CJR, an episode begins with a hospital stay in a qualifying diagnosis-related group (DRG) and ends 90 days after hospital discharge. Hospitals are paid a fixed target price and are responsible for medical spending over the entire episode (except for spending that is deemed as obviously unrelated). There are measures to limit hospital risk exposure and to make sure that hospitals meet minimum quality thresholds. By contrast, under the status quo fee-for-service regime, hospitals are paid a fixed amount for the hospital stay, while the surgical procedure and postdischarge care are reimbursed separately.

We study the impact of this bundled payment reform on the healthcare of non-TM patients. There are multiple insurers in the US healthcare system. On the public side, there is both TM and Medicaid, the health insurance program for low-income individuals. On the private side, a large number of private insurers provide MA coverage to Medicare enrollees and “commercial insurance” to the nonelderly through employer-sponsored insurance (ESI) or the plans available on the health insurance exchanges set up under the Affordable Care Act. TM is large enough that changes in its payment structure may potentially spill over onto the care of other patients; for example, in 2016, TM accounted for 28% of overall hospital admissions and 40% of LEJR admissions.[§]

We focus on potential spillovers of the TM bundled payment program to privately insured MA patients. As mentioned, under MA, private insurers receive capitated payments from Medicare for providing Medicare beneficiaries with health insurance that roughly mimics commercial health insurance for the under-65 population. The idea is that because the private insurer is exposed to enrollees’ healthcare costs, it will have incentives to reduce care utilization relative to fee-for-service TM. Prior work has consistently demonstrated that MA is associated with lower healthcare utilization for patients relative to similar patients in TM (31–34). Again, as with bundled payments, the concern with MA is that because private insurers’ reimbursement does not vary with how much care their enrollees use, they may cut back on necessary care or cherry-pick patients with lower relative costs (35).

Data. The main data we use are Medicare data from the CMS. We also use private insurer data from the Health Care Cost Institute (HCCI). We analyze claims and enrollment data from four years: 2013, 2014, 2016, and 2017. We study outcomes in 2016 and 2017, which are the first 2 years of the program. We use data from 2013 and 2014, which was before the program started, in order to construct control variables. We omit data from 2015, since assignment of MSAs to treatment and control was announced in July 2015 (36), potentially causing contaminating anticipatory effects. *SI Appendix, Appendix A* provides details on the data and sample definition.

We use three main files from CMS: the Master Beneficiary Summary File (MBSF), the Inpatient file (IP), and the Medicare Provider Analysis and Review file (MedPAR). The MBSF provides basic demographics on all Medicare enrollees, including age, race, sex, Medicaid enrollment, and whether the individual receives Medicare through TM or MA. To measure outcomes for TM patients, we use the IP file. These data contain claims-level data on hospital inpatient services, including information on the provider, diagnoses, admission and discharge dates, and discharge destinations for each inpatient admission. These data

are based on standard, widely used Medicare claims data submitted by providers for reimbursement by CMS.[¶]

To measure outcomes for MA patients, we use the MedPAR file. For MA patients, these data contain information on inpatient hospital admissions.[#] MedPAR files have been used in several previous studies of the MA population (6, 7, 37, 38). They have the advantage of covering the universe of MA patients but also have two potential limitations. First, because these claims are not submitted for reimbursement purposes, there exists some uncertainty about the reliability of these “information-only” claims (39). Second, they do not contain any information on postdischarge healthcare utilization for MA patients, or information on healthcare spending for MA patients.

To address these limitations, we also obtained private insurance data from HCCI. The HCCI data are provided by three large insurers: UnitedHealthcare, Aetna, and Humana. HCCI pools these data—masking insurer identities—and makes them available for research. The three insurers cover approximately half of MA enrollees and about a quarter of enrollees with ESI (40, 41). HCCI data have been previously used to analyze both MA claims and claims for ESI (34, 41–47). The data on claims paid to healthcare providers appear comprehensive and high-quality, with certain known exclusions (specifically, enrollees in highly capitated plans and special needs plans) (34). The enrollment file includes monthly enrollment indicators and limited demographic information, including age in 10-year bins, gender, whether the enrollee is also receiving Medicaid, and indicators that allow us to identify whether an enrollee is in a MA or a commercial plan.

Experimental Design and Baseline Sample. CJR was initially designed by CMS as a 5-year, mandatory participation, randomized trial. Year 1 was defined as 1 April to 31 December 2016, and years 2 through 5 were defined as the calendar years 2017 through 2020. CMS randomized eligible MSAs into treatment (bundled payments) or control (status quo fee-for-service) by strata based on MSA population and historical LEJR spending.^{||} CMS announced assignment to treatment in the July 2015 *Federal Register* (36). Fig. 1 shows the final assignment. After exclusions, the program covered 67 treatment MSAs and 104 control MSAs.^{**} Within the 171 MSAs assigned to treatment or control, a small number of hospital types and episode types were further excluded from eligibility (1).

We analyze the first and second years of the CJR program, which include episodes that begin between 1 April 2016 and 15 September 2017, virtually guaranteeing that they end by 31 December 2017 (the end of the second year of the program). An episode is defined as an acute-care hospital stay that results in a discharge in one of the two included DRGs (469 and 470) and ends 90 days after discharge from the acute-care hospital. We

[¶]The claims-level data remain the same even under bundled payments. Under bundled payment, providers continue to bill as if they are paid under fee-for-service, and hospitals receive a “reconciliation payment” at the end of the year based on the difference between their target price and actual billing.

[#]Hospitals are required to report data on MA patients to CMS in order to receive indirect medical education (IME), graduate medical education (GME), or disproportionate share hospital (DSH) payments <http://www.medpac.gov/docs/default-source/default-document-library/ma-encounter-data-april18.pdf>.

