THE EXTENT OF EXTERNALITIES FROM MEDICARE PAYMENT POLICY

ALICE J. CHEN
MICHAEL R. RICHARDS
CHRISTOPHER M. WHALEY
XIAOXI ZHAO

ABSTRACT

Medicare accounts for roughly 20 percent of medical expenditures in the United States and is the dominant payer for many treatments. Consequently, Medicare payment policy may have diffuse consequences. Using a contemporary bundled payment reform (the "CJR" program) and a difference-in-differences research design, we estimate Medicare's spillover reach. We find that altered treatment decisions for targeted joint replacement procedures are closely, though not perfectly, mirrored between traditional Medicare, Medicare Advantage, and the nonelderly commercially insured populations. Results for untargeted procedures performed by CJR-affected physicians also show suggestive evidence consistent with a secondary spillover effect; however, this behavior change does not extend to less related procedures. Our findings align with the "norms hypothesis" for physician decision making but do not imply rigid and uniform treatment choices. Instead, key decision nodes appear to gain greater salience under Medicare's new incentive structure, which leads to revised treatment choices for different payer-procedure combinations. Ignoring the breadth of externalities from Medicare policies risks understating their social welfare impact.

KEYWORDS: bundled payment, episode-based payment, Comprehensive Care for Joint Replacement, value-based purchasing, Affordable Care Act
JEL CLASSIFICATION: H44, I11, I13, I18

I. Introduction

Externalities represent fundamental concepts of economic activity, but quantifying them is often absent or incomplete when evaluating many policies. Moreover, mixed markets, with both private and public components, can be important sources of such externalities since policies affecting one segment of the market (e.g., the public side) can generate spillovers for the other (e.g., the private side). A prominent example of such a context is health care. Multiple public and private insurers often act as intermediaries for patients and establish their own bespoke incentive schemes for local health-care providers (Relman and Reinhardt 1986; Hughes Tuohy, Flood, and Stabile 2004). In the US, the Medicare public insurance

Alice J. Chen, University of Southern California. Michael R. Richards (corresponding author, michael_richards@baylor.edu), Baylor University. Christopher M. Whaley and Xiaoxi Zhao, RAND Corporation.

Electronically published April 4, 2022.

American Journal of Health Economics, volume 8, number 2, spring 2022.
 https://doi.org/10.1086/718769

program is responsible for more than one out of every five medical expenditure dollars,[1] and it directly or indirectly covers nearly 60 million individuals.[2] The program's size and prominence suggest that Medicare policy could wield significant influence over private actors within US health-care markets—especially since nearly all hospitals and related health-care providers will have a blend of Medicare and non-Medicare business.

Accompanying research demonstrates that commercial (i.e., private) health insurers partially anchor their respective provider service fees to Medicare's administered price schedule (White 2013; Clemens and Gottlieb 2017; Clemens, Gottlieb, and Molnár 2017; Trish et al. 2017; Cooper et al. 2019), and similarly, changes in Medicare reimbursement levels seem to alter providers' effort devoted to the non-Medicare, private market (Sloan, Morrisey, and Valvona 1988; Yip 1998; He and Mellor 2012; White 2014). Recent work also finds that Medicare regulations can lead to changes in provider decisions for non-Medicare patients (Geruso and Richards 2021; Richards, Seward, and Whaley 2021).[3] Yet, the potentially broad impacts of the Medicare program go beyond price and rule setting. For instance, Medicare is often the source of payment model experimentation for health-care providers—a role that has accelerated since the passing of the Affordable Care Act (ACA) in 2010 (Abrams et al. 2015).

Traditional fee-for-service (FFS) payments for providers are widely believed to contribute to inefficiencies and excessive spending; see Burns and Pauly (2018) for a recent commentary. Yet making providers more cost-conscious through alternative incentive structures has proven difficult—partly because of the small influence of any individual insurer within the multi-payer landscape (Frandsen, Powell, and Rebitzer 2019). To overcome these inherent challenges, the traditional FFS Medicare program has at times unilaterally introduced new approaches to provider reimbursements. The direct effects from these Medicare policies in terms of curbing Medicare spending and/or improving health outcomes of Medicare beneficiaries are of first-order importance, but indirect effects (i.e., spillovers onto

1 These and related annual national spending statistics are provided and regularly updated by the Centers for Medicare and Medicaid Services and can be found here: https://www.cms.gov/Research-Statistics-Data-and-Systems/Statistics-Trends-and-Reports/NationalHealthExpendData/NHE-Fact-Sheet.

2 Medicare enrollment counts for traditional Medicare and Medicare Advantage are available from the Kaiser Family Foundation. Note, there are roughly 40 million beneficiaries in traditional (fee-for-service) Medicare and another 20 million Medicare-eligible individuals that have opted for the privatized Medicare Advantage route. These statistics can be found here: https://www.kff.org/medicare/state-indicator/total-medicare-beneficiaries.

3 We also note a modest but growing literature that has explored the reverse direction for spillover influence among payers (i.e., other insurers affecting Medicare). A few studies have taken interest in how the growth in the Medicare Advantage program affects the care and spending for traditional Medicare patients (Baicker, Chernew, and Robbins 2013; Baicker and Robbins 2015; Callison 2016), and other recent research has examined Medicaid (Bond and White 2013; Joynt et al. 2013; Joynt et al. 2015; McInerney, Mellor, and Sabik 2017; Glied and Hong 2018; Carey, Miller, and Wherry 2020) as well as non-Medicare commercial insurance (He, McInerney, and Mellor 2015; Richards and Tello-Trillo 2019) demand and contracting externalities on the Medicare fee-for-service population. Though indirectly related to our context, the findings from this complementary work help underscore the frequency and importance of cross-payer influences within multi-payer US health-care markets.

other payers and/or treatments) could be comparable in magnitude and consequently reveal more diffuse social welfare implications for a given payment reform. To date, there is considerable evidence related to the former (i.e., direct effects) but scarce evidence for the latter (i.e., indirect effects) when it comes to Medicare reimbursement policy.

In this paper, we focus our attention on Medicare's Comprehensive Care for Joint Replacement (CJR) Model that was announced in mid-2015 and began in 2016. The CJR program is devoted to lower extremity joint replacements (LEJRs), that is, hip and knee arthroplasty for beneficiaries enrolled in traditional Medicare. While the scope of the payment reform is narrow, LEJRs are the most common inpatient surgery for Medicare patients, with over 400,000 procedures and more than $7 billion in spending in 2014 alone. Moreover, the total cost to Medicare for a LEJR procedure and its associated recovery care demonstrates wide variation—for example, as much as double the average cost between the highest and lowest spending geographies, despite the public insurer setting its own national price schedule.[4] For these reasons, opportunities for greater care standardization to shrink the LEJR cost variance appear plentiful.

The CJR program (discussed in detail in Section II) relies on a version of bundled or "episode-based" reimbursement, where providers are held accountable for the costs of the entire episode of care (i.e., surgery and recovery periods). This contrasts with the FFS model where each increment of care is reimbursed and providers face no downside risk from choosing more expensive treatment options. Importantly, provider participation in the CJR program was not voluntary. Instead, a geographic randomization procedure sharply defined implementation areas, and all providers practicing within the designated areas were required to participate in the payment model. This procedure removes selection effects from nonrandom participation—an issue that has challenged evaluations and scale-ups for many other contemporary payment reforms.

With these programmatic features, we leverage data on the universe of LEJR procedures in Florida from the beginning of 2013 through the end of 2018 and a difference-in-differences (DD) and event time estimation framework to identify the spillover effects of a targeted Medicare policy onto other payers and services. First, we demonstrate that first-order CJR policy effects for Medicare LEJR cases in Florida align with the existing literature. Second, we determine whether the observed changes in treatment decisions extend to Medicare Advantage and non-Medicare commercially insured patients—consistent with a cross-payer spillover for the same procedure. We then further exploit the richness of our data to execute identical analyses on non-LEJR procedures performed by LEJR surgeons, as well as non-LEJR procedures typically outside of the scope of a LEJR surgeons. Doing so allows us to assess whether any policy-induced adjustments to provider clinical behavior and decision making extend beyond the policy's remit—an empirical test for the "norms hypothesis" of physician behavior (Newhouse and Marquis 1978; Glied and Zivin 2002; Frank and Zeckhauser 2007; Landon 2017).[5]

4 These background details as well as program details are provided by the Centers for Medicare and Medicaid Services and are available here: https://innovation.cms.gov/innovation-models/cjr.

5 The "norms hypothesis" for physician behavior states that tangible and/or cognitive costs limit the amount of treatment customization that providers will engage in for otherwise similar patients (Newhouse

Overall, our estimates show that provider behavior toward Medicare Advantage and commercially insured LEJR patients is meaningfully altered following the payment reform for traditional FFS Medicare patients. Specifically, we observe that operating charges decline by roughly 10 percent across each of the three payer groups, which is at least consistent with stronger cost-containment efforts (e.g., standardizing operating room staffing and use of supplies). There are also pronounced changes in post-acute-care treatment decisions across the three payer populations. Use of inpatient rehabilitation facilities (IRFs) following the LEJR hospital stay declined sharply for both Medicare and Medicare Advantage patients once the CJR program was announced (i.e., pre-implementation).[6] The lower reliance on IRFs is sustained over the three years of post-bundled payment implementation, with estimated payment reform effect sizes reflecting a 22 to 44 percent decrease in IRF use for traditional Medicare LEJR cases and a 50 to 63 percent decrease among Medicare Advantage cases when compared with their pre-CJR levels. The nonelderly, commercially insured LEJR patients do not show an immediate spillover effect, but instead begin to receive roughly 25 to 38 percent less IRF care after the CJR program has been in place for one year. Interestingly, while the use of home health post-acute services only suggestively climbs for traditional Medicare LEJR patients, both Medicare Advantage and commercially insured LEJR patients show marked increases in home health service utilization of 32 percent and 16 percent, respectively. For the nonelderly commercially insured group, specifically, much of the substitution toward home health is largely at the expense of discharging patients with no post-acute-care services following hospitalization (i.e., the policy spillover is increasing treatment intensity for this group). When examining non-LEJR procedures performed by LEJR surgeons, we find suggestive evidence of secondary spillover effects for these untargeted procedures; however, a wider breadth of procedures shows no provider behavior changes.

Our findings reveal statistically significant and economically meaningful externalities from the Medicare CJR program that have not been fully captured in previous studies.[7] The effects we document are more comprehensive and larger than the existing literature, perhaps because of our all-payer universe of patient encounters for each hospital, as well as our study setting that includes some of the highest Medicare spending areas in the country—that is, where legislated payment reforms may have their greatest bite. The strongest and most important effects localize to post-acute-care services, as opposed to within-hospital

and Marquis 1978; Glied and Zivin 2002; Frank and Zeckhauser 2007; Landon 2017). Well-identified empirical evidence for the presence or absence of such provider behavior is often lacking, however.

6 The cost of IRF services in the Medicare program alone totals approximately $8.7 billion per year (MedPAC 2021).

7 Specifically, we are aware of three very recent studies of the effect of the CJR program on LEJR cases for Medicare Advantage enrollees—Meyers et al. (2019); Wilcock et al. (2020); and Einav et al. (2020a)—which we discuss in detail within Section II. Our paper is distinct from these existing works in several respects. First, we have a complete universe of encounters for all payers within our health-care markets of interest. Second, we benefit from a longer time series (2013–18) and a deliberate estimation approach that transparently demonstrates the marked dynamics as well as persistency in providers' policy responses over time. Finally, and as previously noted, we broaden our analyses to non-LEJR procedures.

management (e.g., length of stay), and our estimates offer empirical support for the long-debated "norms hypothesis" of provider behavior. That said, providers do not seem to rigidly apply their treatment approaches to all procedure-payer combinations, and their post-reform changes in behavior evolve with time.

II. Background

Improvements to provider payment structures to achieve better patient outcomes and better align providers' incentives with those of patients, payers, and society have long been sought. While the approaches and nomenclature have evolved over time (e.g., pay-for-performance, value-based payment), the rates of adoption and ultimate impacts have been underwhelming (e.g., see Rosenthal and Frank 2006; Mullen, Frank, and Rosenthal 2010; Conrad 2016; Markovitz and Ryan 2017; Burns and Pauly 2018).