^{||}MSAs were divided into eight strata based on the interaction of historical LEJR spending quartile and above- vs. below- median MSA population. Within each stratum, MSAs were assigned to treatment with probabilities that ranged from 30 to 45%, with higher treatment probabilities for strata with higher historical LEJR payments.

^{**}After the initial assignment, Medicare realized that they did not exclude some hospitals that were already (prior to assignment) signed up for BPCI (a different Medicare program) and subsequently excluded an additional eight MSAs from the treatment group. Medicare later identified the 17 MSAs in the control group that would have been excluded based on these criteria. Since these exclusions were based on hospital decisions made prior to assignment we simply drop these 25 MSAs from the study.

[‡]Authors’ analysis of the 2016 MedPAR file.

[§]Based on authors’ calculations from the 2016 National Inpatient Sample and MedPAR files.

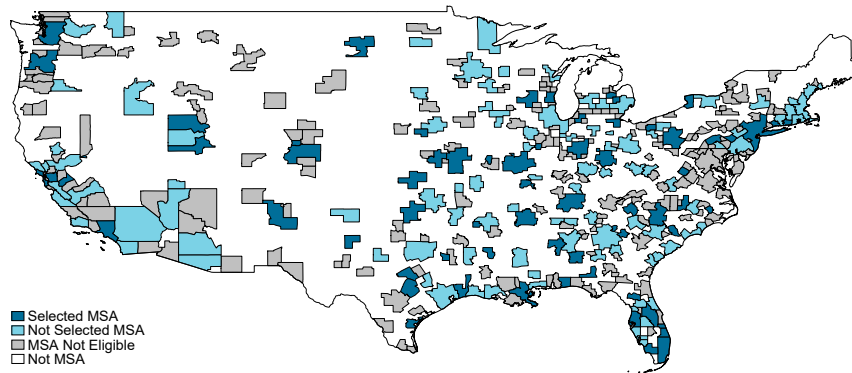


Fig. 1. Map of MSAs by whether selected for CJR, in the mainland United States. The map shows MSAs that are selected for final treatment, eligible but not selected, and not eligible, in the mainland United States. In addition, there is one not selected and one not eligible MSA in Alaska, one not selected and one not eligible MSA in Hawaii, and one not selected and six not eligible MSAs in Puerto Rico.

analyze the final set of 171 MSAs included in the CJR program. Following Finkelstein et al. (1), we impose hospital and patient eligibility criteria on LEJR admissions on TM, MA, and ESI patients to restrict our sample to LEJR episodes that would have otherwise been eligible for CJR. The criteria for MSA eligibility were determined by CMS and described in detail in prior work [see Finkelstein et al. (1) appendix, section 1]; *SI Appendix, Appendix A* provides more detail.

Following Finkelstein et al. (1), our primary outcome of interest is the share of CJR-eligible episodes discharged to institutional postacute care. Institutional postacute care consists of three types of care—skilled nursing facilities, inpatient rehabilitation facilities, and long-term care hospitals—of which skilled nursing facilities are by far the most common. We also analyze share discharged home with home health care services and the share discharged home without home health. Secondary outcomes—which are not available for MA patients in the CMS data—include the number of days in skilled nursing facilities during the episode, spending in skilled nursing facilities during the episode, and total episode spending.^{††} Finally, we examine the impact of bundled payment on patient volume. *SI Appendix, Appendix B* provides details on the construction of specific variables we analyze.

Econometric Framework. We study the impact of CJR during the first 2 years of the program. We focus on the first 2 years of the program because starting in year 3 the program was unexpectedly modified to be voluntary for some treatment MSAs, with hospitals endogenously selecting out of the program (4).

Since participation in the program was mandatory during the years we study it, and was randomly assigned at the MSA level, estimation of its impact is straightforward. Essentially, we compare average outcomes in MSAs randomly assigned to the bundled payment regime to average outcomes in MSAs randomly assigned to remain under the status quo fee-for-service payment regime. Specifically, we estimate effects of the program with regressions of the form

$$y_{j,(2016-2017)} = \beta_0 + \beta_1 BP_j + \beta_2 y_{j,2014} + \beta_3 y_{j,2013} + \delta_{s(j)} + \epsilon_j, \quad [1]$$

where $y_{j,(2016-2017)}$ is the average per-episode outcome in MSA j during the first 2 years of the program, BP_j is an indicator for the MSA being randomly assigned to bundled payments, and β_1 is the average treatment effect of bundled payment. As in

Finkelstein et al. (1), we include lagged outcomes from years 2014 and 2013 to improve statistical power. Because the randomization was conducted within strata, with differences in the probability of treatment assignment, we include strata fixed effects $[\delta_{s(j)}]$ to isolate the experimental variation. In all tables, we report heterogeneity-robust SEs.

Results

Summary Statistics. Table 1 reports summary statistics for CJR-eligible patients in control MSAs in 2016 and 2017. Part A reports demographics, and part B reports healthcare utilization and spending.

We show CMS data on CJR-eligible patients in TM (column 1) and in MA (column 2). We observe about 340,000 CJR-eligible episodes in control MSAs; about 65% are in TM, with the remaining 35% in MA. TM and MA patient demographics are similar: The average age is 72, over 60% are female, about 10% are eligible for Medicaid, and the vast majority of admissions are without a major complication or comorbidity.

Discharge locations for TM and MA patients are also strikingly similar: About 30% are discharged to institutional postacute care, a little over one-third are discharged home with home health care services, and about one-third are discharged home without home health care. The similarity in discharge destinations is somewhat surprising, given that prior work has shown that MA patients are much less likely to be discharged to institutional postacute care relative to similar TM patients, for the broader population of hospital patients (34).

As a point of comparison, column 3 shows summary statistics for the ~35,000 CJR-eligible episodes covered by MA in the HCCI data. MA patient demographics for the three insurers providing MA in the HCCI data are generally similar to the demographics for the total MA patient sample (column 2); the exception is that the share eligible for Medicaid is substantially lower in the HCCI data. The share discharged to different locations is also very similar for the MA patients covered by the HCCI insurers and the total MA patient sample. This suggests that the HCCI data are reasonably representative of the overall MA population.

The HCCI data in column 3 also allow us to examine length of stay and spending postdischarge and compare it to TM patients (column 1). In both populations, the vast majority of patients discharged to institutional postacute care are sent to skilled nursing facilities. Despite similar rates of discharge to skilled nursing facilities (25% of patients in both TM and MA), MA patients have shorter length of stay (5.0 days for MA patients vs. 7.3 days for TM patients) and they have substantially lower average

^{††}We were unable to reliably, separately identify long-term-care hospital or inpatient rehabilitation facility claims in the HCCI data, which is why our postdischarge utilization and spending measures focus on skilled nursing facilities only.

Table 1. Summary statistics (control MSAs 2016–2017)

	CMS data		HCCI data	
	TM (1)	MA (2)	MA (3)	ESI (4)
A: Demographics and diagnosis characteristics				
Average age	72.3 (0.83)	72 (1.61)	—*	—
Share 65 and older	0.89 (0.03)	0.87 (0.07)	0.89 (0.07)	—
Share female	0.64 (0.03)	0.64 (0.04)	0.63 (0.07)	0.54 (0.10)
Share eligible for Medicaid	0.1 (0.05)	0.12 (0.08)	0.06 (0.06)	—†
DRG breakdown‡				
469 (with major complication or comorbidity)	0.04 (0.01)	0.04 (0.02)	0.03 (0.02)	0.01 (0.03)
470 (without major complication or comorbidity)	0.96 (0.01)	0.96 (0.02)	0.97 (0.02)	0.99 (0.03)
B: Healthcare utilization and spending				
Share discharged to				
Institutional postacute care§	0.31 (0.1)	0.28 (0.11)	0.27 (0.12)	0.02 (0.04)
Home (with home health care)	0.34 (0.2)	0.37 (0.2)	0.37 (0.21)	0.23 (0.20)
Home (without home health care)	0.33 (0.23)	0.34 (0.23)	0.35 (0.24)	0.71 (0.22)
Other destinations	0.02 (0.03)	0.02 (0.03)	0.01 (0.03)	0.02 (0.05)
Share discharged to skilled nursing facilities	0.25 (0.11)	0.26 (0.11)	0.25 (0.12)	0.02 (0.04)
No. of days in skilled nursing facilities¶	7.3 (2.4)		5.0 (2.1)	0.6 (0.8)
Episode spending in skilled nursing facilities¶	3,164 (1,190)		2,081 (938)	180 (223)
Total episode spending	22,662 (3,544)		20,765 (2,999)	39,454 (14,243)
No. of CJR-eligible episodes	221,814	120,967	34,804	21,126
No. of MSAs	104	104	104	103

Table reports mean and SD for the summary statistics of CJR-eligible LEJR patients at the MSA level for the control MSAs. All outcomes are measured during the episode. All measures are based on the 2016–2017 Medicare and HCCI claims data. Episodes admitted between 1 April 2016 and 15 September 2017 are included. Number of MSAs count the number of MSAs with any CJR-eligible LEJR episode, among the final set of 104 control MSAs.

*Age is only available in 10-year bins in HCCI data: 45 to 54, 55 to 64, and so on.

†Dual eligibility indicator is not available in ESI data.

‡DRG: diagnosis-related group, a patient-classification system used in hospital reimbursements.

§Institutional postacute care includes skilled nursing facilities, long-term-care hospitals, and inpatient rehabilitation facilities.

¶Number of days and spending in skilled nursing facilities are averaged across all episodes, not just episodes with skilled nursing facility use.

spending (\$2,081 for MA patients vs. \$3,164 for TM patients) at skilled nursing facilities.

Finally, we use the HCCI data to examine the treatment of CJR-eligible patients in the commercial (under 65) privately insured population covered with ESI by one of the HCCI insurers. A comparison of columns 3 and 4 indicates that ESI patients are naturally much younger (ESI patients in our sample are under 65 by definition, while 90% of MA patients are 65 and over). Interestingly, despite being younger, average episode

spending on ESI patients is about twice as large than on MA patients (\$40,000 compared to \$20,000), presumably reflecting the much higher rates paid by insurers relative to Medicare (41).

ESI patients are also associated with a very different distribution over postdischarge destinations relative to MA or TM patients. In particular, institutional postacute care use is rare among ESI patients; only 2% are discharged to institutional postacute care compared to about 30% of patients in TM and in MA. The low usage of institutional postacute care among ESI

Table 2. Impact of CJR on TM and MA patients

Outcome	TM				MA			
	Control mean	Treatment effect	95% CI	P value	Control mean	Treatment effect	95% CI	P value
Share discharged to Institutional postacute care	0.313	−0.034	[−0.051, −0.017]	0.0001	0.283	−0.033	[−0.050, −0.015]	0.0003
Home (with home health care)	0.339	0.004	[−0.031, 0.039]	0.81	0.365	0.004	[−0.033, 0.040]	0.843
Home (without home health care)	0.329	0.042	[0.007, 0.077]	0.02	0.336	0.042	[0.004, 0.080]	0.030
Other destinations	0.020	−0.004	[−0.008, −0.0001]	0.05	0.016	−0.003	[−0.007, 0.001]	0.184
No. of CJR episodes	2,133	−10	[−155, 135]	0.89	1,163	−5	[−125, 116]	0.937
No. of treatment/control/all MSAs	67/104/171				67/104/171			

Table reports results from estimating Eq. 1 using the CMS data. Specifically, it reports MSA-level estimates from a regression of the row outcome on an indicator for CJR, controlling for strata fixed effects, and two lags of the outcome variable. CI and P value are based on heteroskedasticity robust SEs. In CMS TM and CMS MA columns, the outcomes are measured using all LEJR admissions between 1 April 2016 and 15 September 2017 that would have qualified for CJR, among TM and MA enrollees, respectively. See notes to Table 1 for variable definitions.

patients limits the scope and magnitude of any potential spillover to ESI patients, since decreased rates of discharge to institutional postacute care is the main channel by which the payment reform affects the treatment of TM patients (1–3, 5). We therefore do not explore spillovers to the ESI population.