Medicare has recently used its large presence and financial leverage to overcome some of the inertia belonging to provider-payer contracting. As part of the Alternative Payment Models launched by the ACA, LEJR procedures (namely knee and hip arthroplasty) became focal targets of two separate bundled (episode-based) payment reforms: the voluntary Bundled Payments for Care Improvement (BPCI) program beginning in 2013 and the mandatory CJR program beginning in 2016 (Siddiqi et al. 2017).[8] Exhaustive descriptions of the CJR program can be found among existing studies as well as accompanying Centers for Medicare and Medicaid Services (CMS) resources (e.g., see footnote 4). We, in turn, only highlight key features of the payment reform that motivate our empirical aims and inform our identification strategy.

As previously remarked, the CJR program relies on a version of bundled service reimbursement, meaning that providers are incentivized to manage the full costs of the care episode, which spans the hospital stay and 90 days post-hospitalization. Medicare did not implement a posted price approach for LEJR bundled payments. Each hospital was ex ante given a target spending amount derived from its own historical Medicare LEJR costs and historical Medicare LEJR costs from its region. FFS payments would prevail during the course of treatment and recovery, and then an end-of-year reconciliation process would determine whether the hospitals' average total episode spending met or deviated from the benchmark amount. Higher spending would trigger financial penalties, and lower spending would trigger financial rewards from the public payer.[9] Additionally, provider participation was not voluntary, which is in stark contrast to other ACA-driven incentive structure

8 Even prior to Medicare's payment model experiments, researchers had advocated for a bundled approach for joint replacements as an opportunity to reduce health-care costs (Sood et al., 2011).

9 For the first year of the program, providers faced only upside risk, which they knew up front. As planned at the CJR program outset, downside risk was introduced beginning in 2017. Also, despite the enthusiasm for this Medicare policy development, several studies preemptively raised concerns over CJR implementation (Ellimoottil et al. 2016, 2017; Ibrahim, Kim, and McConnell 2016), and since its inception, others have highlighted an uneven financial impact across providers—particularly with respect to safety net and other lower-resourced hospitals (Kim et al. 2019; Navathe, Liao, Shah, et al. 2018; Thirukumaran, Glance, Cai, Balkissoon, et al. 2019; Thirukumaran, Glance, Cai, Kim, et al. 2019).

experiments emanating from Medicare—including other episode-based payment programs, such as the BPCI initiative. Outside of the CJR program, we are aware of only one other contemporary and large-scale bundled payment model with mandatory participation.[10] And finally, CJR implementation was not universal. As opposed to a blanket rollout, CMS administrators leveraged a randomization approach to select clusters of counties (i.e., metropolitan statistical areas or MSAs) across the US that would be exclusively incorporated into the CJR's five-year demonstration model. By design, this programmatic element creates precisely demarcated treatment and control classifications for health-care providers based on their practice location and the pre-implementation geographic randomization.

In terms of care delivery, existing evaluations of the CJR program find modest declines in spending overall (e.g., around 3 percent), with changes most pronounced in the use of post-acute-care institutional stays, such as skilled nursing facilities (SNFs) and IRFs (Finkelstein et al. 2018; Barnett et al. 2019; Haas et al. 2019). Relatedly, a convenience sample survey of affected orthopedic surgeons commented that post-acute-care services and spending became top priorities under the CJR regime (Sood et al. 2019).[11] Because of post-reconciliation transfers (i.e., CJR provider "bonuses"), however, estimates of net savings to the Medicare program range from zero to small (Finkelstein et al. 2018; Barnett et al. 2019; Haas et al. 2019; Einav et al. 2020b). But as previously noted, we have more limited evidence as to whether, and how, the introduction of the CJR program may have influenced provider behavior outside of Medicare LEJR cases, which could generate private gains for other market participants (i.e., Medicare Advantage and commercial insurers).

Among the three existing studies of non-Medicare LEJR patients, each focuses on a subset of Medicare Advantage patients found within the CMS MedPAR data. These data have the benefit of being national in scope but are also inherently incomplete records of the case mix, payer mix, and treatment behavior for a given hospital.[12] Using these data, Wilcock et al. (2020) asses the first year of the CJR program (2016) and find a small but symmetrical decline in the use of institutional post-acute-care facilities between traditional Medicare and Medicare Advantage patients. However, they are unable to detect a statistically significant change among specific types of post-acute-care facilities. Meyers et al. (2019) extend the analyses of Wilcock et al. (2020) by adding an additional year of data and incorporating more detailed post-acute-care databases. The authors likewise see modest reductions

10 In 2013, the state of Arkansas introduced an analogously structured initiative that focused on perinatal care and applied to all payers within the state (see Carroll et al. 2018).

11 These findings also align with what was previously documented for the voluntary bundled payment initiatives tied to LEJRs—though some BPCI participating hospitals also reported shorter length of stay and lower joint implant costs (Dummit et al. 2016; Iorio et al. 2016; Navathe et al. 2017).

12 For instance, because managed care organizations do not submit all claims to CMS, the Research Data Assistance Center (ResDAC) typically recommends removing these data from analyses, unless studying hospice care or a subset of added facility payments. Additional details regarding the limitations of analyses within the MedPAR database are described here: https://www.resdac.org/articles/differences-between-inpatient-and-medpar-files. Additionally, Meyers et al. (2019) formally assessed the post-acute discharge information from MedPAR and deemed it inaccurate, which motivated their inclusion of alternative data sources.

in institutional post-acute care across the two payers, though they also remark that there is no effect on home health receipt for Medicare Advantage patients and that the Medicare Advantage payer group typically demonstrates small effects overall. Conversely, Einav et al. (2020a) find a nearly equivalent decline (10–12 percent) in post-acute-care facility utilization across both payers but also no change in home health post-acute services for Medicare Advantage patients. Importantly, when the authors repeat their spillover analyses within a national claims database from three large insurers,[13] the Medicare Advantage results become equivocal—perhaps because of power issues and/or an incomplete census of Medicare Advantage patients belonging to a given hospital. Taken together, the current and small literature on this aspect of the CJR program lacks consensus on Medicare Advantage market effects. It also offers no empirical evidence for the nonelderly commercial market or for the non-LEJR cases performed within CJR-affected hospitals. Thus, the extent of Medicare bundled payment externalities remains an open empirical question.

III. Data

Our encounter-level data encompass the universe of inpatient discharge records from the state of Florida, which we obtained from the Florida Agency for Health Care Administration (AHCA). The detailed discharge records include a rich set of variables, such as diagnosis and procedure codes, type of insurance, patient demographic information, the specific facility and geographic location where the procedure was performed, ancillary care services provided during the inpatient stay, and post-hospitalization discharge disposition. We use the administrative data over a relatively long time series, starting in 2013 and ending in 2018. Unlike many other data resources, the data also capture all payers (including self-pay) in Florida markets over our full study period, rather than a specific payer or a subset of payers in the market.[14] Florida is also home to the second-largest Medicare population in the country and some of the highest Medicare spending areas,[15] and it is somewhat overrepresented in the CJR demonstration model since the randomization process placed meaningful weight on areas' historical Medicare spending levels. Notably, Florida accounts for 21 percent of the pre-policy LEJR cases among CJR treated MSAs nationwide.[16]

13 The authors use Health Care Cost Institute (HCCI) data, which include Aetna, Humana, and UnitedHealth Medicare Advantage enrollees from across the country.

14 For example, existing studies of the Medicare CJR program largely use claims data that contain data from the Medicare fee-for-service population. Other commonly used sources of data include data from three private insurers offered through the Health Care Cost Institute, or data from employer-sponsored plans provided by IBM Watson (formerly Truven).

15 State-level statistics on aggregate Medicare populations can be found here: https://www.kff.org/medicare/state-indicator/total-medicare-beneficiaries. Regional variation in spending per beneficiary can be found here: https://www.dartmouthatlas.org/interactive-apps/medicare-reimbursements/.

16 Meyers et al. (2019) report 269,723 Medicare LEJR procedures from 2013 to 2016Q1 in treated MSAs nationwide. Over that same period, Florida had 57,424 Medicare LEJR procedures in its treated MSAs.

IV. Empirical Strategy

A. TREATMENT DECISIONS FOR LEJR CASES

Our primary analyses focus on the specific procedures (i.e., LEJR cases) targeted by the CJR bundled payment program. We consequently restrict to inpatient stays with Medicare severity diagnosis-related group (MS-DRG) classifications 469 and 470 for hip and knee replacement hospitalizations—consistent with the program parameters and other studies in this literature. We then further restrict to three main payer groups: (1) Medicare FFS, (2) Medicare Advantage, and (3) non-Medicare commercial insurance. These three payer sources account for 93 percent of all LEJR cases observed in Florida in a typical year.[17]

Classifying a given provider (i.e., hospital) as part of the treatment or control group for our subsequent difference-in-differences (DD) analyses is accomplished by applying the sharp geographic boundaries set by CMS for the CJR payment model demonstration experiment. Using the list of included counties distributed by CMS, we map Florida counties belonging to the CJR program in Figure 1. The darker-shaded counties are those mandating CJR participation for all hospitals present in those areas. Hospitals operating in the lighter-shaded counties in Figure 1 are explicitly excluded from bundled payment model and therefore maintain the traditional FFS incentive structure for their Medicare LEJR patients. In this way, we ultimately capture the effects of the Medicare CJR program on both Medicare and non-Medicare patients and how they evolve over time by leveraging the quasi-random assignment of Florida hospitals to the bundled payment model.[18]

While we include all relevant Florida LEJR hospitals in the control group that were not assigned to a Medicare CJR treatment county, we do note that some MSAs (including in Florida) were not eligible to be randomized by CMS—typically because of low aggregate LEJR volumes. In our main analyses, we retain hospitals from these areas to enhance our estimates' precision; however, in Online Appendix B we repeat our estimation for key findings when excluding this subset of hospitals from the analytic sample. The findings and inferences are qualitatively unchanged, with the exception of some additional noise for a few payer-outcome combinations (see Online Appendix B).

17 The remaining payers include self-pay, workers' compensation, Veterans Administration, and other government payers.

18 Our sample consists of 182 acute care facilities; 120 facilities were in treated MSAs and 62 were in control MSAs (Figure 1). Of the facilities in treated MSAs, 98 matched directly to the CMS list of CJR participants, and another 8 hospitals were part of a health system (i.e., an organization with multiple hospitals) included in the CJR program list. Nine of the hospitals were part of the voluntary BPCI program; however, the hospitals opted into the BPCI model that mirrored the CJR program and followed a similar timeline. Only five treatment group hospitals could not be cleanly matched to CMS participation lists, which likely reflects changes in naming convention and/or system affiliation over time. Including these hospitals should, if anything, bias our results downward since we are strictly using geography for treatment-control classifications and therefore generating an intent-to-treat DD estimate that captures the market-wide effects of geographically assigned Medicare payment policy.

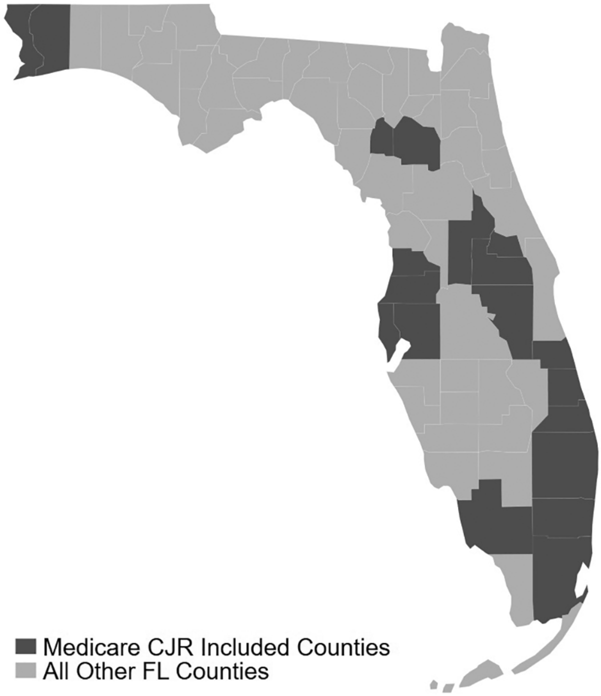

FIGURE 1. Medicare CJR included versus excluded counties in Florida. Counties that belonged to a treated MSA are in dark gray.