Average Spillover Effects. Table 2 shows the estimated impact of the payment reform on discharge location for CJR-eligible patients. The left-hand side shows the direct effect on targeted TM patients; this has already been extensively studied using these same data (1–3, 5). Consistent with these prior studies, we find that the payment reform decreases the probability a TM patient is discharged to institutional postacute care. Specifically, we estimate a 3.4 percentage point (95% confidential interval –5.1 to –1.7 percentage points) decline in the probability a TM patient is discharged to institutional postacute care, or about 10% relative to the control mean. As in prior work, there is no effect on the probability of being discharged home with home health care and an 4.2 percentage point increase in discharges home (without home health), which is of similar magnitude to the decline in discharges to institutional postacute care.^{††} Also consistent with prior work, there is no evidence of an impact of CJR on the volume of CJR-eligible episodes.

The right-hand side shows the spillover effects of the payment reform on the discharge locations of CJR-eligible patients covered by MA. The spillover effects are strikingly similar to the direct effects. The payment reform for TM patients reduces the share of MA patients discharged to institutional postacute care by 3.3 percentage points (95% confidence interval –5.0 to –1.5 percentage points). This is the same sign and similar magnitude (in both levels and proportions) to the direct effect on TM patients; Wilcock et al. (6) and Meyers et al. (7) previously documented comparable findings.

As a confirmatory check on our spillover estimates using the CMS data, we also estimated the spillover effects on MA patients in the HCCI data. We found that the results were broadly similar, but less precisely estimated. Further analysis suggested the imprecision reflected the smaller sample size, rather than differential spillovers for the subset of insurers represented in the HCCI data; *SI Appendix, Appendix C* describes these analyses in more detail.

Heterogeneity across Hospitals. We examine heterogeneity across hospitals in both direct and spillover effects. These analyses inform our discussion of underlying mechanisms. We focus this analysis on the primary outcome and key margin of response: the probability a patient is discharged to institutional postacute care.

To estimate hospital-specific treatment effects, we follow Einav et al. (4) and use the following regression:

$$y_{h,(2016-2017)} = \beta_0 + \sum_h \beta_{1,h} BP_h + \beta_2 y_{h,2014} + \beta_3 y_{h,2013} + \delta_{s(h)} + \epsilon_h, \quad [2]$$

where h denotes a hospital, $y_{h,(2016-2017)}$ denotes the share of patients in 2016 and 2017 that were discharged to institutional postacute care, BP_h is an indicator for the hospital being randomly assigned to bundled payments, and $\beta_{1,h}$ is the hospital-specific treatment effect. Similar to our baseline specification, we include (hospital-level) lagged outcomes as covariates to improve statistical power. As before, we include strata fixed effects because randomization was conducted within strata. In the

regressions, we weight hospital-level observations by the number of CJR-eligible episodes to reduce noise from hospitals with very few episodes.

We estimate Eq. 2 separately on the TM patients and on the MA patients, producing hospital-specific estimates of the impact of the payment reform on the share of TM patients discharged to institutional postacute care and the share of MA patients discharged to institutional postacute care. The estimates vary quite a bit across hospitals, presumably reflecting both true underlying heterogeneity across hospitals in the treatment effect as well as nontrivial sampling variation, especially in the context of some of the smaller hospitals.

More importantly, we estimate a positive and large correlation of 0.60 (SE 0.026) between the TM and MA hospital-specific estimates (*SI Appendix, Table S2*). Given the noise in the TM and MA estimates, we would not expect the correlations to be close to one, so we view a correlation estimate of 0.60 as indicating a fairly strong positive relationship between the TM and MA treatment effects at the hospital level. This is not surprising. As we discuss in more detail in the next section, any of the mechanisms that might produce same-signed spillover effects would predict that hospital-level estimates of the direct effects and indirect effects should be positively correlated.

To interpret the $\beta_{1,h}$ s from Eq. 2 as heterogeneous causal effects of the payment reform, the (admittedly strong) identifying assumption is that, conditional on the covariates, there are no hospital-specific time trends, and thus any heterogeneity in the change in outcomes across hospitals reflects heterogeneous treatment effects. To probe the sensitivity to this assumption, we estimate an alternative specification where we include hospital-specific linear time trends as controls. These time trends are identified from the hospital-specific outcomes in 2013 and 2014. In this specification, the identifying assumption is that, conditional on covariates, there are no hospital-specific deviations from the time trend. As a result, the correlation between the estimates is reduced to 0.17 (SE 0.039) but still positive and statistically significant (*SI Appendix, Table S2*).

We next examine how hospital-specific direct and spillover effects vary by specific hospital characteristics. This analysis is necessarily more speculative than the result on the existence of spillovers, since the hospital characteristics we examine are not randomly assigned. Nonetheless, it provides interesting insights on the characteristics of hospitals associated with greater direct and spillover effects.

Table 3 reports the average hospital-specific treatment effects by hospital characteristics. Fig. 2 then visually illustrates how closely related the TM effect is to the MA effect by plotting each one of the effects reported in Table 3 for MA against its analogous estimate for TM. Perhaps unsurprisingly, given the previous results, hospital-level treatment effects and spillover effects often point in the same direction within the binary groups defined by hospital characteristics.

In particular, both the direct and spillover effects of bundled payment reform are higher for hospitals with an above-median volume of TM CJR-eligible episodes compared to hospitals with a below-median number of TM CJR-eligible episodes.^{§§} Indeed, the hospitals with above-median volume of TM patients seem to account for the entirety of the direct effect as well as the entirety of the spillover effect. Concentration of the direct effect and the

^{††}This can be either because the patients who would have been sent to institutional postacute care are being sent home without home health or because there is a cascading effect where the patients who would have been sent to institutional postacute care are being sent home with home health, and patients who would have been sent home with home health are now being sent home without home health supports. A cascading effect seems (to us) more likely, but we cannot differentiate between these two channels.

^{§§}For TM patients, we estimate that bundled payment reduces the share of patients discharged to institutional postacute care by 0.03 percentage points on average for hospitals with above median TM CJR volume, compared to a 0.05 percentage point increase for hospitals with below median volume; these estimates of statistically distinguishable ($P = -0.001$). For MA patients, the impact of bundled payments is, on average, a –0.019 percentage point decline for hospitals with above-median TM CJR volume, compared to a 0.019 percentage point increase for hospitals with below-median volume; again, these estimates are statistically distinguishable ($P = 0.001$).

Table 3. Heterogeneity in impact by hospital characteristics

Hospital characteristics	TM			MA		
	Yes	No	<i>P</i> value of difference	Yes	No	<i>P</i> value of difference
Above-median TM CJR volume	−0.030 (0.093)	0.053 (0.160)	0.001	−0.019 (0.097)	0.019 (0.134)	0.005
Above-median TM CJR share	−0.021 (0.100)	−0.027 (0.111)	0.278	−0.004 (0.115)	−0.014 (0.104)	0.166
Nonprofit	−0.021 (0.101)	−0.025 (0.108)	0.348	−0.013 (0.103)	0.002 (0.123)	0.100
Teaching	−0.044 (0.103)	−0.017 (0.103)	0.054	−0.018 (0.117)	−0.008 (0.107)	0.292
Above-median no. of beds	−0.017 (0.101)	−0.035 (0.109)	0.033	−0.005 (0.105)	−0.023 (0.114)	0.057
Above-median quality score	−0.039 (0.094)	0.010 (0.114)	0.001	−0.024 (0.097)	0.018 (0.122)	0.001
Geographical regions						
South	−0.037 (0.106)	−0.014 (0.101)	0.019	−0.019 (0.122)	−0.006 (0.101)	0.154
Northeast	−0.017 (0.124)	−0.024 (0.099)	0.340	0.010 (0.112)	−0.015 (0.107)	0.040
West	−0.017 (0.089)	−0.024 (0.107)	0.247	−0.019 (0.098)	−0.006 (0.112)	0.121
Midwest	−0.010 (0.095)	−0.027 (0.106)	0.065	−0.003 (0.094)	−0.012 (0.112)	0.216

The outcome is share discharged to institutional postacute care in all panels. Table is based on the hospital-specific treatment effects from estimating Eq. 2 using the CMS data. It reports these hospital-specific treatment effects by hospital characteristics for TM and MA, respectively. For each hospital characteristics in each row, the table reports the mean and SD of the hospital-specific treatment effects, separately for hospitals that satisfy the given characteristic (Yes) and hospitals that do not satisfy the given characteristic (No). The *P* value from a two-samples *t* test is reported. All regressions and summary statistics are weighted by the number of CJR (TM or MA) episodes in hospital.

spillover effect in high-volume hospitals is similar to Clemens et al.'s (48) finding that the linkage between private insurance payments and public insurance payments was also larger for larger physician groups.

Hospitals with an above-median quality score also have larger treatment and spillover effects than hospitals with below-median scores (−0.039 vs. 0.010, *P* < 0.001 for TM, −0.024 vs. 0.018, *P* < 0.001 for MA).¹¹ On the other hand, hospital-specific treatment and spillover effects are less closely aligned when comparing hospitals by their nonprofit status, teaching status, number of hospital beds, or the share of TM CJR patients (among TM and MA CJR patients). In separate analysis, we explore the role of doctors in spillovers by looking at changes in other non-LEJR procedures performed by LEJR doctors but do not detect any spillovers. *SI Appendix, Table S3* reports this exercise.

Discussion and Implications

We use a payment reform in TM, which was randomly applied to some markets but not others, to study spillovers of healthcare payment reform. We find spillovers of the same sign and similar magnitude on privately insured MA patients. Naturally, our findings are specific to our setting; the existence, sign and magnitude of any spillovers may well vary across contexts.

¹¹ Quality is measured by a composite quality score (CQS), which is the official quality measure that determines bonus payments in CJR. The measure ranges from 0 to 20 points, and hospitals must score at least 5 points to be eligible for bonus payments. Up to 10 points are given based on a hospital's quality performance percentile on a complication measure for total hip arthroplasty and total knee arthroplasty; up to 8 points are given based on a standardized national patient experience survey; up to 2 points are given for submitting the patient reported outcomes and risk variable data. Finally, up to 1.8 points can be added to the final score for improvement in either of the first two measures relative to the previous performance year, as long as the final score does not exceed 20. See <https://innovation.cms.gov/Files/x/cjr-qualsup.pdf> for details on this measure.

Understanding the economic forces driving the spillovers we detected is a useful first step in thinking about how it may generalize to other settings.