We operationalize our DD estimation by using two complementary specifications. We first estimate the standard two-way fixed-effects model for each LEJR patient discharge record (i) in hospital (h) at time (t):

$$Y_{iht} = \delta(\mathbf{1}(CJR_h) \times \mathbf{1}(Post_t)) + \tau_t + \lambda_h + \epsilon_{iht} \quad (1).$$

The DD model in equation 1 includes full vectors of half-year fixed effects (τ) as well as hospital fixed effects (λ) and a binary indicator for CJR inclusion status (i.e., the *CJR* variable) at the hospital level based on the original participant structure of the CJR bundled payment model (described above). The *Post* variable is equal to 1 for all time points after the announcement of the payment reform demonstration project (i.e., 2015H2), meaning we allow for anticipatory behavior in providers' responses. The δ parameter consequently reveals a summary DD estimate that averages over any immediate and longer-term direct and indirect effects of the Medicare policy.

Our focal outcomes of interest (Y_{iht}) include inpatient length of stay, operating room charges as proxy for surgical costs, and post-hospitalization discharge decisions (i.e., sent directly to home with no further care, a SNF discharge, an IRF discharge, or a discharge with home health post-acute-care services). Length of stay and operating room charges are continuous measures, while the discharge status outcomes are all binary. Of note, 99 percent of all FFS Medicare LEJR cases in the pre-CJR period (2013–15H1) have one of the four discharge outcomes that belong to our DD estimation. We cluster standard errors at the hospital level, and we estimate equation 1 separately for each specific payer group of interest to determine the presence and extent of first-order policy effects on traditional Medicare patients and spillover effects on Medicare Advantage and commercially insured patients.[19]

For our second modeling approach, we intentionally adapt our specification to an event study setup at the half-year level (across 2013–18).[20] We do so for two key reasons. First, we wish to leverage our granular and relatively long time series to examine any differential behavior between treated and control hospitals prior to, as well as following, the introduction of the CJR program. The former informs the validity of the DD research design in our study setting and hence the credibility of our resulting inferences. The latter allows us to observe any dynamic effects from the policy intervention, which could result from provider adjustment frictions and/or provider learning over time. Second, in practice, there are three distinct "post" periods that are analytically relevant in the context of the CJR program. The demonstration model structure, including hospital participants, was announced in mid-2015; however, the official start date was April 2016. Additionally, the program underwent a dramatic change at the end of 2017—leading to one-half of the national treatment areas being switched into a voluntary participation model (based on their status as having historically lower LEJR per episode spending). Also, among the remaining mandatory participation areas, opt-out provisions were introduced for qualifying "low-volume" or "rural" hospitals.[21] Approximately three-fourths of eligible hospitals nationwide withdrew from the CJR program when allowed to do so in 2018. Both Kim et al. (2018) and Einav et al. (2020b) unsurprisingly document strategic staying for hospitals electing to do so once the CJR program shifts to voluntary participation among the affected geographic areas.

On a practical level, these changes have little bearing on participation in the CJR program among Florida-specific hospitals—allowing Florida health-care markets to effectively represent the counterfactual scenario where the CJR bundled payment demonstration continues as originally intended.[22] However, 2018 is also unique in that Medicare made

19 We have also explored clustering the standard errors at the geographic treatment (i.e., MSA) level. With only 29 clusters available in Florida, we needed to use random inference approaches to account for small numbers of geographic clusters and small numbers of treated geographic units (eight in total in our context). This alternative standard error approach actually yielded tighter estimates, meaning that our standard errors from hospital-level clustering are more conservative than MSA clustering or no clustering at all within our analytic context.

20 The discharge records are quarterly; however, we implement the estimation at the half-year level to smooth out some strong seasonality in procedure volumes for these elective surgeries. Event studies at the quarterly level reveal qualitatively the same patterns of findings and inferences.

21 See the Federal Register, 82 (158) August 17, 2017, for full descriptions of these specific programmatic changes.

22 Only a single CJR-included county in Florida (affecting two hospitals in total) was granted the voluntary participation option beginning in 2018.

a long-awaited ruling that removed LEJRs from its inpatient-only (IPO) list to facilitate outpatient surgical delivery for these procedures for the first time (e.g., see Richards, Seward, and Whaley 2021). For these reasons, we want to adopt a model that allows us to observe whether hospital responses between 2015 and 2017 differ from 2018, should these two separate Medicare LEJR policies (CJR and IPO) inadvertently interact in the latter year.[23]

Our flexible event study approach consequently estimates CJR-affected hospitals' differential behavior for our outcomes of interest during the three years prior to the CJR program announcement, during the announcement, and for three years of post-implementation. The accompanying specification for each LEJR patient (i) in hospital (h) at time (t) is a slight adaptation of equation 1:

$$Y_{iht} = \sum_{\substack{j=2013H1 \\ j\neq 2015H1}}^{2018H2} \theta_j(\mathbf{1}(\tau_t = j)) + \sum_{\substack{k=2013H1 \\ k\neq 2015H1}}^{2018H2} \delta_k(\mathbf{1}(CJR_h) \times \mathbf{1}(\tau_t = k)) + \lambda_h + \epsilon_{iht} \quad (2).$$

The modified DD model in equation 2 includes full vectors of half-year fixed effects (τ) as well as hospital fixed effects (λ) and a binary indicator for CJR inclusion status (i.e., the *CJR* variable) at the hospital level based on the original participant structure of the CJR bundled payment model, just as before. The omitted time point is the six-month period immediately preceding the CJR program's announcement by CMS (i.e., when $\tau_t = 2015\text{H1}$). The set of δ_k parameters capture our focal DD estimates for our treated hospitals (i.e., those located in MSAs randomized to treatment by the CJR program in 2015, making $CJR = 1$) over our full study period. Of note, we also estimate equation 2 with a set of patient demographics (age, sex, race, and number of comorbid conditions) on the left-hand side. The patient demographic outcomes are examined to test for evidence of post-policy strategic screening of LEJR patients (i.e., "cream skimming") by hospitals to improve their likelihood of meeting or beating their respective Medicare spending targets.

A causal interpretation of the estimated post-CJR coefficients from equation 2—that is, δ_{2015H2} through δ_{2018H2}—relies on the standard difference-in-differences assumptions. We believe our approach is valid for three reasons. First, while we cannot explicitly test the absence of contemporaneous shocks, CJR inclusion areas were randomly assigned, which makes the potential for contemporaneous shocks specific to just these geographies unlikely. Second, our event study approach allows us to formally examine pre-implementation trends in our outcomes of interest for several years. Finally, as noted above, to test whether changes in treatment approaches reflect post-policy LEJR patient selection, we estimate equation 2 with patient characteristics as outcomes. We do not find that implementation of the CJR

23 For example, mandated participation in the CJR program could blunt the incentive to shift clinically appropriate cases to the outpatient setting insofar as those patients would be healthier, on average, and therefore help hold down average episode spending on the inpatient side, which raises the probability of receiving a reconciliation bonus payment at the end of 2018.

program leads to changes in observable patient characteristics among LEJR cases, which suggests that our main findings are attributable to changes in provider behavior and not differences in patient composition (fully discussed in Section V).

B. TREATMENT DECISIONS FOR NON-LEJR CASES

Our secondary DD analyses merely apply the estimation approaches developed in Section IV.A to inpatient procedure cases that are beyond the scope of the CJR program. We do this in two separate ways.

First, we examine treatment decisions for non-LEJR cases performed by policy-affected surgeons and hospitals. We can think of this set of procedures as plausibly the most proximate to the CJR program since the specialized (orthopedic) surgeons are common to both groups (i.e., LEJR and non-LEJR cases) and the cases all fall into the same family of surgeries. To identify the relevant cases for each of our three key payer groups, we focus on orthopedic surgeons who performed both LEJR Medicare cases in our primary analytic sample and any non-LEJR cases also associated with the musculoskeletal systems and connective tissues diagnoses.[24] The mapping is accomplished via the included 10-digit National Provider Identification (NPI) number for the primary surgeon attached to each inpatient discharge record. The resulting set of cases represents non-LEJR procedures that are most similar diagnostically to LEJR cases and most commonly inpatient procedures performed by LEJR surgeons.[25] We do not use all non-LEJR cases belonging to these surgeons since the greater heterogeneity in the types of procedures belonging to the mix of non-LEJR cases can mask spillover effects for cases more similar to LEJR cases in terms of case complexity, patient severity, and clinical decision making. For instance, Ryan (2018) remarks that CJR payment reform effects may not extend to other surgery types where post-acute-care use is low at baseline, even in the presence of direct incentives (e.g., a bundled payment model) for these other surgery types. We provide additional details on our sample selection in Online Appendix A (part A).

Our second approach examines provider behavior outside of the musculoskeletal surgical space. Specifically, we use all inpatient stays with a major procedure that is not related to the musculoskeletal systems and connective tissues and is not performed by a LEJR surgeon. This wider scope of inpatient surgeries can inform whether the targeted Medicare incentive program is sufficient to induce sweeping changes in hospital and physician behavior related to the costs of care, inpatient management of patients, and discharge planning that apply to a hospital's diverse set of surgical service lines.

24 All diagnoses codes belong to 25 mutually exclusive groups, known as Major Diagnostic Categories (MDCs). The two LEJR diagnosis codes in our primary analysis belong to MDC-8: disease and disorders of musculoskeletal system and connective tissues. Our non-LEJR cases focus on the remaining 97 diagnosis codes within MDC-8 that are performed by LEJR surgeons. In total, MDC-8 cases comprise 36 percent of all non-LEJR cases performed by LEJR surgeons.

25 Nine of the top 10 most performed non-LEJR procedures belong to MDC-8.

For each of these two (mutually exclusive) non-LEJR analytic samples, we then adapt equation 1 and equation 2 from Section IV.A to include a vector of DRG fixed effects (η):

$$Y_{idht} = \delta(\mathbf{1}(CJR_h) \times \mathbf{1}(Post_t)) + \tau_t + \lambda_h + \eta_d + \epsilon_{idht} \quad (3),$$

$$Y_{idht} = \sum_{\substack{j=2013H1 \\ j \neq 2015H1}}^{2018H2} \theta_j(\mathbf{1}(\tau_t = j)) + \sum_{\substack{k=2013H1 \\ k \neq 2015H1}}^{2018H2} \delta_k(\mathbf{1}(CJR_h) \times \mathbf{1}(\tau_t = k)) + \lambda_h + \eta_d + \epsilon_{idht} \quad (4).$$

All other features of the prior specifications are preserved, including the requirements for valid inferences from equation 3 and equation 4.

V. Results

A. POLICY EFFECTS ON TARGETED LEJR CASES

We first examine aggregated volumes of LEJR procedures by payer type across treated and control Florida hospitals over our study period (Figure 2). Notwithstanding seasonality-driven variation, all three patient groups demonstrate smooth upward trends in case volumes for both treatment and control hospitals. Figure 2 also offers prima facie evidence that hospitals mandated to participate in the CJR program did not respond by shrinking their set of CJR-eligible cases.[26] Importantly, Figure 2 establishes traditional Medicare as the dominant payer for LEJR cases facing hospitals. At any given time point during our study period, Medicare accounts for roughly 40 to 50 percent of all LEJR cases belonging to these three prominent payer groups. Furthermore, the shares of LEJR cases belonging to Medicare Advantage and commercially insured markets are allocated over many rival insurers and plans, rather than concentrated among one insurer as is the case for traditional Medicare.

In Table 1, we present the summary statistics for LEJR cases in the pre-policy announcement period (i.e., before 2015H2) along two dimensions: payer and CJR status. The patient demographics closely parallel each other across treatment-control statuses within a payer type in the top portion of Table 1. The one exception is the number of comorbidities for Medicare FFS patients: hospitals in CJR-included areas list nearly one additional diagnosis, on average, than control group hospitals (i.e., 8.55 relative to 7.59). Moving to the bottom portion of Table 1, we see that length of stay is also broadly similar for each of the three treatment-control comparisons and is approximately three to three and a half days, on average, for LEJR procedures. Operating room charges are likewise similar across treatment and control groups as well as payer type. There is greater, though not always consistent, variation in discharge-related treatment decisions between CJR-included and CJR-excluded hospitals for a given payer type. For example, across all three payers, CJR hospitals have

26 The later dips in inpatient LEJR cases seen across payers is consistent with the findings from Richards, Seward, and Whaley (2021), which documents providers shifting a subset of hip and knee replacements to the outpatient surgical setting following the Medicare IPO ruling change. However, Figure 2 also demonstrates that such behavior in 2018 is common across CJR treatment status in Florida—indicating that these policies do not obviously interact.