A priori, spillovers (if they exist) may be of the same sign or opposite sign from the direct effects. Same-signed spillovers could arise if changes in healthcare delivery arise from a “high fixed cost, low marginal cost” underlying mechanism. For example, in our setting, if hospitals build a computer model that

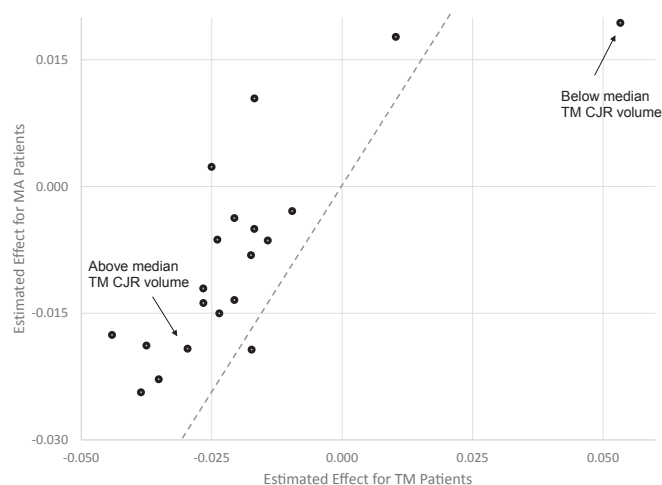


Fig. 2. Correlation between estimated effects for MA and TM patients. Figure plots the estimates reported in Table 3, which are hospital-specific treatment effects by hospital characteristics. The estimated effects for MA patients are shown on the y axis, and the estimated effects for TM patients are shown on the x axis. The dashed line is the 45° line.

uses available data to guide discharge decisions for targeted TM patients, this guidance could be extended to nontargeted MA patients with little additional cost. Same-signed spillovers could also arise if providers are constrained to provide similar care to patients irrespective of insurance plan; these constraints could reflect operational constraints (e.g., providers do not know a patients' insurance type) or provider preferences (e.g., doctors might consider it unethical to treat patients differentially).

Opposite-signed spillovers could arise if the cost of delivering a particular type of care is increasing in the number of patients treated, what economists call an upward sloping supply curve for care delivery. For example, in our setting, if the supply of beds at institutional postacute care facilities is an important constraint on discharges, a decrease in discharges of TM patients might be offset by an increase in discharges of MA patients. Similarly, if hospitals have a fixed number of personnel who can help patients transition to their homes, then an increase in the share of TM patients discharged home, which would occupy these resources, might be offset by fewer discharges home of MA patients. While the same-signed estimates we find do not allow us to completely rule out the existence of such effects, they indicate that such effects are not large enough to offset factors that push in the other direction.

Several pieces of evidence suggest that a fixed-cost model of changing healthcare provision is a key driver of spillovers in our setting. Such a model would predict the same-signed and similar-magnitude spillovers that we find. It would also predict our finding that both the direct and spillover effects are concentrated in hospitals with a larger number of directly affected patients. This type of fixed-cost explanation is also supported by qualitative survey evidence of hospital administrators on how hospitals in the treatment group responded to the payment reform (49); the administrators report reducing discharges to institutional postacute care by using risk stratification and by forming networks of preferred skilled nursing facilities to influence quality and costs, conditional on discharge. Of course, the evidence supporting a fixed-cost model does not rule out a role for constraints on treating similar patients with different insurance differently, but we are unable to examine this directly. More broadly, given the variety of different potential mechanisms that may be at play, spillovers may well differ across contexts.

Nonetheless, the results from our setting have several general implications. First, they suggest that the large empirical literature on the impacts of public healthcare policies may miss an

important set of costs and benefits of these policies, since the literature focuses almost exclusively on the direct impacts on those covered by the policies. This includes estimates of the impact of public health insurance coverage [see Finkelstein et al. (50) for a recent review] and the impact of public health insurance payment policies (22, 51–54). Second, the finding of substantial spillovers suggests that in many empirical settings, estimates of the direct effects of health insurance policies may be biased since “control group” patients not directly affected by the policy may be indirectly affected. Examples of the type of work subject to these concerns are surveyed in Finkelstein et al. (50).

Third, our results raise important questions about the optimal design of healthcare payment policies. When there are multiple insurance providers, spillovers imply that single insurers may have suboptimal incentives to change provider behavior. Prior theoretical work has explored how common agency problems—such as free-riding, coordination failures, or capacity constraints—can lead to inefficient payment policies (55–57). A key assumption underlying these models is that the health insurance payments for one group of patients may affect provider behavior for other patients. Our findings provide empirical support for this assumption.

Most broadly, our findings suggest that researchers and policymakers should take a broader perspective when considering the design of healthcare policy. Even if the policy objective is narrowly focused on TM patients, the fact the providers change their treatment of nontargeted patients suggests that these nontargeted patients need to be accounted for in analyses of a policy's impact or optimal design. How exactly healthcare payment design should be affected depends on the economic model underlying the spillovers. Our findings highlight the need for further work on this topic.

ACKNOWLEDGMENTS. We thank Sam Wang for extraordinary research assistance and Leila Agha, Adam Sacarny, and Jonathan Skinner for helpful comments. We gratefully acknowledge support from J-PAL North America's Health Care Delivery Initiative (A.F. and N.M.), the National Institute of Aging grant P01AG019783-15, and the Laura and John Arnold Foundation (L.E., A.F., and N.M.). The authors acknowledge the assistance of the HCCI and its data contributors, Aetna, Humana, and UnitedHealthcare, in providing the claims data analyzed in this study. The data used in this article are not publicly available. To obtain the CMS data, there is a standard application process for applying for these data (described at <https://www.resdac.org/>). To inquire about potential access to the HCCI data, emails should be directed to data@healthcostinstitute.org.