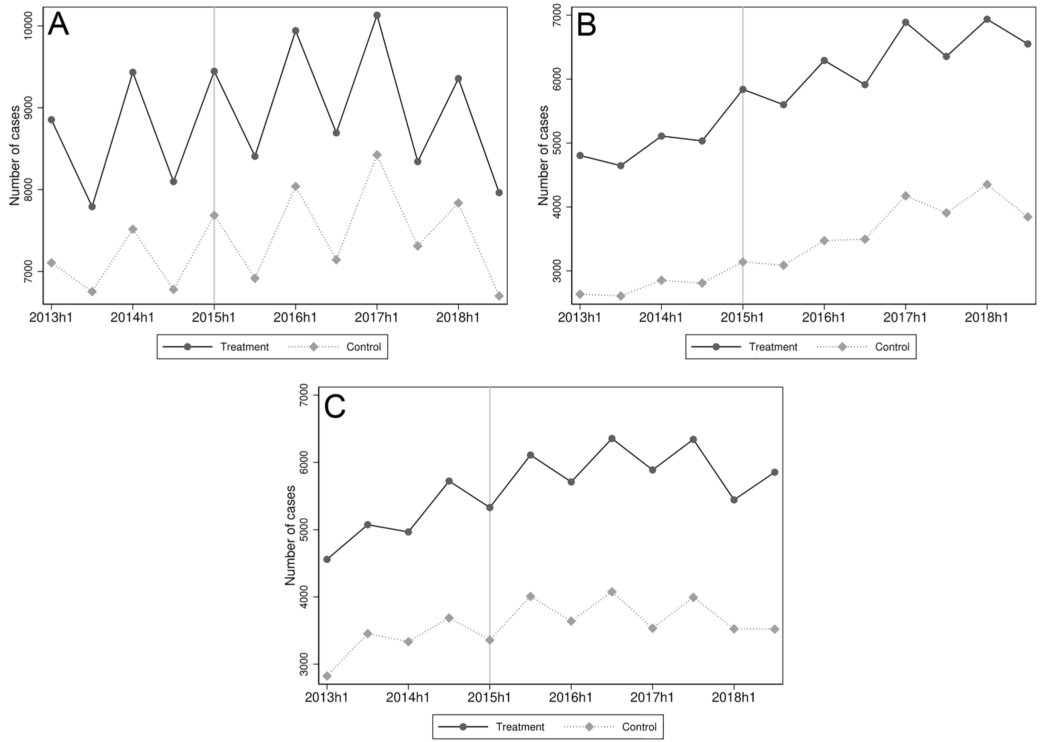

FIGURE 2. Average number of cases per hospital over time, by payer. Each plot shows the average number of LEJR procedures performed per hospital, as measured by Medicare severity diagnosis-related groups (MS-DRGs) 469 and 470. A: Medicare FFS cases. B: Medicare Advantage cases. C: Commercial cases.

TABLE 1. Summary statistics for LEJR procedures in pre-CJR policy period

	Medicare		Medicare Advantage		Commercial	
	Control (1)	Treatment (2)	Control (3)	Treatment (4)	Control (5)	Treatment (6)
Patient demographics						
Age	73.95	74.50	72.33	72.78	58.69	58.21
	(8.66)	(8.90)	(8.49)	(8.85)	(7.49)	(7.99)
1(White)	0.93	0.91	0.89	0.82	0.88	0.85
	(0.26)	(0.28)	(0.31)	(0.39)	(0.32)	(0.36)
1(Female)	0.62	0.63	0.62	0.65	0.57	0.54
	(0.49)	(0.48)	(0.48)	(0.48)	(0.50)	(0.50)
Number of other diagnoses	7.59	8.55	7.13	7.80	5.41	5.94
	(4.87)	(5.17)	(4.54)	(4.82)	(3.81)	(4.01)
LEJR outcomes						
Length of stay	3.47	3.64	3.44	3.69	2.82	2.90
	(2.02)	(2.35)	(1.97)	(2.37)	(1.24)	(1.59)
Operating room charges	22.92	22.09	23.49	24.43	22.94	22.67
(nominal dollars '000)	(12.08)	(13.90)	(12.68)	(14.30)	(11.76)	(12.46)
1(Discharged to SNF)	0.45	0.42	0.41	0.49	0.15	0.13
	(0.50)	(0.49)	(0.49)	(0.50)	(0.36)	(0.34)
1(Discharged to IRF)	0.06	0.09	0.03	0.08	0.02	0.04
	(0.24)	(0.29)	(0.17)	(0.27)	(0.15)	(0.19)
1(Discharged to home health)	0.42	0.40	0.48	0.34	0.68	0.64
	(0.49)	(0.49)	(0.50)	(0.47)	(0.47)	(0.48)
1(Discharged home)	0.05	0.07	0.06	0.08	0.14	0.18
	(0.22)	(0.26)	(0.24)	(0.27)	(0.35)	(0.39)
County characteristics[1]						
Per capita personal income	39,139	45,013	37,725	42,622	38,844	42,651
	(6,700)	(11,926)	(5,619)	(9,272)	(5,828)	(9,920)
Unemployment rate	7.31	7.22	7.43	7.33	7.28	7.10
	(0.86)	(0.93)	(0.80)	(0.92)	(0.80)	(0.88)
Observations (*N*)	35,843	43,623	14,044	25,432	16,659	25,650

Note: Data are from the Florida Agency for Healthcare Administration inpatient discharge records from 2013H1 to 2015H1. Control and treatment status is based on the CJR geographic boundaries described in Figure 1. [1]Weighted by number of procedures.

a higher rate of IRF use and lower rate of home health use prior to the CJR program's announcement. Given the stark cost differences between these two post-acute-care modalities, this greater tendency to utilize the more expensive (i.e., IRF) option may be partly responsible for the high historical Medicare spending found within these specific geographies. Relatedly, while CJR hospitals in Florida send more than 50 percent of their traditional Medicare and Medicare Advantage patients to IRFs or SNFs at baseline, the comparable national average for CJR hospitals from Meyers et al. (2019) is only 43 percent and 40 percent, respectively. This finding again highlights Florida hospitals' strong tendency to choose higher-cost treatment options when addressing an identical medical problem and further indicates the potential threat of financial penalties facing Florida hospitals assigned to the bundled payment model for all Medicare LEJR cases. The departures in pre-period discharge treatment decisions among Florida hospitals in Table 1 also underscore the value of estimating the flexible DD event study specification. Namely, we want to ensure that any differences between CJR and non-CJR hospitals are fixed throughout the 2013–15 pre-period, as opposed to creating divergent trends in provider behavior, which would invalidate the DD research design.[27]

Table 2 displays the summary DD estimates for the average effect of the CJR program during the announcement and three-year post-implementation periods. Length of stay shows no change with the introduction of Medicare's CJR bundled payment initiative for any of the three payer types. The estimates are small in magnitude (i.e., less than five one-hundredths of a day) and far from statistical significance at conventional levels. However, operating room charges decline by a statistically significant \$2,100–2,200, on average, across the three payer types. Relative to the baseline charges reported in Table 1, this translates to roughly a 10 percent decrease. Charges are an imperfect proxy for actual costs of care delivery, but the pattern is at least consistent with greater cost-containment efforts, such as standardizing staffing and/or other resource uses for LEJR cases (e.g., eliminating high-cost physician preference items). We do not observe any indications of direct or indirect policy effects on the probability of being discharged to SNF care (column 3, Table 2). The provider behavior that does change across all three payer groups is the propensity to discharge LEJR patients to IRFs following their hospitalization (column 4, Table 2). CJR hospitals decrease their use of IRFs for post-acute care by 2 percentage points (22 percent) for traditional Medicare, by 3.6 percentage points (45 percent) for Medicare Advantage, and by 0.8 percentage points (19 percent) for the nonelderly, commercially insured.[28] The DD estimates are all statistically significant for the IRF outcome in Table 2—though only at the 10 percent level for the commercially insured patients (panel C). These findings for FFS Medicare patients align with the existing literature. For example, Barnett et al. (2019) show a nearly 6 percent decline in post-acute-care discharges during the first two years of the CJR program, and more specifically, Haas et al. (2019) demonstrate a five-times larger relative decline in IRF spending

27 We also note that some simple county characteristics are broadly similar across the treatment and control groups in the bottom panel of Table 1.

28 The relative (i.e., percentage) changes discussed throughout Section V.A are based on comparing the DD estimates to the corresponding pre-CJR rates for the CJR-included hospitals found in Table 1.

TABLE 2. Effect of Medicare bundled payment policy on LEJR procedure outcomes, by payer

			Discharge disposition for LEJR procedures			
	Length of stay (days) (1)	Operating room charges (2)	SNF (3)	IRF (4)	Home health (5)	Home (routine) (6)
	A. Medicare FFS (N = 194,669)					
1(CJR) × **1**(Post)	−0.0177	−2,169[b]	0.0120	−0.0202[b]	0.0221	−0.0189
	(0.0733)	(948.6)	(0.0185)	(0.0089)	(0.0236)	(0.0230)
	B. Medicare Advantage (N = 110,341)					
1(CJR) × **1**(Post)	−0.0110	−2,222[b]	−0.0043	−0.0360[a]	0.0688[a]	−0.0292
	(0.0854)	(981.9)	(0.0198)	(0.0122)	(0.0232)	(0.0192)
	C. Commercial health insurance (N = 110,307)					
1(CJR) × **1**(Post)	−0.0434	−2,113[b]	−0.0100	−0.0076[c]	0.0764[b]	−0.0606[b]
	(0.0632)	(981.0)	(0.0117)	(0.0039)	(0.0315)	(0.0295)

Note: Analytic data are from 2013H1 to 2018H2. *CJR* equals 1 if the LEJR procedure occurred in a hospital affected by the CJR program. *Post* equals 1 in 2015H2 and beyond. Each coefficient is a separate difference-in-differences regression, which includes hospital and half-year fixed effects. Standard errors are clustered at the hospital level. [a]*p*-value at 0.01, [b]*p*-value at 0.05, [c]*p*-value at 0.10.

following CJR implementation relative to SNF spending. Our policy effect sizes are potentially larger, however, because of Florida's higher spending levels at baseline (and hence greater room for improvement) and the inclusion of a longer implementation period (i.e., through 2018).

Importantly, the remaining columns of Table 2 offer some insight as to where patients are increasingly being discharged if not to an IRF. For traditional Medicare patients, there is an approximately 2 percentage point increase in the probability of being discharged to home health post-acute-care services following their LEJR procedure. Although the estimate is imprecise, the magnitude of the home health increase is virtually identical to the observed IRF reduction. Among Medicare Advantage LEJR patients, there is a statistically significant increase of almost 7 percentage points (20 percent) in the likelihood of being discharged with home health, which more than offsets the decline in IRF utilization. The estimate for routine discharges to home (i.e., without post-acute care) in column 6 of panel B suggests that the remaining substitution toward home health services is coming from this margin, though the decline in routine discharges is not quite significant with a t-statistic of 1.52. This shift is more clearly present among the commercially insured patient population.

These patients are 7.6 percentage points (12 percent) more likely to receive home health care and 6 percentage points (34 percent) less likely to receive a routine discharge to home.

We show the corresponding event study results for length of stay, operating room charges, as well as discharges to IRFs, home health, and routine home discharges in Figure 3.[29] These graphs add important context to the results and inferences from Table 2. There is no clear pre-announcement differential trending among CJR hospitals for any of the outcomes in Figure 3, meaning that the parallel trends assumption underlying the DD strategy appears to be satisfied. While length of stay appears unchanged with a hospital's exposure to the Medicare CJR program (panel A, Figure 3), operating room charges sharply fall after the program is announced and remain depressed through 2017 for LEJR cases belonging to each of the three payers (panel B, Figure 3).[30] Likewise, following the announcement of the CJR program during the latter half of 2015, we observe a sharp (2 percentage point) drop in IRF use for traditional Medicare beneficiaries (panel C, Figure 3). The effect approximately doubles in the later years of the payment demonstration model—reflecting a 40 percent decline in IRF utilization compared with the pre-CJR level.[31] The Medicare Advantage population shows no response with CJR announcement but then an immediate reduction in IRF utilization once the bundled payment incentives are rolled out in 2016. The post-period point estimates range from -3.5 to -5.0 percentage points and demonstrate a fairly stable and sustained effect over the three years of the CJR program. Providers do not shift their IRF use for the nonelderly, commercially insured LEJR patients until the second year of the Medicare bundled payment initiative. At that time, as much as a 1.4 percentage point decrease emerges, which for the younger LEJR patient population translates to a 35 percent reduction compared with the pre-policy behavior. Importantly, across all three payer groups in Figure 3, IRF use remains lower through the end of 2018, which is consistent with the CJR-inclusion restrictions continuing to bind for the vast majority of Florida hospitals after the 2017 administrative changes by the Department of Health and Human Services and CMS.

The event study results for home health post-acute-care services for traditional Medicare patients (panel D, Figure 3) show a post-CJR upward movement, but the estimates are noisy. On the other hand, Medicare Advantage and commercially insured LEJR patients show marked and growing policy effects for this outcome in Figure 3. By the end of our study period (i.e., three years out), the likelihood of being discharged with home health services increases by approximately 10 percentage points for each of these patient groups,

29 The corresponding event study figures for SNF discharge outcomes are shown in Online Appendix Figure A1.

30 In Online Appendix Figure A1, we provide event study results for total charges. There is a similar pattern to what is observed for operating room charges, but the estimates are much noisier overall.

31 Haas et al. (2019) unadjusted trends for treatment versus control hospital LEJR average spending per case show suggestive—though somewhat noisy—indications of the spend differential narrowing in the months leading up to the official start of the CJR program. Likewise, Barnett et al. (2019) demonstrate some suggestive anticipatory behavior in terms of converging average post-acute-care institutional spending across treatment and control hospitals as well. And Meyers et al. (2019) observe evidence of announcement effects for traditional Medicare and Medicare Advantage patients in their study.

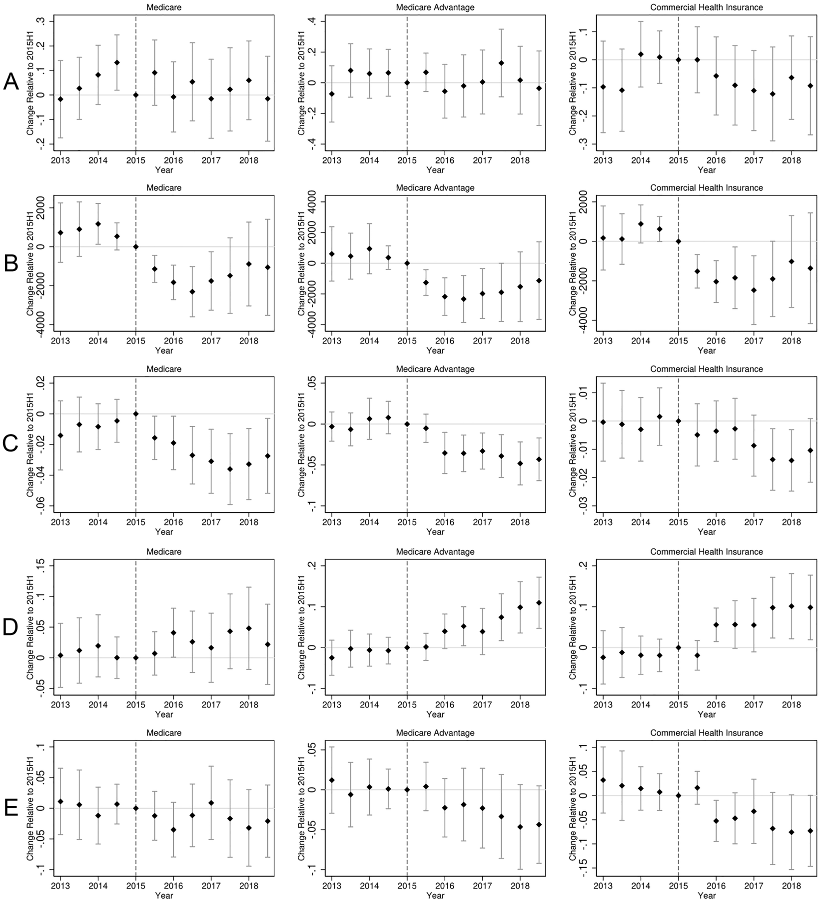

FIGURE 3. Effect of bundled payment on LEJR outcomes. Each plot shows the event study coefficients and 95 percent confidence interval for outcomes pertinent to LEJR procedures. All changes are measured relative to the first half of 2015 (2015H1). Each regression includes hospital fixed effects. Standard errors are clustered at the hospital level. Dashed vertical line represents the six months immediately preceding the announcement of the CJR program and hospital assignment to the CJR program. A: Length of stay. B: Operating room charges. C: Probability discharged to IRF. D: Probability discharged to home health. E: Probability discharged home (routine).

representing a 29 percent and 16 percent relative change for Medicare Advantage and commercially insured, respectively. Finally, there is no obvious change in the likelihood of being discharged with no further recovery services for traditional Medicare patients (panel E, Figure 3), but there is a suggestive decline over time for Medicare Advantage patients and a pronounced drop for the commercially insured. The latter group is 5 to 8 percentage points less likely to have a routine discharge home among CJR-affected hospitals, which translates to an approximately 28 to 42 percent decrease over their baseline (pre-policy) rate in Table 1.

In Online Appendix Figure A2, we do not find any evidence for strategic patient selection in terms of observable patient characteristics (age, sex, and race) for any of the three payer groups. The number of reported other diagnoses does not obviously change for Medicare or commercially insured LEJR patients, though there is a 0.4 to 0.8 percentage point decrease in the number of reported additional diagnoses among the Medicare Advantage group following the policy change. This translates to only a 5 percent to 10 percent relative decline from a baseline mean of nearly eight additional diagnoses reported per patient (Table 1). It therefore seems unlikely that this singular diagnosis reporting change limited to Medicare Advantage patients—which likely represents minimal changes in patient risk—can explain the pronounced spillover effects documented in Table 2 and Figure 3.

Taken together, our results on spillover effects onto other insurers indicate that both Medicare Advantage and commercial patients were affected by the traditional Medicare bundled payment incentive structure. Across all payers, providers favored discharges home with home health care over more intensive and expensive care settings (i.e., IRFs). However, providers also reduced their discharges home with just routine care, particularly among the commercially insured—indicating that nonelderly commercially insured LEJR patients at CJR-affected hospitals actually receive greater post-acute care following Medicare's bundled payment initiative. Because of data constraints, however, we are unable to assess whether the higher intensity of services for this patient group translates to net savings for their insurers, such as through faster recovery times or fewer follow-on procedures.[32]

B. POLICY EFFECTS ON UNTARGETED NON-LEJR CASES

Next, we consider spillovers for provider behavior toward non-LEJR cases. Recall from Section IV.B that we are ultimately considering two types of non-LEJR cases: those likely to be highly similar to LEJRs (i.e., performed by the same surgeons and belonging to the same MDC-8 classification) and those without clear overlap with LEJRs (i.e., performed by non-LEJR surgeons and not belonging to the MDC-8 classification).

Table 3 shows the summary statistics for our comparable non-LEJR cases among surgeons who also perform LEJR procedures. Across payer-specific control and treatment groups, patient demographics are again similar, with the exception of the commercially insured having a slightly younger population among CJR hospitals (47 versus 52 years of age, on average). The average length of stay for non-LEJR procedures is slightly longer among CJR hospitals. However, operating room charges are similar, and discharge propensities to SNFs, IRFs, home health care, and home with no post-acute care all occur with almost equal frequency across treated and control populations. As before, we rely on the event study estimates to demonstrate whether the trends in these variables are similar, in spite of any small (and fixed) level differences in the pre-period. We also note the discrepancies in

32 Specifically, there are no patient identifiers in the Florida AHCA discharge data that would allow longitudinal tracking of health-care utilization at the patient level.

TABLE 3. Summary statistics for non-LEJR procedures in pre-CJR policy period

	Medicare		Medicare Advantage		Commercial	
	Control	Treatment	Control	Treatment	Control	Treatment
Patient demographics						
Age	75.57	75.86	73.24	74.71	52.14	47.41
	(11.24)	(11.86)	(10.74)	(11.31)	(15.30)	(18.42)
1(White)	0.94	0.90	0.90	0.84	0.88	0.82
	(0.25)	(0.30)	(0.29)	(0.37)	(0.32)	(0.38)
1(Female)	0.65	0.65	0.64	0.66	0.53	0.50
	(0.48)	(0.48)	(0.48)	(0.47)	(0.50)	(0.50)
Number of other diagnoses	10.02	10.61	9.24	9.67	5.89	5.59
	(5.82)	(6.02)	(5.55)	(5.82)	(4.99)	(4.86)
Non-LEJR outcomes						
Length of stay	4.46	4.94	4.51	4.98	3.49	3.75
	(3.80)	(4.31)	(3.98)	(4.13)	(3.71)	(4.06)
Operating room charges	22.36	21.13	22.34	21.95	24.43	23.96
(nominal dollars '000)	(17.35)	(17.89)	(17.32)	(18.75)	(18.33)	(20.61)
1(Discharged to SNF)	0.47	0.45	0.46	0.48	0.11	0.08
	(0.50)	(0.50)	(0.50)	(0.50)	(0.32)	(0.27)
1(Discharged to IRF)	0.11	0.12	0.04	0.06	0.05	0.05
	(0.31)	(0.33)	(0.20)	(0.25)	(0.22)	(0.22)
1(Discharged to home health)	0.23	0.23	0.28	0.24	0.29	0.28
	(0.42)	(0.42)	(0.45)	(0.43)	(0.45)	(0.45)
1(Discharged home)	0.16	0.15	0.20	0.19	0.53	0.57
	(0.37)	(0.36)	(0.40)	(0.39)	(0.50)	(0.50)
Observations (*N*)	30,037	43,020	10,634	23,386	12,418	25,821

Note: Data are from the Florida AHCA inpatient discharge records for non-LEJR musculoskeletal and connective tissue procedures from 2013H1 to 2015H1. Sample is restricted to orthopedic surgeons who also perform LEJR procedures. Control and treatment status is based on the CJR geographic boundaries described in Figure 1.

the discharge destination rates and patient characteristics between LEJR and orthopedic non-LEJR cases. Compared with Table 1, the summary statistics in Table 3 show that non-LEJR patients, particularly those insured with traditional Medicare, tend to have more additional diagnoses (i.e., comorbidities), are more likely to be discharged home without additional care, and are less likely to receive home health care, specifically.

The DD results from the two-way fixed-effects estimations for these non-LEJR cases are displayed in Table 4. While length of stay is again unaffected, there is a $2,074 drop in operating room charges that closely approximates what was observed in Table 2 for LEJR cases for each of the three payer types. Among traditional Medicare patients (panel A), there is a 2.4 percentage point (16 percent) reduction in discharges home without any post-acute support and a corresponding 2.4 percentage point (5 percent) increase in discharges to SNFs. Among Medicare Advantage patients (panel B), the decrease in home discharges is smaller and not statistically significant; however, there is evidence of a substantive spillover reduction in IRF post-acute care, which falls by 1.9 percentage points (32 percent) following the introduction of the LEJR bundled payment reforms. Substituting away from IRFs, CJR-affected providers increase their reliance on home health care by 2 percentage points (9 percent) and SNF care by 1.6 percentage points (though this estimate is less precise) for non-LEJR Medicare Advantage cases. Finally, among the commercially insured population (panel C), there is no clear change in SNF or IRF care, but there is a decrease in routine

TABLE 4. Spillover effect of LEJR bundled payment on non-LEJR procedures, by payer

			Discharge disposition for LEJR procedures			
	Length of stay (days) (1)	Operating room charges (2)	SNF (3)	IRF (4)	Home health (5)	Home (routine) (6)
	A. Medicare FFS (N = 175,809)					
1(CJR) × **1**(Post)	0.0861	−2,074[b]	0.0240[b]	−0.0087	0.0088	−0.0241[b]
	(0.0808)	(924.7)	(0.0093)	(0.0093)	(0.0090)	(0.0097)
	B. Medicare Advantage (N = 93,362)					
1(CJR) × **1**(Post)	0.1170	−1,883[c]	0.0163	−0.0192[a]	0.0213[c]	−0.0194
	(0.0908)	(989.5)	(0.0103)	(0.0068)	(0.0114)	(0.0125)
	C. Commercial health insurance (N = 92,906)					
1(CJR) × **1**(Post)	0.0825	−2,243[b]	−0.0037	0.0019	0.0296[b]	−0.0278[b]
	(0.0884)	(1,137)	(0.0055)	(0.0042)	(0.0145)	(0.0134)

Note: Analytic data are from 2013H1 to 2018H2, among physicians who perform both LEJR and related non-LEJR (i.e., MDC-8) procedures. *CJR* equals 1 if the non-LEJR procedure occurred in a hospital affected by the CJR program. *Post* equals 1 in 2015H2 and beyond. Each coefficient is a separate difference-in-differences regression, which includes hospital, MS-DRG, and half-year fixed effects. Standard errors are clustered at the hospital level. [a]p-value at 0.01, [b]p-value at 0.05, [c]p-value at 0.10.

discharges to home by 2.8 percentage points (5 percent) and a commensurate increase in discharges with home health post-acute-care services by 3.0 percentage points (10 percent).