1. A. Finkelstein, Y. Ji, N. Mahoney, J. Skinner, Mandatory Medicare bundled payment program for lower extremity joint replacement and discharge to institutional postacute care: Interim analysis of the first year of a 5-year randomized trial. *JAMA* **320**, 892–900 (2018).
2. The Lewin Group, “CMS comprehensive care for joint replacement model: Performance year 1 evaluation report” (The Lewin Group, Falls Church, VA, 2018).
3. M. L. Barnett et al., Two-year evaluation of mandatory bundled payments for joint replacement. *N. Engl. J. Med.* **380**, 252–262 (2019).
4. L. Einav, A. Finkelstein, Y. Ji, N. Mahoney, Voluntary regulation: Evidence from Medicare payment reform (NBER Working Paper 27223, National Bureau of Economic Research, Cambridge, MA, 2019).
5. D. A. Haas, X. Zhang, R. S. Kaplan, Z. Song, Evaluation of economic and clinical outcomes under Centers for Medicare & Medicaid Services mandatory bundled payments for joint replacements. *JAMA Intern. Med.* **179**, 924–931 (2019).
6. A. D. Wilcock, M. L. Barnett, J. M. McWilliams, D. C. Grabowski, A. Mehrotra, Association between Medicare's mandatory joint replacement bundled payment program and post-acute care use in Medicare Advantage. *JAMA Surg.* **155**, 82–84 (2019).
7. D. J. Meyers, C. M. Kosar, M. Rahman, V. Mor, A. N. Trivedi, Association of mandatory bundled payments for joint replacement with use of postacute care among Medicare Advantage enrollees. *JAMA Netw. Open* **2**, e1918535 (2019).
8. A. Finkelstein, The aggregate effects of health insurance: Evidence from the introduction of Medicare. *Q. J. Econ.* **122**, 1–37 (2007).
9. K. Baicker, D. Staiger, Fiscal shenanigans, targeted federal health care funds, and patient mortality. *Q. J. Econ.* **120**, 345–386 (2005).
10. L. C. Baker, K. S. Corts, HMO penetration and the cost of health care: Market discipline or market segmentation? *Am. Econ. Rev.* **86**, 389–394 (1996).
11. S. Glied, J. G. Zivin, How do doctors behave when some (but not all) of their patients are in managed care? *J. Health Econ.* **21**, 337–353 (2002).
12. K. Baicker, M. E. Chernew, J. A. Robbins, The spillover effects of Medicare managed care: Medicare Advantage and hospital utilization. *J. Health Econ.* **32**, 1289–1300 (2013).
13. K. Baicker, J. A. Robbins, Medicare payments and system-level health-care use: The spillover effects of Medicare managed care. *Am. J. Health Econ.* **1**, 399–431 (2015).
14. J. Clemens, J. D. Gottlieb, In the shadow of a giant: Medicare's influence on private physician payments. *J. Polit. Econ.* **125**, 1–39 (2017).
15. E. Miguel, M. Kremer, Worms: Identifying impacts on education and health in the presence of treatment externalities. *Econometrica* **72**, 159–217 (2004).
16. B. Crépon, E. Duflo, M. Gurgand, R. Rathelot, P. Zamora, Do labor market policies have displacement effects? Evidence from a clustered randomized experiment. *Q. J. Econ.* **128**, 531–580 (2013).
17. J. P. Newhouse, *Free for All?* (Harvard University Press, 1993).
18. K. Baicker, A. Finkelstein, “Insuring the uninsured” (J-Pal Policy Briefcase, 2014).
19. C. Michalopoulos et al., “The Accelerated Benefits Demonstration and Evaluation Project: Impacts on health and employment at twelve months” (MDRC, 2011).
20. J. Cubanski, T. Neuman, M. Freed, “The facts on Medicare spending and financing” (Henry J. Kaiser Family Foundation, San Francisco, 2019).
21. Kaiser Family Foundation, “Total number of Medicare beneficiaries” (Henry J. Kaiser Family Foundation, San Francisco, 2019).
22. D. Cutler, The incidence of adverse medical outcomes under prospective payment. *Econometrica* **63**, 29–50 (1995).
23. D. M. Cutler, K. Ghosh, The potential for cost savings through bundled episode payments. *N. Engl. J. Med.* **366**, 1075–1077 (2012).
24. E. S. Fisher, Medicare's bundled payment program for joint replacement: Promise and peril? *JAMA* **316**, 1262–1264 (2016).
25. Centers for Medicare & Medicaid Services (CMS), “Bundled Payments for Care Improvement (BPCI) initiative: General information” (CMS, 2019).