In Online Appendix Figure A3, we present the event study regression results for these non-LEJR cases. For each outcome and patient population, we do not observe strong indications of pre-CJR differential behavior between CJR and non-CJR hospitals. Similar to the DD regression results in Table 4, we observe post-announcement reductions in operating charges across all three payers (panel B) as well as in the probability of discharge to an IRF for traditional Medicare and Medicare Advantage patients (panel C). However, the decline in IRF use among traditional Medicare non-LEJR cases is more ephemeral, with a rebound toward the end of our study period. Conversely, the Medicare Advantage population experiences an approximately 2 percentage point decrease that is fairly stable over the three post-policy years. Consistent with the summary DD result in Table 4, we do not find a compelling change in IRF discharges for commercially insured patients in panel C. For the two other primary outcomes that align with the LEJR results from Table 2 and Figure 3, we find suggestive increases in the use of home health services and declines in routine discharges to home without post-acute-care services. However, the event studies for the commercially insured populations in panels D and E indicate that CJR hospitals differentially changed their behavior starting in the earlier part of 2015. In panel F of Online Appendix Figure A3, the event study patterns of traditional Medicare and Medicare Advantage non-LEJR procedures offer additional suggestive evidence (consistent with Table 4) of increased SNF use among CJR-impacted hospitals after the CJR program is in place.

Table 5 reports the pre-period summary statistics for our second collection of non-LEJR cases (i.e., those outside the MDC-8 classification and not performed by LEJR surgeons).[33] Just as before, the patient characteristics and provider treatment patterns are not markedly different across the treatment-control divide for any of the three payer types. The corresponding DD estimates for this analytic sample of surgical cases is found in Table 6. For all six outcomes and all three payer types, the coefficients are uniformly small and not statistically different from zero at conventional levels, with one exception: the probability of SNF discharge for Medicare Advantage cases in panel B—which is only marginally significant. In Online Appendix Table A3, we perform an analogous empirical exercise as Table 5; however, we restrict to non-MDC 8 classifications that also had relatively high rates of post-acute-care use at baseline and therefore might be plausibly more affected by broad changes in hospital behavior following the introduction of the CJR program. Even with this narrowing of procedure types, there is still no evidence of a spillover effect extending to cases performed by non-LEJR surgeons (Online Appendix Table A3).

Taken together, these complementary results offer some evidence that the impact of the Medicare CJR program can extend to treatment decisions for non-LEJR cases, but such externalities appear confined to cases more similar to LEJRs and performed by common surgeons (and perhaps common surgical teams). Specifically, CJR-affected providers appear to

33 Online Appendix Table A2 provides a detailed breakdown of all non-LEJR procedures performed by non-LEJR physicians.

TABLE 5. Summary statistics in pre-CJR policy period for all inpatient stays with a major procedure when excluding MDC-8 cases

	Medicare		Medicare Advantage		Commercial	
	Control	Treatment	Control	Treatment	Control	Treatment
Patient demographics						
Age	72.73	72.37	71.77	72.54	38.48	36.12
	(12.73)	(14.00)	(10.91)	(11.49)	(21.70)	(20.99)
1(White)	0.85	0.80	0.84	0.75	0.81	0.73
	(0.36)	(0.40)	(0.37)	(0.43)	(0.39)	(0.45)
1(Female)	0.50	0.50	0.48	0.50	0.62	0.63
	(0.50)	(0.50)	(0.50)	(0.50)	(0.49)	(0.48)
Number of other diagnoses	14.17	14.27	13.06	12.83	6.87	6.27
	(6.88)	(7.11)	(6.60)	(6.97)	(6.15)	(5.78)
Outcomes						
Length of stay	6.55	7.32	6.27	6.58	4.41	4.77
	(7.19)	(9.18)	(6.52)	(7.46)	(7.48)	(8.66)
Operating room charges	9.37	8.76	10.71	10.84	8.42	8.30
(nominal dollars '000)	(20.36)	(18.38)	(21.87)	(21.38)	(17.91)	(16.65)
1(Discharged to SNF)	0.19	0.18	0.14	0.14	0.02	0.01
	(0.39)	(0.39)	(0.35)	(0.34)	(0.13)	(0.11)
1(Discharged to IRF)	0.03	0.03	0.01	0.01	0.01	0.01
	(0.17)	(0.17)	(0.12)	(0.12)	(0.09)	(0.09)
1(Discharged to home health)	0.19	0.20	0.20	0.21	0.07	0.07
	(0.39)	(0.40)	(0.40)	(0.41)	(0.25)	(0.25)
1(Discharged home)	0.45	0.43	0.52	0.52	0.86	0.87
	(0.50)	(0.50)	(0.50)	(0.50)	(0.35)	(0.33)
Observations (N)	322,164	496,668	112,402	279,031	226,566	560,435

Note: Data are from the Florida AHCA inpatient discharge records for all inpatient stays including major procedures from 2013H1 to 2015H1 that do not belong to the MDC-8 grouping (musculoskeletal system and connective tissue diseases and disorders). Control and treatment status is based on the CJR geographic boundaries described in Figure 1.

reduce their reliance on post-acute-care IRFs, especially among the Medicare Advantage population, and to view home health services as a more attractive post-acute-care option for their other non-LEJR cases. That said, the magnitude of these changes for non-LEJR cases tends to be roughly half of the change in discharge location decisions observed for LEJR cases, which indicates that even if such spillovers are present, the impact is likely to be

TABLE 6. Spillover effect of LEJR bundled payment on non-MDC-8 procedures, by payer

			Discharge disposition for LEJR procedures			
	Length of stay (days) (1)	Operating room charges (2)	SNF (3)	IRF (4)	Home health (5)	Home (routine) (6)
	A. Medicare FFS (N = 2,036,059)					
1(CJR) × **1**(Post)	0.0196	−90.01	0.0014	0.0012	−0.0031	−0.0005
	(0.0787)	(311.7)	(0.0033)	(0.0017)	(0.0054)	(0.0060)
	B. Medicare Advantage (N = 1,139,360)					
1(CJR) × **1**(Post)	0.1120	69.49	−0.0056[c]	−0.0022	0.0012	0.0051
	(0.0838)	(367.5)	(0.0031)	(0.0013)	(0.0056)	(0.0062)
	C. Commercial health insurance (N = 2,080,254)					
1(CJR) × **1**(Post)	0.0342	−207.90	−0.0005	−0.0003	−6.11e-05	−0.0008
	(0.0554)	(328.6)	(0.0007)	(0.0005)	(0.0028)	(0.0031)

Note: Analytic data are from 2013H1 to 2018H2 and restricted to inpatient stays for procedures that do not belong to the MDC-8 category. *CJR* equals 1 if the non-LEJR procedure occurred in a hospital affected by the CJR program. *Post* equals 1 in 2015H2 and beyond. Each coefficient is a separate difference-in-differences regression, which includes hospital, MS-DRG, and half-year fixed effects. Standard errors are clustered at the hospital level. [c]*p*-value at 0.10.

considerably weaker relative to procedures explicitly targeted by the payment reform. Unlike LEJR cases, the non-LEJR cases demonstrate a modest increase in SNF use among traditional Medicare population; however, the increase may result from underlying differences in care needs and post-acute-care options among non-LEJR patients. Moreover, while the finding for this outcome contrasts with the LEJR findings from Table 2, it can still be consistent with the Medicare CJR bundled payment model creating the necessary impetus for affected providers to reevaluate their post-acute-care options and previous clinical decision-making tendencies for a wider set of procedures.

VI. Heterogeneous Effects for LEJR Cases

Our final empirical exercise examines any heterogeneity in the effects for the two outcomes showing the most pronounced and consistent post-CJR changes in Table 2 and Figure 3 (i.e., operating room charges and IRF discharge probability). Specifically, we use a triple differences (DDD) model to test whether certain LEJR surgeons drive the overall effects previously estimated in Section V.[34] To do so, we adapt equation 1 to a DDD specification:

$$\begin{aligned} Y_{iht} = \tau_t + \lambda_h + \varsigma PhysicianType_i + \gamma_1(\mathbf{1}(CJR_h) \times \mathbf{1}(PhysicianType_i)) \\ + \gamma_2(\mathbf{1}(Post_t) \times \mathbf{1}(PhysicianType_i)) + \delta_1(\mathbf{1}(CJR_h) \times \mathbf{1}(Post_t)) \\ + \delta_2(\mathbf{1}(CJR_h) \times \mathbf{1}(Post_t) \times \mathbf{1}(PhysicianType_i)) + \epsilon_{iht} \end{aligned} \quad (5).$$

The key difference for equation 5 when compared with equation 1 is the presence of the *PhysicianType* variable and accompanying triple interaction term and sub-interaction terms. We consider two separate binary indicators for *PhysicianType* and run the model separately for each. The first indicator option is equal to 1 if the LEJR surgeon is an international medical graduate (IMG) for the specific LEJR patient discharge record (i) in hospital (h) at time (t) and 0 otherwise. The second physician characteristics option is equal to 1 if the LEJR surgeon completed medical school by 1995 or earlier (and 0 otherwise) in order to serve as a proxy for a surgeon with greater experience—and perhaps greater inertia with respect to changing clinical decision making and treatment style. Both physician-level characteristics are from the publicly available Physician Compare database maintained by CMS.

The corresponding DDD results are displayed in Table 7. There is no clear and consistent pattern when examining heterogeneity in terms of LEJR surgeon experience (panels C and D). More experienced surgeons do not drive the results seen in Table 2 and Figure 3; however, there is no clear indication that they are not responding to the CJR incentives. This contrasts to the heterogeneity results for IMGs (panels A and B). Panel A of Table 7 indicates that the IMG LEJR surgeons are the drivers of the post-CJR implementation decreases in average operating room charges. If these declines are reflective of real behavioral changes in terms of more judicious use of surgical room staffing and supplies and/or greater restraints on physician preference items (e.g., higher-cost joint implants), the estimates are at least consistent with IMGs being more amenable to such changes in the short run. The DDD estimates for the likelihood of being discharged to IRF care are less precise in panel B, but they do offer suggestive evidence that the IMG LEJR surgeons demonstrated a stronger response to the CJR program along this margin—at least among the traditional Medicare and Medicare Advantage patients.

VII. Discussion and Conclusions

Industry experts and policy makers regularly lament the state of US health care and advocate for more cost-conscious approaches to financing medical service delivery and population health outcomes. One path toward such ends is to use the frequently largest payer (i.e.,

34 Within Online Appendix Table A4, we perform identical DDD empirical estimations using hospital characteristics—namely, the hospital having high average LEJR charges at baseline, the hospital being not-for-profit (NFP), or the hospital having any vertically integrated LEJR surgeons. There are no consistent patterns across the two key outcomes of interest for any of the three hospital characteristics. These results indicate that the IMG heterogeneity revealed in Table 7 is not merely a reflection of IMG LEJR surgeons being concentrated in a particular type of Florida hospital.

TABLE 7. Heterogeneous effects of Medicare bundled payment policy on LEJR procedure outcomes according to surgeon characteristics, by payer

	Medicare FFS (1)	Medicare Advantage (2)	Commercial (3)	Medicare FFS (1)	Medicare Advantage (2)	Commercial (3)
	A. Operating room charges by surgeon foreign education			B. Discharge to IRF by surgeon foreign education		
1(CJR) × **1**(Post)	−1,073	−1,375	−808	−0.017[b]	−0.022[b]	−0.009[b]
	(1,043)	(1,072)	(1,059)	(0.009)	(0.009)	(0.004)
1(CJR) × **1**(Post) × **1**(IMG)	−3,653[a]	−2,632[c]	−4,714[a]	−0.014	−0.035[c]	−0.001
	(1,187)	(1,465)	(1,409)	(0.011)	(0.021)	(0.009)
N	192,176	108,102	107,546	192,176	108,102	107,546
	C. Operating room charges by surgeon age			D. Discharge to IRF by surgeon age		
1(CJR) × **1**(Post)	−2,873[a]	−3,276[a]	−3,560[a]	−0.028[b]	−0.042[b]	−0.004
	(999)	(1,055)	(1,124)	(0.011)	(0.017)	(0.006)
1(CJR) × **1**(Post) × **1**(Older)	1,022	1,731[c]	2,468[b]	0.012	0.010	−0.007
	(1,086)	(968)	(1,082)	(0.010)	(0.012)	(0.007)
N	191,593	107,912	107,272	191,593	107,912	107,272

Note: Analytic data are from 2013H1 to 2018H2. *CJR* equals 1 if the LEJR procedure occurred in a hospital affected by the CJR program. *Post* equals 1 in 2015H2 and beyond. "IMG" stands for completing medical school outside of the US. "Older" are physicians that graduated from medical school in 1995 or earlier. These physician characteristics are from the Physician Compare database maintained by CMS. Each regression is a triple differences model and includes hospital and half-year fixed effects. Interaction between the physician characteristics and post dummy are included but not shown in the table. Standard errors are clustered at the hospital level. [a]*p*-value at 0.01, [b]*p*-value at 0.05, [c]*p*-value at 0.10.

Medicare) facing a given provider as a vehicle for driving non-FFS incentive structures. Despite a ballooning number of Medicare-based payment model demonstrations, externalities tied to a given model are rarely considered or quantified. We leverage an innovative Medicare payment reform and comprehensive data from the state of Florida to estimate the presence and reach of Medicare spillovers.

We find that the program leads to meaningful direct and indirect changes in provider behavior. Among the traditional Medicare population, discharges to high-cost inpatient rehab facilities decrease by 22 percent, but this understates the impact of the CJR program along this margin. IRF discharges also decline by as much as 44 percent for Medicare Advantage patients and as much as 38 percent for nonelderly commercially insured patients. For all three LEJR patient populations, the reduction in IRF discharges is offset by greater use of home health post-acute-care services. IRF care is markedly more expensive than home health services, and importantly, there is no clear clinical evidence that joint replacement patients receiving IRF care, rather than home health care, fare better (Buhagiar et al. 2019). For these reasons, the observed post-acute-care substitution behavior indicates a beneficial re-optimization of providers' LEJR treatment choices across payers following the introduction of stronger incentives from one prominent payer (i.e., Medicare) to economize on post-acute costs per episode. Additionally, we find suggestive evidence that the program creates spillovers for more similar surgical procedures that are not covered by the CJR program, which has the potential to benefit patients and payers beyond the scope of a targeted payment reform intervention.[35]

We further note that the provider behavior changes we empirically document are more consistent and substantively larger than what has been found among existing CJR studies—including those examining the LEJR Medicare Advantage market. As previously remarked, our approach has the analytic advantage of capturing more payers as well as the full universe of accompanying cases over a longer time period, which can help in identifying effects (especially those that evolve over time). However, another, and not mutually exclusive, potential explanation is that we focus on Florida, which contains some of the highest-spending areas and providers in the country. For LEJR cases, specifically, comparing our data with the national data used in other work reveals that CJR-targeted Florida hospitals had a much stronger propensity to use high-cost post-acute-care options at baseline when compared with their CJR peers from elsewhere around the country. Intuitively, we would expect more pronounced changes (i.e., greater policy bite) among high-cost providers once a new incentive structure has been introduced. This difference also likely means that we are capturing an upper bound of direct and indirect bundled payment policy effects, but this is still informative for health policy debates regarding interventions that seek to curb spending in very high-spending areas.

Our observed provider behavior changes—concentrated among post-acute-care decision making—are in concordance with the existing CJR literature and also providers' private financial interests within the CJR regime. Carroll et al. (2018) illustrate the importance

35 Recall that we found no evidence in Section V.B that indicated hospital and physician behavior changed for a wider set of non-LEJR cases (i.e., those not performed by LEJR surgeons exposed to the CJR program).

of conflicting incentives facing a given provider under a bundled payment structure that relies on ex post reconciliation rewards or penalties—as the CJR program does. Namely, all care under the direct control of the provider is still paid FFS, so any billable effort reductions that help in achieving the benchmark total spend simultaneously reduce the provider's own payments. Conversely, all care delivered by other providers (but still part of the episode) can be restrained to make it more likely that overall spending reductions are accomplished (generating a financial reward) without lowering total FFS payments to the focal provider at risk. Our results are consistent with the Carroll et al. (2018) theoretical model and estimates for perinatal care in Arkansas. Specifically, the authors show an approximately 3 percent decline in total episode spending, which seems to reflect referrals to less expensive birth facilities by the providers directly incentivized by the bundled payment initiative. Our findings also support the argument made by Frandsen, Powell, and Rebitzer (2019) in which the authors describe the importance of a given payer's relative share of the market for influencing provider contracting and behavior.[36] It also seems unlikely that pre-CJR LEJR providers experienced prohibitive fixed-cost investments (e.g., specialized information technology needs) in order to improve discharge planning and hence participate in risk-bearing contracts for LEJR cases—as would be the case in the context of a payer coordination failure (see Frandsen, Powell, and Rebitzer 2019).

Importantly, our estimates offer compelling empirical support for the "norms hypothesis" of provider behavior and complement the results from Barnett, Olenski, and Sacarny (2020). Our work is arguably most similar in spirit to Barnett, Olenski, and Sacarny (2020) as well. The authors examine antipsychotic prescribing to commercially insured patients among physicians receiving a behavioral nudge about overprescribing for their Medicare patients. Crucially, the authors also assess changes in prescribing behavior for the intervention targeted drug as well as closely related pharmaceutical therapies. Despite finding nearly symmetrical reductions in prescribing for the targeted pharmaceutical therapy for Medicare beneficiaries and older adult commercially insured patients, Barnett, Olenski, and Sacarny (2020) do not find changes in prescribing patterns beyond the focal drug tied to the Medicare information intervention. The authors also show that the cutbacks in prescribing for the focal drug appear indiscriminate (i.e., not reflective of underlying patient appropriateness). In comparison, we demonstrate more nuanced treatment decision making, which even leads to greater post-acute-care service provision for some patients. CJR-affected providers seem to devote more attention to post-hospitalization treatment choices and outcomes across payers and procedures but not rigidly impose the same changes to all clinical contexts. Our evidence is therefore more consistent with specific decision nodes in the treatment pathway (e.g., post-surgical discharge planning) gaining new and greater salience as well as potential provider learning about the cost-quality trade-offs from different treatment approaches after the introduction of new Medicare financial incentives. In these ways, our findings juxtaposed to Barnett, Olenski, and Sacarny (2020) offer a useful comparison for the spillover reach for two different styles of Medicare intervention (i.e., provider payment

36 Recall that within our analytic data, traditional Medicare LEJR cases typically outnumber each of the other two prominent payers' (i.e., Medicare Advantage and non-Medicare commercial) cases by 50 to 100 percent in any given half-year.

reforms versus provider information campaigns) that apply to two different types of physicians (i.e., surgeons versus primary care specialists). The evidence from each study indicates that Medicare policy can be the source of substantive externalities for other patients and payers.

One pre-implementation CJR study (Maniya et al. 2017) arrived at the conclusion that hospitals should by and large ignore the program and accept any resulting penalties. Clearly, many providers, especially in Florida markets, did not take such advice and instead used the CJR program as an impetus for rethinking treatment approaches for other (non-Medicare) patients. To date, no studies have documented perverse consequences from the CJR or BPCI programs, such as greater complication rates, altered case volumes, or patient cream-skimming (Dummit et al. 2016; Navathe, Liao, Dykstra, et al. 2018; Barnett et al. 2019; Meyers et al. 2019; Einav et al. 2020a), and previously, Sood et al. (2011) expressed a perhaps common view that other payers might eventually learn from the Medicare CJR experience. While still potentially true, our findings suggest that many of them received an immediate private benefit. Relatedly, existing studies of the CJR program (Finkelstein et al. 2018; Barnett et al. 2019; Haas et al. 2019; Einav et al. 2020b) remark that the net savings has been minimal; however, the more holistic view we take adds important qualifications. Our empirical evidence is consistent with uncontracted and uncompensated savings for Medicare Advantage insurers over a three-year period—meaning more efficient care delivery for the broader health-care system during this time.[37] That said, bundled payments are not easy to construct or implement, especially for a wide variety of medical problems and clinical contexts (Burns and Pauly 2018). The iterative and circumscribed process to provider payment reforms inside and outside of Medicare is therefore likely to continue; however, implementers and evaluators should pay closer attention to the range of externalities that may surface from a given intervention. Social welfare estimates and policy makers' ex post decisions could be distorted if the analytic view is too narrow, especially when the involved payer looms large within the relevant health-care market.

ACKNOWLEDGEMENTS

The authors thank the Florida Agency for Healthcare Administration (AHCA) for providing valuable data resources. AHCA was not responsible for any data analyses or interpretations. All opinions and remaining errors belong solely to the authors.

FUNDING INFORMATION

Funding was provided by NIA K01AG061274 (Whaley).

REFERENCES

Abrams, M., R. Nuzum, M. Zezza, J. Ryan, J. Kiszla, and S. Guterman. 2015. "The Affordable Care Act's Payment and Delivery System Reforms: A Progress Report at Five Years." *Commonwealth Fund Issue Brief* 12:1–16.

37 The spending implications for the commercially insured group are less clear because of data constraints.

Baicker, K., M. E. Chernew, and J. A. Robbins. 2013. "The Spillover Effects of Medicare Managed Care: Medicare Advantage and Hospital Utilization." *Journal of Health Economics* 32 (6): 1289–300.

Baicker, K., and J. A. Robbins. 2015. "Medicare Payments and System-Level Health Care Use: The Spillover Effects of Medicare Managed Care." *American Journal of Health Economics* 1 (4): 399–431.

Barnett, M. L., A. Olenski, and A. Sacarny. 2020. "Common Practice: Spillovers from Medicare on Private Health Care." NBER Working Paper No. 27270. https://doi.org/10.3386/w27270.

Barnett, M. L., A. Wilcock, J. M. McWilliams, A. Epstein, K. E. J. Maddox, E. J. Orav, D. C. Grabowski, and A. Mehrota. 2019. "Two-Year Evaluation of Mandatory Bundled Payments for Joint Replacement." *New England Journal of Medicine* 380 (3): 252–62.

Bond, A. M., and C. White. 2013. "Massachusetts Coverage Expansion Associated with Reduction in Primary Care Utilization among Medicare Beneficiaries." *Health Services Research* 48 (6): 1826–39.

Buhagiar, M. A., J. M. Naylor, I. A. Harris, W. Xuan, S. Adie, and A. Lewin. 2019. "Assessment of Outcomes of Inpatient or Clinic-Based vs. Home-Based Rehabilitation after Total Knee Arthroplasty: A Systematic Review and Meta-Analysis." *JAMA Network Open* 2 (4): e192810. https://doi.org/10.1001/jamanetworkopen.2019.2810.

Burns, L. R., and M. V. Pauly. 2018. "Transformation of the Health Care Industry: Curb Your Enthusiasm?" *Milbank Quarterly* 96 (1): 57–109.

Callison, K. 2016. "Medicare Managed Care Spillovers and Treatment Intensity." *Health Economics* 25 (7): 873–87.

Carey, C. M., S. Miller, and L. R. Wherry. 2020. "The Impact of Insurance Expansions on the Already Insured: The Affordable Care Act and Medicare." *American Economic Journal: Applied Economics* 12 (4): 288–318.

Carroll, C., M. Chernew, A. M. Fendrick, J. Thompson, and S. Rose. 2018. "Effects of Episode-Based Payment on Health Care Spending and Utilization: Evidence from Perinatal Care in Arkansas." *Journal of Health Economics* 61:47–62.

Clemens, J., and J. D. Gottlieb. 2017. "In the Shadow of a Giant: Medicare's Influence on Private Physician Payments." *Journal of Political Economy* 125 (1): 1–39.

Clemens, J., J. D. Gottlieb, and T. L. Molnár. 2017. "Do Health Insurers Innovate? Evidence from the Anatomy of Physician Payments." *Journal of Health Economics* 55:153–67.