26. J. Cromwell, D. A. Dayhoff, A. H. Thomaian, Cost savings and physician responses to global bundled payments for Medicare heart bypass surgery. *Health Care Financ. Rev.* **19**, 41–57 (1997).
27. J. P. Doran, S. J. Zabinski, Bundled payment initiatives for Medicare and non-Medicare total joint arthroplasty patients at a community hospital: Bundles in the real world. *J. Arthroplasty* **30**, 353–355 (2015).
28. L. A. Dummit *et al.*, Association between hospital participation in a Medicare bundled payment initiative and payments and quality outcomes for lower extremity joint replacement episodes. *JAMA* **316**, 1267–1278 (2016).
29. A. S. Navathe *et al.*, Cost of joint replacement using bundled payment models. *JAMA Intern. Med.* **177**, 214–222 (2017).
30. B. C. Maughan *et al.*, Medicare's Bundled Payments for Care Improvement initiative maintained quality of care for vulnerable patients. *Health Aff.* **38**, 561–568 (2019).
31. B. E. Landon *et al.*, Analysis of Medicare Advantage HMOs compared with traditional Medicare shows lower use of many services during 2003–09. *Health Aff.* **31**, 2609–2617 (2012).
32. J. Z. Ayanian, B. E. Landon, R. C. Saunders, L. G. Pawlson, J. P. Newhouse, Medicare beneficiaries more likely to receive appropriate ambulatory services in HMOs than in traditional Medicare. *Health Aff.* **32**, 1228–1235 (2013).
33. M. Duggan, J. Gruber, B. Vabson, The consequences of health care privatization: Evidence from Medicare Advantage exits. *Am. Econ. J. Econ. Policy* **10**, 153–186 (2018).
34. V. Curto, L. Einav, A. Finkelstein, J. Levin, J. Bhattacharya, Health care spending and utilization in public and private Medicare. *Am. Econ. J. Appl. Econ.* **11**, 302–332 (2019).
35. J. Brown, M. Duggan, I. Kuziemko, W. Woolston, How does risk selection respond to risk adjustment? New evidence from the Medicare Advantage program. *Am. Econ. Rev.* **104**, 3335–3364 (2014).
36. Centers for Medicare and Medicaid Services (CMS), HHS, Medicare program; Hospital inpatient prospective payment systems for acute care hospitals and the long-term care hospital prospective payment system policy changes and fiscal year 2016 rates; Revisions of quality reporting requirements for specific providers, including changes related to the electronic health record incentive program; Extensions of the Medicare-dependent, small rural hospital program and the low-volume payment adjustment for hospitals. Final rule; Interim final rule with comment period. *Fed. Regist.* **80**, 49325–49886 (2015).
37. C. C. Afendulis, M. E. Chernew, D. P. Kessler, The effect of Medicare Advantage on hospital admissions and mortality (NBER Working Paper 19101, National Bureau of Economic Research, Cambridge, MA, 2013).
38. A. Kumar *et al.*, Comparing post-acute rehabilitation use, length of stay, and outcomes experienced by Medicare fee-for-service and Medicare Advantage beneficiaries with hip fracture in the United States: A secondary analysis of administrative data. *PLoS Med.* **15**, e1002592 (2018).
39. MedPAC, “Ensuring the accuracy and completeness of Medicare Advantage encounter data” in *Report to the Congress: Medicare and the Health Care Delivery System* (2019), pp. 203–239.
40. G. Jacobson, G. Casillas, A. Damico, T. Neuman, M. Gold, “Medicare Advantage 2016 spotlight: Enrollment market update” (Henry J. Kaiser Family Foundation, 2016).
41. Z. Cooper, S. V. Craig, M. Gaynor, J. Van Reenen, The price ain't right? Hospital prices and health spending on the privately insured. *Q. J. Econ.* **134**, 51–107 (2019).
42. E. R. Adrion *et al.*, Out-of-pocket spending for hospitalizations among nonelderly adults. *JAMA Intern. Med.* **176**, 1325–1332 (2016).
43. L. Baker, M. K. Bundorf, A. Devlin, D. P. Kessler, Why don't commercial health plans use prospective payment? *Am. J. Health Econ.* **5**, 465–480 (2019).
44. Z. Cooper *et al.*, Hospital prices grew substantially faster than physician prices for hospital-based care in 2007–14. *Health Aff.* **38**, 184–189 (2019).
45. Z. Cooper, S. Craig, C. Gray, M. Gaynor, J. Van Reenen, Variation in health spending growth for the privately insured from 2007 to 2014. *Health Aff.* **38**, 230–236 (2019).
46. D. Dranove, C. Garthwaite, C. Ody, Health spending slowdown is mostly due to economic factors, not structural change in the health care sector. *Health Aff.* **33**, 1399–1406 (2014).
47. D. Pelech, T. Hayford, Medicare Advantage and commercial prices for mental health services. *Health Aff.* **38**, 262–267 (2019).
48. J. Clemens, J. D. Gottlieb, T. L. Molnár, Do health insurers innovate? Evidence from the anatomy of physician payments. *J. Health Econ.* **55**, 153–167 (2017).
49. J. M. Zhu, V. Patel, J. A. Shea, M. D. Neuman, R. M. Werner, Hospitals using bundled payment report reducing skilled nursing facility use and improving care integration. *Health Aff.* **37**, 1282–1289 (2018).
50. A. Finkelstein, N. Mahoney, M. J. Notowidigdo, What does (formal) health insurance do, and for whom? *Annu. Rev. Econ.* **10**, 261–286 (2018).
51. J. Clemens, J. D. Gottlieb, Do physicians' financial incentives affect medical treatment and patient health? *Am. Econ. Rev.* **104**, 1320–1349 (2014).
52. K. Ho, A. Pakes, Hospital choices, hospital prices, and financial incentives to physicians. *Am. Econ. Rev.* **104**, 3841–3884 (2014).
53. P. J. Eliason, P. Grieco, R. C. McDevitt, J. W. Roberts, Strategic patient discharge: The case of long-term care hospitals. *Am. Econ. Rev.* **108**, 3232–3265 (2018).
54. L. Einav, A. Finkelstein, N. Mahoney, Provider incentives and healthcare costs: Evidence from long-term care hospitals. *Econometrica* **86**, 2161–2219 (2018).
55. J. Glazer, T. G. McGuire, Multiple payers, commonality and free-riding in health care: Medicare and private payers. *J. Health Econ.* **21**, 1049–1069 (2002).
56. B. Frandsen, M. Powell, J. B. Rebitzer, Sticking points: Common-agency problems and contracting in the US healthcare system. *RAND J. Econ.* **50**, 251–285 (2019).
57. P. T. Léger, R. J. Town, J. Wu, “A theory of geographic variations in medical care” (2019). https://pdfs.semanticscholar.org/1ca4/c54f654ccf61d4d5098c3277e42c62568149.pdf?_ga=2.80319070.877234021.1593703033-88243013.1584979178. Accessed 12 March 2020.