Conrad, D. A. 2016. "The Theory of Value-Based Payment Incentives and Their Application to Health Care." *Health Services Research* 50 (S2): 2057–89.

Cooper, Z., S. V. Craig, M. Gaynor, and J. Van Reenen. 2019. "The Price Ain't Right? Hospital Prices and Health Spending on the Privately Insured." *Quarterly Journal of Economics* 134 (1): 51–107.

Dummit, L. A., D. Kahvecioglu, G. Marrufo, R. Rajkumar, J. Marshall, E. Tan, M. J. Press, et al. 2016. "Association between Hospital Participation in a Medicare Bundled Payment Initiative and Payments and Quality Outcomes for Lower Extremity Joint Replacement Episodes." *Journal of the American Medical Association* 316 (12): 1267–78.

Einav, L., A. Finkelstein, Y. Ji, and N. Mahoney. 2020a. "Randomized Trial Shows Healthcare Payment Reform Has Equal-Sized Spillover Effects on Patients Not Targeted by Reform." *Proceedings of the National Academy of Sciences* 117 (32) 18939–47. https://doi.org/10.1073/pnas.2004759117.

———. 2020b. "Voluntary Regulation: Evidence from Medicare Payment Reform." NBER Working Paper No. 27223. https://doi.org/10.3386/w27223.

Ellimoottil, C., A. M. Ryan, H. Hou, J. Dupree, B. Hallstrom, and D. C. Miller. 2016. "Medicare's New Bundled Payment for Joint Replacement May Penalize Hospitals That Treat Medically Complex Patients." *Health Affairs* 35 (9): 1651–57.

———. 2017. "Implications of the Definition of an Episode of Care Used in the Comprehensive Care for Joint Replacement." *JAMA Surgery* 152 (1): 49–54.

Finkelstein, A., Y. Ji, N. Mahoney, and J. Skinner. 2018. "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." *Journal of the American Medical Association* 320 (9): 892–900.

Frandsen, B., M. Powell, and J. B. Rebitzer. 2019. "Sticking Points: Common-Agency Problems and Contracting in the US Healthcare System." *RAND Journal of Economics* 50 (2): 251–85.

Frank, R. G., and R. Z. Zeckhauser. 2007. "Custom-Made versus Ready-to-Wear Treatments: Behavioral Propensities in Physicians' Choices." *Journal of Health Economics* 26 (6): 1101–27.

Geruso, M., and M. R. Richards. 2021. "Trading Spaces: Medicare's Regulatory Spillovers on Treatment Setting for Non-Medicare Patients." NBER Working Paper No. 28576. https://doi.org/10.3386/w28576.

Glied, S., and K. Hong. 2018. "Health Care in a Multi-payer System: Spillovers of Health Care Service Demand among Adults under 65 on Utilization and Outcomes in Medicare." *Journal of Health Economics* 60:165–76.

Glied, S,. and J. G. Zivin. 2002. "How Do Doctors Behave When Some (But Not All) of Their Patients Are in Managed Care?" *Journal of Health Economics* 21 (2): 337–53.

Haas, D. A., X. Zhang, R. S. Kaplan, and Z. Song. 2019. "Evaluation of Economic and Clinical Outcomes under Centers for Medicare & Medicaid Services Mandatory Bundled Payments for Joint Replacements." *JAMA Internal Medicine* 179 (7): 924–31.

He, D., M. McInerney, and J. M. Mellor. 2015. "Physician Responses to Rising Local Unemployment Rates: Healthcare Provision to Medicare and Privately Insured Patients." *Journal of Health Economics* 40:97–108.

He, D., and J. M. Mellor. 2012. "Hospital Volume Responses to Medicare's Outpatient Prospective Payment System: Evidence from Florida." *Journal of Health Economics* 31 (5): 730–43.

Hughes Tuohy, C., C. M. Flood, and M. Stabile. 2004. "How Does Private Finance Affect Public Health Care Systems? Marshalling the Evidence from OECD Nations." *Journal of Health Politics, Policy and Law* 29 (3): 359–96.

Ibrahim, S. A., H. Kim, and K. J. McConnell. 2016. "The CMS Comprehensive Care Model and Racial Disparity in Joint Replacement." *Journal of the American Medical Association* 316 (12): 1258–59.

Iorio, R., A. J. Clair, I. A. Inneh, J. D. Slover, J. A. Bosco, and J. D. Zuckerman. 2016. "Early Results of Medicare's Bundled Payment Initiative for a 90-Day Total Joint Arthroplasty Episode of Care." *Journal of Arthroplasty* 31 (2): 343–50.

Joynt, K. E., D. Chan, E. J. Orav, and A. K. Jha. 2013. "Insurance Expansion in Massachusetts Did Not Reduce Access among Previously Insured Medicare Patients." *Health Affairs* 32 (3): 571–78.

Joynt, K. E., D. C. Chan, J. Zheng, E. J. Orav, and A. Jha. 2015. "The Impact of Massachusetts Health Care Reform on Access, Quality, and Costs of Care for the Already-Insured." *Health Services Research* 50 (2): 599–613.

Kim, H., J. I. Grunditz, T. H. A. Meath, A. R. Quiñones, S. A. Ibrahim, and K. J. McConnell. 2019. "Level of Reconciliation Payments by Safety-Net Hospitals under the First Year of the Comprehensive Care for Joint Replacement Program." *JAMA Surgery* 154 (2): 178–79.

Kim, H., T. H. A. Meath, J. I. Grunditz, A.R. Quiñones, S. A. Ibrahim, and K. J. McConnell. 2018. "Characteristics of Hospitals Exiting the Newly Voluntary Comprehensive Care for Joint Replacement Program." *JAMA Internal Medicine* 178 (12): 1715–17. https://doi.org/10.1001/jamainternmed.2018.4743.

Landon, B. E. 2017. "Tipping the Scale—The Norms Hypothesis and Primary Care Physician Behavior." *New England Journal of Medicine* 376 (9): 810–11.

Maniya, O. Z., R. C. Mather III, D. E. Attarian, B. Mistry, A. Chopra, M. Strickland, and K. Schulman. 2017. "Modeling the Potential Impact of the Medicare Comprehensive Care for Joint Replacement Episode-Based Payment Model." *Journal of Arthroplasty* 32 (11): 3268–73.

Markovitz, A. A., and A. M. Ryan. 2017. "Pay-for-Performance: Disappointing Results or Masked Heterogeneity?" *Medical Care Research and Review* 74 (1): 3–78.

McInerney, M., J. M. Mellor, and L. M. Sabik. 2017. "The Effects of State Medicaid Expansions for Working-Age Adults on Senior Medicare Beneficiaries." *American Economic Journal: Economic Policy* 9 (3): 408–38.

MedPAC. 2021. "Inpatient Rehabilitation Facilities Payment System." Accessed February 2022. https://www.medpac.gov/wp-content/uploads/2021/11/medpac_payment_basics_21_irf_final_sec.pdf.

Meyers, D. J., C. M. Kosar, M. Rahman, V. Mor, and A. N. Trivedi. 2019. "Association of Mandatory Bundled Payments for Joint Replacements with Use of Postacute Care among Medicare Advantage Enrollees." *JAMA Network Open* 2 (12): e1918535.

Mullen, K. J., R. G. Frank, and M. B. Rosenthal. 2010. "Can You Get What You Pay For? Pay-for-Performance and the Quality of Healthcare Providers." *RAND Journal of Economics* 41 (1): 64–91.

Navathe, A. S., J. M. Liao, S. E. Dykstra, E. Wang, Z. M. Lyon, Y. Shah, J. Martinez, et al. 2018. "Association of Hospital Participation in a Medicare Bundled Payment Program with Volume and Case Mix of Lower Extremity Joint Replacement Episodes." *Journal of the American Medical Association* 320 (9): 901–10.

Navathe, A. S., J. M. Liao, Y. Shah, Z. Lyon, P. Chatterjee, D. Polsky, and E. J. Emanuel. 2018. "Characteristics of Hospitals Earning Savings in the First Year of Mandatory

Bundled Payment for Hip and Knee Surgery." *Journal of the American Medical Association* 319 (9): 930–32.

Navathe, A. S., A. B. Troxel, J. M. Liao, N. Nan, J. Zhu, W. Zhong, and E. J. Emanuel. 2017. "Cost of Joint Replacement Using Bundled Payment Models." *JAMA Internal Medicine* 177 (2): 214–22.

Newhouse, J. P., and M. S. Marquis. 1978. "The Norms Hypothesis and the Demand for Medical Care." *Journal of Human Resources* 13 Suppl.:159–82.

Relman, A. S., and U. E. Reinhardt. 1986. "Debating For-Profit Health Care and the Ethics of Physicians." *Health Affairs* 5 (2): 5–31.

Richards, M. R., J. Seward, and C. Whaley. 2021. "Removing Medicare's Outpatient Ban and Medicare and Private Surgical Trends." *American Journal of Managed Care* 27 (3): 445–52.

Richards, M. R., and S. Tello-Trillo. 2019. "Public Spillovers from Private Insurance Contracting: Physician Responses to Managed Care." *American Economic Journal: Economic Policy* 11 (4): 375–403.

Rosenthal, M. B., and R. G. Frank. 2006. "What Is the Empirical Basis for Paying for Quality in Health Care?" *Medical Care Research and Review* 63 (2): 135–57.

Ryan, A. M. 2018. "Medicare Bundled Payment Programs for Joint Replacement: Anatomy of a Successful Payment Reform." *Journal of the American Medical Association* 320 (9): 877–79.

Siddiqi, A., P. B. White, J. B. Mistry, C. U. Gwam, J. Nace, M. A. Mont, and R. E. Delanois. 2017. "Effect of Bundled Payments and Health Care Reform as Alternative Payment Models in Total Joint Arthroplasty: A Clinical Review." *Journal of Arthroplasty* 32 (8): 2590–97.

Sloan, F. A., M. A. Morrisey, and J. Valvona. 1988. "Effects of the Medicare Prospective Payment System on Hospital Cost Containment: An Early Appraisal." *Milbank Quarterly* 66 (2): 191–220.

Sood, N., P. J. Huckfeldt, J. J. Escarce, D. C. Grabowski, and J. P. Newhouse. 2011. "Medicare's Bundled Payment Pilot for Acute and Postacute Care: Analysis and Recommendations on Where to Begin." *Health Affairs* 30 (9): 1708–17.

Sood, N., V. L. Shier, H. Nakata, R. Iorio, and J. R. Lieberman. 2019. "The Impact of Comprehensive Care for Joint Replacement Bundled Payment Program on Care Delivery." *Journal of Arthroplasty* 34 (4): 609–12.

Thirukumaran, C. P., L. G. Glance, X. Cai, R. Balkissoon, A. Mesfin, and Y. Li. 2019. "Performance of Safety-Net Hospitals in Year 1 of the Comprehensive Care for Joint Replacement Model." *Health Affairs* 38 (2): 190–96.

Thirukumaran, C. P., L. G. Glance, X. Cai, Y. Kim, and Y. Li. 2019. "Penalties and Rewards for Safety Net vs. Non–Safety Net Hospitals in the First 2 Years of the Comprehensive Care for Joint Replacement Model." *Journal of the American Medical Association* 321 (20): 2027–30.

Trish, E., P. Ginsburg, L. Gascue, and G. Joyce. 2017. "Physician Reimbursement in Medicare Advantage Compared with Traditional Medicare and Commercial Health Insurance." *JAMA Internal Medicine* 177 (9): 1287–95.

White, C. 2013. “Contrary to Cost-Shift Theory, Lower Medicare Hospital Payment Rates for Inpatient Care Lead to Lower Private Payment Rates.” *Health Affairs* 32 (5): 935–43.

———. 2014. “Cutting Medicare Hospital Prices Leads to a Spillover Reduction in Hospital Discharges for the Nonelderly.” *Health Services Research* 49 (5): 1578–95.

Wilcock, A. D., M. L. Barnett, J. M. McWilliams, D. C. Grabowski, and A. Mehrotra. 2020. “Association between Medicare’s Mandatory Joint Replacement Bundled Payment Program and Post-Acute Care Use in Medicare Advantage.” *JAMA Surgery* 155 (1): 82–84.

Yip, W. 1998. “Physician Response to Medicare Fee Reductions: Changes in the Volume of Coronary Artery Bypass Graft (CABG) Surgeries in the Medicare and Private Sectors.” *Journal of Health Economics* 17 (6): 675–99.